Perceptual Discriminability Drives Overinformative Reference, But Colour Information is Special

Merrick Giles[1†], Paula Rubio-Fernandez[2], Francis Mollica[1]
[1] Melbourne School of Psychological Sciences, University of Melbourne, Australia
[2] Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
[†] merrickdgiles@gmail.com

Abstract

When speakers refer to objects in the world, they often *overinform*: provide their listener with redundant adjectival information. Contrary to classical theories in linguistics, recent theories have framed overinformativeness as an efficient means of grounding reference in perceptual information of high discriminability to facilitate listener comprehension. However, the generalisability of such theories is constrained by the methodological challenge associated with reliably manipulating the perceptual discriminability of naturalistic stimuli. Here, we overcome these methodological challenges, using methods from psychophysics to manipulate the perceptual discriminability of colour and material attributes in a reference-production experiment. We provide a robust validation of the view that overinformative reference is driven by speakers grounding expressions in attributes of high discriminability. However, we also find that colour information is privileged above and beyond such factors of discriminability.

Keywords: reference production; overinformative reference; experimental pragmatics.

Introduction

Reference in the physical world is frequently overinformative: speakers provide their listeners with more information than required to isolate an object of reference (Pechmann, 1989; Sedivy, 2003). A speaker commonly says "pass me the blue cup" in a situation with just one cup: here, the modifier *blue* is redundant; they could have just said "the cup". The persistence of redundant, overinformative modification might at first seem inconsequential. However, it provides a striking counter example to classical notions of communicative efficiency (Zipf, 1949; Grice, 1975): efficient speakers say as much as necessary, but no more. In contrast, overinformative real world speakers seem generally blind to requirements of informativeness, expending unnecessary effort.

More recent theories of reference, however, have reassessed these notions of informational necessity and redundancy. They argue that an expression is communicatively efficient if it facilitates listener comprehension: this is the case even if such an expression exceeds the strictly necessary informational requirements of the exchange (Jara-Ettinger & Rubio-Fernandez, 2022; Kursat & Degen, 2021; Rehrig, Cullimore, Henderson, & Ferreira, 2021; Rubio-Fernandez, 2019; Tourtouri, Delogu, & Crocker, 2021). In the case of physical reference, a key factor in listener comprehension is the perceptual processing of a scene (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; McMurray, Clayards, Tanenhaus, & Aslin, 2008). The new theories therefore argue that expressions with redundant modifiers such as *blue* in "the blue cup" should be considered efficient if the informationally redundant modifier facilitates the listener's perceptual search for the object of reference. In turn, this so-called redundant modifier is likely to facilitate the referential search process when the modifier grounds reference in an easy-to-perceptually-discriminate object attribute. From the perspective of comprehension efficiency, then, overinformativeness might naturally arise as speakers seek to ground reference in highly discriminable object properties: redundancy is simply a worthwhile trade-off to achieve efficient comprehension. We will call this the *perceptual efficiency view*.

The reference production literature supports the assumptions of informational redundancy as perceptually efficient, operationalising perceptual efficiency as *perceptual distinctiveness*. While perceptual distinctiveness is itself implemented in varied ways, a range of experimental paradigms support the conclusion that speakers use redundant *colour* modifiers to ground reference in visually distinctive object attributes. For example, redundant colour modifiers increase as the density of the display increases (e.g., nine objects as opposed to two; Clarke, Elsner, & Rohde, 2013; Gatt, Krahmer, Van Deemter, & van Gompel, 2017; Rubio-Fernandez, 2019). These density effects are particularly pronounced when the target referent has a *pop-out* colour: whereby the colour of the referent differs dramatically from a uniform set. Additionally, rates of redundant colour modifiers are higher when referential environments are polychromatic as opposed to monochromatic—i.e., colour is more distinctive of the referent (Koolen, Goudbeek, & Krahmer, 2013; Rubio-Fernandez, 2016; Long, Moore, Mollica, & Rubio-Fernandez, 2021). The general visual search literature dovetail these findings as visual search for target objects is much faster in displays wherein the search-target is similarly distinctive in colour (Nagy & Sanchez, 1990; Wolfe & Horowitz, 2017). Finally, the comprehension literature supports the claim that overinformative expressions that convey perceptually useful information facilitate referential comprehension, whether measured by EEG (Tourtouri, Delogu, Sikos, & Crocker, 2019; Tourtouri et al., 2021), eye-tracking (Rehrig et al., 2021; Rubio-Fernandez, 2021), or response times (Paraboni, Van Deemter, & Masthoff, 2007; Rubio-Fernandez, 2021).

However, the generalisability of the perceptual efficiency view of reference remains constrained. The vast majority of experimental research supporting the view manipulates the perceptual distinctiveness of *colour* as the attribute of re-

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society* ©2025 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

dundancy ("the blue cup" in a display with just one cup). Likely due to the inherent challenge in finding naturalistic attributes with no colour association, no research has manipulated perceptual discriminability in psychophysical terms using alternative, non-relative object attributes such as material constitution. Operationalised in these psychophysical terms, perceptual discriminability corresponds to the ease or difficulty of perceptual search. This operationalisation corresponds closely to the perceptual efficiency view, while avoiding the vague and varied implementations of *perceptual distinctiveness*. Nevertheless, such manipulations have to date only been invoked using colour-redundant attributes (Rubio-Fernandez, 2021).

The lack of research into the effect of discriminability for alternative attributes presents a further difficulty for the perceptual efficiency view: a ubiquitous experimental finding is that speakers redundantly mention an objects' colour far more often than alternative attributes such as size or material constitution (Degen, Hawkins, Graf, Kreiss, & Goodman, 2020; Jara-Ettinger & Rubio-Fernandez, 2022; Kursat & Degen, 2021; Pechmann, 1989; Sedivy, 2003, 2004; Van Gompel, Van Deemter, Gatt, Snoeren, & Krahmer, 2019; Viethen, van Vessem, Goudbeek, & Krahmer, 2016; Koolen, Gatt, Goudbeek, & Krahmer, 2011). Proponents of the perceptual efficiency view argue that this *colour asymmetry* must be attributed to colours high perceptual distinctiveness relative to alternative perceptual attributes such as material (Rubio-Fernandez, 2021; Koolen et al., 2013; Van Gompel et al., 2019; Kursat & Degen, 2021). However, in principle it is unlikely that colour information is more perceptually distinctive than alternative perceptual information sources in every environment; yet no work has investigated whether the colour-asymmetry is reversed when an alternative attribute (such as material) is made more or less perceptually discriminable than colour in a referential environment. These limitations suggest that the hypothesised perceptual drivers of reference production may only apply to reference about *colour*, rather than being a domain-general principle.

The Present Study

Manipulating the perceptual discriminability of non-relative naturalistic attributes presents a methodological challenge, likely explaining the lack of psychophysical manipulation of non-colour attributes in prior studies. While researchers can exploit the continuous nature of colour space to manipulate its perceptual discriminability (Viethen et al., 2016), there is no analogous continuum for, e.g., material constitution, through which discriminability can be psychophysically manipulated.

We introduce several experimental innovations to meet this challenge. First, we manipulate the perceptual discriminability of material constitution through the auditory modality; this enables us to create a continuous auditory perceptual space between two distinct material impact sounds. Further, we adopt established experimental techniques from psychophysics to calibrate the perceptual discriminability of colour (visual) and material (auditory) for each individual participant (Leek, 2001). This use of auditory stimuli requires participants to describe referential scenes from working memory as auditory components of the stimuli cannot be presented simultaneously: it must be presented sequentially, with each sound disappearing before the following sound occurs. To ensure equal memory load for both colour and material, we also presented the visual aspects of the stimuli in precisely the same manner. This differs from prior research that has largely made use of *static* director-tasks, wherein all objects are static and perceptually available during description (e.g., Degen et al., 2020; Koolen et al., 2013; Kursat & Degen, 2021; Rubio-Fernandez, 2021).

While employed to address the challenge of creating viable naturalistic stimuli, describing properties inferred from auditory information using working memory has several advantages. First, while visually discriminating material constitution is difficult (as indicated by search times, (Kursat & Degen, 2021; Jara-Ettinger & Rubio-Fernandez, 2022)), humans can easily discriminate material constitution from impact sounds (Fujisaki, Goda, Motoyoshi, Komatsu, & Nishida, 2014). For example, listeners can almost perfectly discriminate between material categories such as wood and metal on the basis of an impact sound (Giordano & McAdams, 2006; Traer, Cusimano, & McDermott, 2019). Second, probing referential patterns for properties inferred from audition allows us to generalise the perceptual efficiency view across modalities, capturing many instances of real-world reference. Finally, we achieve even greater generalisability given the use of working memory, as reference to objects from working memory captures many realistic referential contexts (e.g., "that red bike that just went past"). Thus, we provide the strongest validation of the perceptual efficiency view in explaining the richness of reference processing, extending across modalities and beyond static environments.

Using this audio-visual context and description from working memory, we operationalise and manipulate the perceptual discriminability of colour (visual) and material (auditory). In a reference-production task, speakers describe objects (baseball bats), each with a colour (inferred from visual properties), and a material (inferred from auditory properties). We manipulate two components relevant to the hypothesis that reference is driven by perceptual discriminability.

First, we manipulate the discriminability of attributes; this manipulation is split into three display type conditions: 1) Baseline, where the sufficient attribute[1] *and* the redundant attribute are both of high discriminability; 2) Sufficient-High/Redundant-Low (S-High/R-Low), where the sufficient attribute is of high-discriminability, and the redundant attribute is of low-discriminability; 3) Sufficient-Low/Redundant-High (S-Low/R-High), where the sufficient attribute is of low-discriminability, and the redundant at-

[1] We call an attribute 'sufficient' if the attribute alone provides enough information to successfully demarcate an object: e.g., "the green bat" in a scene with two blue bats and *one* green bat.

tribute is of high-discriminability. The perceptual efficiency view predicts that speakers will overinform most often in S-Low/R-high displays, as overinformativeness is required to ground reference in a high-discriminability attribute in this condition.

Second, we manipulate whether colour or material is the redundant attribute. The perceptual efficiency view posits that the colour-asymmetry in reference is due to colours high perceptual discriminability in a scene. As we equate the perceptual discriminability of material and colour stimuli using psychophysical methods, the view predicts that rates of overinformativity between colour and material will equalise.

Method

All participants completed experiments online. The experiment was programmed using custom Javascript code, using psiTurk (Gureckis et al., 2016) for back-end architecture. To ensure a reliable manipulation of perceptual discriminability calibrated to each participant, we derive a personalised set of high- and low-discriminability stimuli for colour (visual) and material constitution (auditory) by implementing an adaptive perceptual staircase task similar to that described in Leek (2001, see left panels Figure 1). With these derived sets of high- and low-discriminability colours and materials, we presented participants with a gamified scenario: They are working quality control at a baseball bat factory, and their job is to describe the selected bat to a co-worker, using its colour (encoded visually) and/or material (encoded auditorily). In this adaptation of the classic director task (Krauss & Glucksberg, 1977), three bats are animated to fall sequentially from left to right (Figure 1: right panel). Combining the material sound with the animation creates the impression of an auditory impact occurring due to the bats falling on a surface. After displaying all bat animations and sounds, the target is clearly demarcated with an "X" (Figure 1: right). Participants provided written responses to refer to the bat associated with the target, using either a.) colour, b.) material, or c.) both (overinformative). The whole experiment took each participants approximately 40 minutes.

Partcipants

As we planned to use complex psychophysics task online, we conducted a few pilot studies to check that the staircase task was able to determine adequate high- and low-discriminability stimuli for the director task. We expected this might be difficult as we have less control to adapt to differences in people's perceptual ability and less control over presentation equipment than a classic lab study. Fortunately, we were able to find adequate stimuli for $\sim 75\%$ of pilot participants. Our power analysis indicated we should target a sample of 80 participants in the director task. Expecting to exclude participants from the director task for lack of adequate stimuli, we recruited 120 native English speakers with no colour-vision or hearing impairments through Prolific. Consistent with ethical approval from the University of Melbourne. Participants were compensated at a rate of £9.55 per hour.

Materials

We synthesised material-derived impact sounds using the generative model established by Traer et al. (2019) for four material sounds: metal, wood, glass, cardboard. Crucially, Traer et al. demonstrated that listeners cannot distinguish these synthetic material sounds from high-fidelity, real-world recordings of the same material. Participants were assigned to one of two conditions: half to a metal-cardboard condition, and half to a wood-glass condition. From the participant's perspective, materials were always labelled as "wood" (wood and cardboard) or "metal" (metal and glass). The fact that glass sounds were mislabelled as metal and that cardboard sounds were mislabelled as wood was unlikely to be noticed by participants, as humans struggle to distinguish between metal and glass, and between wood and cardboard (Traer et al., 2019). To create stimuli for the psychophysical staircases, we employed a linear interpolation of 10 equally spaced steps between the modal parameters (amplitude, centre frequency, decay rate) of the two sounds (wood-to-glass, cardboard-to-metal). The result was a perceptually smooth transition between each material. The volume of all sounds was normalised to ensure consistency.

For the colour stimuli, we derived four co-ordinates from the CIELAB colour space (blue, green, orange, yellow). Participants were assigned to either the blue-green condition, or the orange-yellow condition, allowing us to generalise across warm and cool poles in colour space. To derive psychophysical staircases for colour, we employed a linear interpolation between co-ordinates of the CIELAB space. The CIELAB space was used for its approximate perceptual uniformity (Brainard et al., 2003). As colour is of a far more gradient nature than material, finding a consistent low-discriminability region of colour space required continuums of a greater number of steps: we employed 50 equally spaced steps for yellow-to-orange, and for blue-to-green. The result was a perceptually smooth transition between the two colours (Figure 2).

Procedure

Participants were randomly assigned to one of four conditions, corresponding to a colour-material combinatorial pair (wood-metal/blue-green, cardboard-glass/blue-green, wood-metal/orange-yellow, cardboard-glass/orange-yellow). First, participants proceeded through the adaptive perceptual staircases to identify perceptual regions of high and low perceptual discriminability for both colour and material. For each attribute, participants completed the two staircases simultaneously–at opposing ends–with interleaved trials. For example, participants began one staircase beginning at wood and ending at metal, and one beginning at metal and ending at wood simultaneously (as labelled in their respective conditions).

Each staircase began with a high-discriminability stimulus. When participants correctly label the material source of

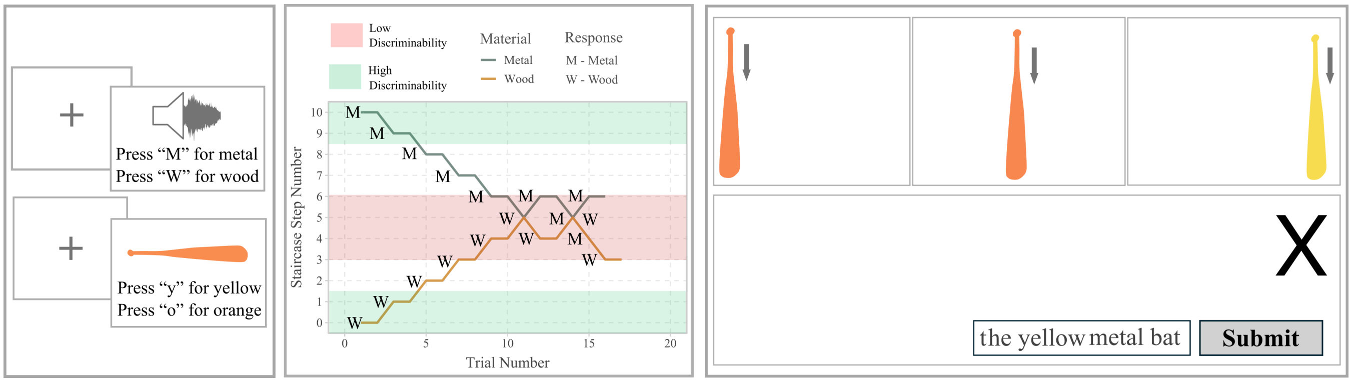

Figure 1: Illustration of Experimental Procedure. Left: participant view showing staircases of material (auditory) and colour (visual). Participants classify sounds as wood or metal, and colours as yellow or orange/blue or green. Middle: example of a single staircase for one participant. Right: participants view of a single example trial during the director task. Participants observe three bats, all of which fall from the top of the screen to the bottom, where they hit an invisible surface. The moment the bats hit the surface, participants hear the auditory impact sound of wood or metal, of high or low perceptual discriminability (depending on the trial condition). In this trial, colour is highly discriminable, and *Sufficient* to establish reference. Supposing that material (Redundant) is high-discriminability, the presented trial would be a Baseline condition. If material was low-discriminability, the trial would be a Sufficient-High/Redundant-Low trial. Note that contrary to the image shown, the bats do *not* remain on the screen. Each individual bat disappears before the following bat appears, presented sequentially left-to-right.

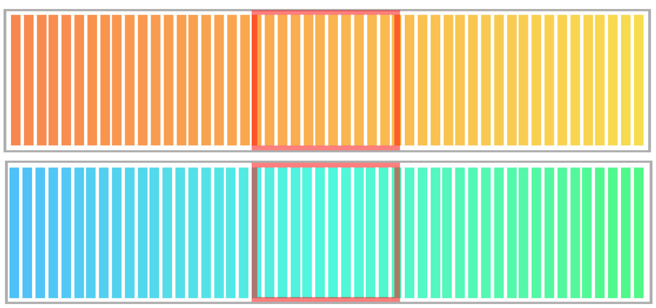

Figure 2: The Two Colour Continuums Used. Red square outline denotes illustrative examples of low-discriminability colour regions: in the actual experiment, these are calibrated to each participants individual region.

the sound, they progress in the staircase (towards the other end). Note that our task is unique in that there is no inherently “correct” identification, but a label is condidered correct if is it the label associated with the starting stimulus. E.g., if the staircase starts with metal, and a participant identifies a stimulus as metal, this is a correct identification; if they identify a stimulus as wood in the same staircase, this is an incorrect identification.

With a correct material identification, participants progressed one step; however, with a correct colour identification, participants progressed five steps (due to the increased number of steps in the colour continuum), but only until their first incorrect identification. After their first incorrect identification, they progress four steps; after their second incorrect identification, two steps; after their third, they move up and down the staircase one step at a time. When participants incorrectly label a stimulus, this is called a reversal. For example, a reversal occurs when a participant identified the previous step as wood, but then identifies the current step as metal. After five reversals, participants must correctly label the stimulus once more (likely when the immediately preceding response was incorrect). This stimulus of the final correct identification was defined as the stimulus of low-discriminability. In contrast, a stimulus was defined as high-discriminability if it met three conditions: 1.) The participant had a reversal count of 0 in the staircase thus far; 2.) The stimulus was “correctly” identified as its source colour/material; 3.) the current step in the continuum was less than or equal to 2.

If the staircase identified high- and low-discriminability exemplars of material and colour, participants continued the experiment and the exemplars were used in the director task. In this task, there were 48 critical trials (24 colour-redundant, 24 material-redundant), which were split into 12 Baseline trials, 12 Sufficient-High/Redundant-Low (S-High/R-Low) trials, 12 Sufficient-Low/Redundant-High (S-Low/R-high) trials, and 12 fillers. Participants described baseball bats with an associated colour and material to a hypothetical listener (Krauss & Glucksberg, 1977, see the *Director Task*). Prior to the experiment, participants were introduced to an image of their hypothetical listener (named Ralph), and were informed that Ralph was equally adept at finding bats using material or using colour.

In each trial, three bats fell from the top of the screen in a sequence from left to right (Figure 1: right panel). Each bat fell for 800 ms before hitting the invisible surface on the bottom of the screen, where they produced the material impact sound (1 s in duration). They remained on the screen (visually) for an additional 1.7 s before disappearing as the following bat began to fall. Once all bats were presented, an “X” appeared at the target’s location, and the participants provided their descriptions as written responses.

We employed moderate feedback during the reference-production phase. If participants failed to uniquely demarcate the referent from the surrounding objects, a thumbs down ap-

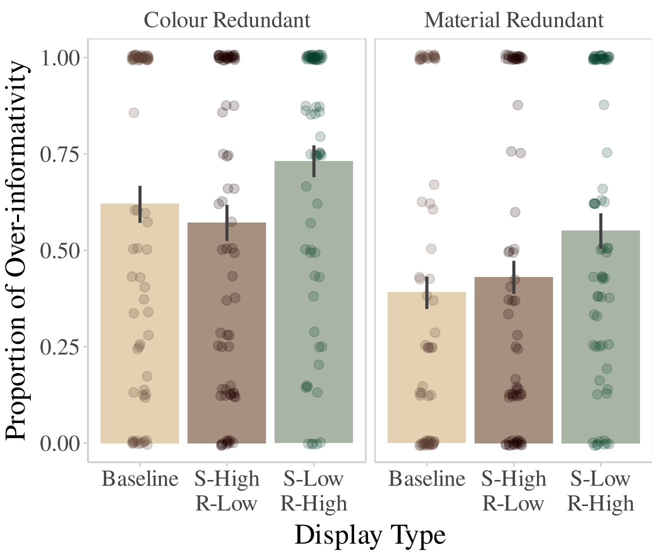

Figure 3: Dots represent individual participant means. Lines represent bootstrapped 95% confidence intervals.

peared on the screen. If participants succeeded in unique demarcation, a thumbs up appeared (regardless of whether the expression was overinformative or not).

Results

From our initial sample of 120 participants, 43 participants did not have adequate stimuli for the director tasks from the staircases, leaving us with 77 participants. We excluded an additional five participants for achieving correct reference on less than 80% of the trials in the director task (which usually has near perfect accuracy)[2]. In line with prior research, we restricted analysis to trials with identifiable references. The excluded non-identifiable references were not systematic across conditions ($<$ 5% of the individual participants' trials), and were therefore attributed to brief lapses of attention and only those individual trials were removed (201 trials total). Additionally, 1 random trial from each participant failed to record due to software error.

The remaining data from critical (non-filler) trials (72 participants, 2785 trials) were analysed using R (R Core Team, 2023, Version 4.3.2). We employed Bayesian Mixed Effects Logistic Regressions (main effects and by-subject intercepts) using `brms` (Bürkner, 2017) with $\mathcal{N}(0,2)$ priors over regression coefficients, chosen using prior-predictive checks. Each analysis used dummy-coded contrasts (references: Material-Redundant, S-Low/R-High).

To recap, the discriminability account predicts that 1) rates of overinformative reference will be higher for S-Low/R-High displays than S-High/R-Low displays and Baseline displays. Assuming discriminability is the only driver of overinformativeness, the account also predicts that 2) rates of overinformativity between attributes (Colour-Redundant vs Material-Redundant) should be no different.

[2]Some responses were also indicative of fatigue effects (e.g., giving up with the experiment and typing *qqq* as the response to all remaining trials). We excluded all trials from such participants.

Effect	β	SE	95% CI	
			LL	UL
Intercept	-1.10	0.40	-1.90	-0.33
Colour Redundant	1.41	0.11	1.19	1.64
Baseline	-0.97	0.13	-1.24	-0.71
S-High/R-Low	-1.13	0.13	-1.39	-0.87

Table 1: Regression Results. SE = Standard Error; 95% CI = Bayesian 95% Credible Intervals.

As can be seen in Figure 3, participants were more likely to overinform in the S-Low/R-High condition compared to baseline (both attributes high-discriminability) and S-High/R-Low. Simply put, participants overinformed more often when doing so was required to ground reference in a high-discriminability attribute. Precisely supporting the perceptual efficiency view, the statistical results are presented in Table 1.

Contrary to the perceptual discriminability hypothesis, rates of overinformativeness were higher when colour, rather than material, was the redundant attribute (Table 1). As the discriminability of redundant colour and material information were controlled using the staircase, the perceptual discriminability hypothesis cannot account for this result.

Discussion

The present study provides a strong empirical test of the perceptual efficiency view (Jara-Ettinger & Rubio-Fernandez, 2022; Kursat & Degen, 2021; Rubio-Fernandez, 2019), addressing key constraints on its generalisability as the driver of overinformativeness. In accordance with the hypothesis, we found that participants overinform more often when the expression of redundant information is required to anchor reference in high discriminability object attributes. Notably, the observed patterns of overinformativeness cannot be attributed to a global reduction in the perceptual discriminability of an object: Speakers did not overinform more often when redundancy would ground reference in a *low-discriminability* attribute. Rather, speakers overinformed more often when the redundancy would ground reference in a *high-discriminability* attribute. Moreover, speakers overinformed more often when the uniquely demarcating attribute was of low discriminability as opposed to high discriminability. In these contexts, grounding reference in an alternative, high-discriminability attribute is hypothesized to facilitate listener comprehension. Contrastingly, in cases where there is a uniquely demarcating high discriminability attribute, there is a lesser need for redundancy, as speakers can ground reference in high-discriminability attributes without the use of overinformativeness. Thus, the results follow the patterns of speakers oriented toward the perceptual facilitation of the listeners search. Additionally, the results demonstrate that perceptual discriminability is an attribute-general driver of reference production, applicable across attributes and modalities. Our results therefore provide more evidence against classical notions of reference production which characterise overinformativeness as inefficient (Grice, 1975; Zipf, 1949).

At a fine-grained level, the influence of perceptual discriminability could be attributed to varied general communicative strategies: first, speakers could follow a general strategy of mentioning all high-discriminability attributes in a scene, largely ignoring whether attributes are sufficient or redundant. The present study shows that this is unlikely: if this were the case, we would observe comparable rates of overinformativeness in both Baseline conditions (all attributes high-discriminability) and Sufficient-Low/Redundant-High (S-Low/R-High) conditions. In both cases, mentioning *all* high-discriminability attributes results in overinformativeness. In contrast, speakers were more likely to overinform in S-Low/R-High, where the discriminability of the sufficient attribute was low. This could suggest two alternative strategies: a.) speakers may follow the strategy of mentioning *at least one* high-discriminability property or b.) speakers may buffer their expressions with high-discriminability redundant attributes when the uniquely demarcating property is unreliable, carrying a high risk of communicative failure. Disentangling these alternatives presents a promising avenue for future work.

Another finding is that speakers overinform more often using redundant colour attributes, as opposed to redundant material attributes. In recent years, colours' asymmetric use in reference has been attributed to its high perceptual distinctiveness in comparison to alternative attributes such as material constitution, pattern, and shape (Jara-Ettinger & Rubio-Fernandez, 2022; Koolen et al., 2013; Kursat & Degen, 2021; Long et al., 2021; Rubio-Fernandez, 2019; Van Gompel et al., 2019). Operationalising distinctiveness in psychophysical terms, the present findings suggest that perceptual distinctiveness alone cannot fully capture this effect. Naturally, this raises the question of why colour should have such a central role in physical reference. We offer two speculations.

First, the distinctiveness of environmental attributes may have influence on perceptual reference beyond the dynamics of an immediate scene. Colour terms are highly optimised for *perceptual* communication, maximising the perceptual distance between and minimising the perceptual distance within basic colour categories (Abbott, Griffiths, & Regier, 2016; Regier, Paul, & Khetarpal, 2007). While functional optimality of this nature may be a general principle of communication across varied semantic domains (Kemp & Regier, 2012, e.g.,), common terms for material categories such as "wood" and "metal" may be suboptimal with respect to perceptual search—this assertion is supported by the comparative superiority of colour over material evidenced by visual search times (Jara-Ettinger & Rubio-Fernandez, 2022; Kursat & Degen, 2021). Learning that colour information is highly useful across many varied perceptual contexts, speakers could come to rely on colour information as a general strategy of reference, giving colour a privileged role in perceptual modification. Therein, the disproportionate use of colour may reflect prior successes in perceptual communication.

Alternatively, colour may simply be more central to the communicative patterns and referential goals of English speakers. The use of colour in language is known to correlate with how much colour-variation within object classes occurs in the environment (E. Gibson et al., 2017), as well as colours' perceptual prominence in the environment (Wnuk, Verkerk, Levinson, & Majid, 2022). Given these facts, it is notable that while all objects have an associated colour, and indeed often have many colours, the relevance of material constitution is largely restricted to specific artifacts. This is likely to make colour more generally useful to the average speaker. This general utility may spill-over into referential patterns when speakers are made to refer in a context such as the present experiment, where colour and material are of equal utility and relevance. It should be noted, however, that the more pronounced relevance of colour in the present experiment might instead be attributed to speakers finding *vision* more relevant to communication than *audition*: speakers could generally consider an objects visual properties more relevant to their communicative goals than they do auditory ones. While we cannot dissociate the possibility that modality (audio–vision) as opposed to attribute (colour–material) is driving the comparative advantage of colour in the present study, this is currently being investigated by the authors in follow-up work, wherein high-discriminability colour is being compared to alternative, highly-discriminable visual properties that are known to guide visual attention in bottom-up processes (Wolfe & Horowitz, 2017).

The two core takeaways of this study are 1.) the colour asymmetry persists beyond discriminability manipulation, and 2.) discriminability is an attribute-general driving factor of referential expressions. Following recent accounts, we have shown that overinformativeness in reference can arise from communicative strategies that facilitate listener comprehension (though they need not be explicitly listener-oriented; see Wu & Gibson, 2021): this is consistent with recent theoretical advances positing that communication is optimised for its functional utility (Ferreira, 2019; Gibson et al., 2019). A further challenge lies in explaining the mechanisms through which speakers learn to deploy efficient communicative strategies, and to what extent these strategies are implicit or deliberative.

References

Abbott, J. T., Griffiths, T. L., & Regier, T. (2016). Focal colors across languages are representative members of color categories. *Proceedings of the National Academy of Sciences*, *113*(40), 11178–11183. doi: 10.1073/pnas.1513298113

Brainard, D. H., et al. (2003). Color appearance and color difference specification. *The science of color*, *2*(191-216), 5. doi: 10.1016/B978-044451251-2/50006-4

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Soft-*

ware, *80*(1), 1–28. doi: 10.18637/jss.v080.i01

Clarke, A. D., Elsner, M., & Rohde, H. (2013). Where's wally: The influence of visual salience on referring expression generation. *Frontiers in psychology*, *4*, 329. doi: 10.3389/fpsyg.2013.00329

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. doi: 10.1037/rev0000186

Ferreira, V. S. (2019). A mechanistic framework for explaining audience design in language production. *Annual review of psychology*, *70*, 29–51. doi: 10.1146/annurev-psych-122216-011653

Fujisaki, W., Goda, N., Motoyoshi, I., Komatsu, H., & Nishida, S. (2014). Audiovisual integration in the human perception of materials. *Journal of Vision*, *14*(4), 12-12. doi: 10.1167/14.4.12

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.

Gibson, Futrell, Piantadosi, Dautriche, Mahowald, Bergen, & Levy. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*(5), 389-407. doi: 10.1073/pnas.1619666114

Gibson, E., Futrell, R., Jara-Ettinger, J., Mahowald, K., Bergen, L., Ratnasingam, S., … Conway, B. R. (2017). Color naming across languages reflects color use. *Proceedings of the National Academy of Sciences*, *114*(40), 10785-10790.

Giordano, B. L., & McAdams, S. (2006, 02). Material identification of real impact sounds: Effects of size variation in steel, glass, wood, and plexiglass plates. *The Journal of the Acoustical Society of America*, *119*(2), 1171-1181. doi: 10.1121/1.2149839

Grice, H. P. (1975). Logic and conversation. *Syntax and semantics*, *3*, 43–58.

Gureckis, T. M., Martin, J., McDonnell, J., Rich, A. S., Markant, D., Coenen, A., … Chan, P. (2016). psiturk: An open-source framework for conducting replicable behavioral experiments online. *Behavior research methods*, *48*, 829–842.

Jara-Ettinger, J., & Rubio-Fernandez, P. (2022). The social basis of referential communication: Speakers construct physical reference based on listeners' expected visual search. *Psychological review*, *129*(6), 1394. doi: doi.org/10.1037/rev0000345

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*(6084), 1049-1054. doi: 10.1126/science.1218811

Koolen, R., Gatt, A., Goudbeek, M., & Krahmer, E. (2011). Factors causing overspecification in definite descriptions. *Journal of Pragmatics*, *43*(13), 3231–3250. doi: 10.1016/j.pragma.2011.06.008

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. doi: 10.1111/cogs.12019

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American*, *236*(2), 100–105.

Kursat, L., & Degen, J. (2021). Perceptual difficulty differences predict asymmetry in redundant modification with color and material adjectives. *Proceedings of the Linguistic Society of America*. doi: 10.3765/plsa.v6i1.5003

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & psychophysics*, *63*(8), 1279–1292. doi: 10.3758/BF03194543

Long, M., Moore, I., Mollica, F., & Rubio-Fernandez, P. (2021). Contrast perception as a visual heuristic in the formulation of referential expressions. *Cognition*, *217*, 104879. doi: 10.1016/j.cognition.2021.104879

McMurray, B., Clayards, M. A., Tanenhaus, M. K., & Aslin, R. N. (2008). Tracking the time course of phonetic cue integration during spoken word recognition. *Psychonomic bulletin & review*, *15*(6), 1064–1071. doi: 10.3758/PBR.15.6.1064

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *JOSA A*, *7*(7), 1209–1217.

Paraboni, I., Van Deemter, K., & Masthoff, J. (2007). Generating referring expressions: Making referents easy to identify. *Computational linguistics*, *33*(2), 229–254. doi: 10.1162/coli.2007.33.2.229

Pechmann, T. (1989). Incremental speech production and referential overspecification. *Linguistics*. doi: 10.1515/ling.1989.27.1.89

R Core Team. (2023). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Regier, Paul, & Khetarpal. (2007). Color naming reflects optimal partitions of color space. *Proceedings of the National Academy of Sciences*, *104*(4), 1436–1441. doi: 10.1073/pnas.0610341104

Rehrig, G., Cullimore, R. A., Henderson, J. M., & Ferreira, F. (2021). When more is more: Redundant modifiers can facilitate visual search. *Cognitive Research: Principles and Implications*, *6*, 1–20. doi: 10.1186/s41235-021-00275-4

Rubio-Fernandez, P. (2016). How redundant are redundant color adjectives? an efficiency-based analysis of color overspecification. *Frontiers in psychology*, *7*, 153. doi: 10.3389/fpsyg.2016.00153

Rubio-Fernandez, P. (2019). Overinformative speakers are cooperative: Revisiting the gricean maxim of quantity. *Cognitive science*, *43*(11), e12797. doi: 10.1111/cogs.12797

Rubio-Fernandez, P. (2021). Color discriminability makes over-specification efficient: Theoretical analysis and empirical evidence. *Humanities and Social Sciences Communications*, *8*(1), 1–15. doi: 10.1057/s41599-021-00818-6

Sedivy. (2003). Pragmatic versus form-based accounts of referential contrast: Evidence for effects of informativity expectations. *Journal of Psycholinguistic Research*, *32*, 3–23. doi: 10.1023/A:1021928914454

Sedivy. (2004). Evaluating explanations for referential context effects: Evidence for gricean meachanisms in online language interpretation. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (p. 345 - 362). Cambridge, MA: MIT Press.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, *268*(5217), 1632-1634. doi: 10.1126/science.7777863

Tourtouri, E. N., Delogu, F., & Crocker, M. W. (2021). Rational redundancy in referring expressions: Evidence from event-related potentials. *Cognitive Science*, *45*(12), e13071. doi: 10.1111/cogs.13071

Tourtouri, E. N., Delogu, F., Sikos, L., & Crocker, M. W. (2019). Rational over-specification in visually-situated comprehension and production. *Journal of Cultural Cognitive Science*, *3*(2), 175–202. doi: 10.1007/s41809-019-00032-6

Traer, J., Cusimano, M., & McDermott, J. H. (2019). A perceptually inspired generative model of rigid-body contact sounds. *Digital Audio Effects (DAFx)*, *1*(2), 3.

Van Gompel, R. P., Van Deemter, K., Gatt, A., Snoeren, R., & Krahmer, E. J. (2019). Conceptualization in reference production: Probabilistic modeling and experimental testing. *Psychological review*, *126*(3), 345. doi: 10.1037/rev0000138

Viethen, J., van Vessem, T., Goudbeek, M., & Krahmer, E. (2016). Color in reference production: the role of color similarity and color codability. *Cognitive Science*, *41*, 1493–1514. doi: 10.1111/cogs.12387

Wnuk, E., Verkerk, A., Levinson, S. C., & Majid, A. (2022). Color technology is not necessary for rich and efficient color language. *Cognition*, *229*, 105223.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature human behaviour*, *1*(3), 0058. doi: 10.1038/s41562-017-0058

Wu, S. A., & Gibson, E. (2021). Word order predicts cross-linguistic differences in the production of redundant color and number modifiers. *Cognitive Science*, *45*(1), e12934. doi: 10.1111/cogs.12934

Zipf, G. K. (1949). *Human behaviour and the principle of least effort*. Addison-Wesley.