Gaze and Pupil Indices of Rational Reference Production

Speakers often encode more information than necessary for specifying referents in the immediate visual context [1], defying Grice's Quantity [2]. Why speakers engage in such (seemingly) irrational behaviour has been the subject of much debate: Some studies suggest that production choices are mainly motivated by a concern to ease planning effort (Egocentric view [3-4]), while others support the view that speakers aim at producing utterances that are efficient given the conditions (Audience-design view [5-8]). Previous work has largely considered the influence of redundancy on listeners' comprehension as an index of production strategy, and linked audience design to increased cognitive effort for the speaker. In this study, we use eye-tracking to directly assess production strategies and cognitive effort. In a referential communication experiment, we examine whether speakers' gaze patterns before speech onset differ based on their production strategy, and whether audience design is associated with increased cognitive effort (measured as pupil dilations).

In a 2x3 within-participants design, we manipulate which adjective is *necessary* for specifying the target referent (colour or pattern), and which adjective *reduces* the uncertainty about the target referent (*referential entropy*) to a greater degree (colour, pattern, equal) (see [8]). We employ 36 experimental items comprising 6 visual displays each (one per condition). In each display, 6 objects are arranged in different configurations based on the condition (Fig.1), while the target object remains constant. In the fillers (N=108), either two or no adjectives are necessary. We control for the perceptibility of the distinguishing feature for each target object (see [9]), based on ratings obtained in an online norming study.

The procedure is as follows: Two participants, randomly assigned the roles of Speaker and Listener, engage in a referential communication game, while the Speaker's eye movements are tracked. Participants' task is to identify whether the objects are arranged in the same configuration on both their screens. The Speaker is instructed to ask questions about the horizontal position of an object marked as the target only on their screen (e.g., 'Is the blue ball on the left?' in Fig. 1a). The Listener has to respond with a 'yes' or 'no' by pressing a key in a keyboard in front of them. In half of the trials, the Listeners see a mirror version of the Speaker display. We plan to collect data from 48 pairs. The experiment is conducted in German.

Participants will be grouped based on their use of redundant adjectives (see [8]). We will analyse the proportions of referentially redundant utterances (binary coded, with GLMMs), and participants' pupil dilations in two time-windows: before and after the reveal of the target referent. We will also use LMMs to analyse proportions of inspections to the referents before and after the reveal of the target object, and log-gaze probability ratios for fixations to the target (e.g., the blue ball in Fig.1a) vs. the contrast (e.g., the green ball in Fig.1a) or competitor (e.g., the blue mitt in Fig.1a) objects in the interval between the reveal of the target referent and speech onset. All models will include 'Necessary adjective' and 'Entropy-reducing adjective' as fixed effects, the perceptibility score as a control factor, and the maximal random effects structure [10]. Models analysing eye-tracking measures will additionally include 'Speaker Group' as a between-subjects factor. We generally expect that, compared to speakers using an audience-design strategy, egocentric speakers will scan the visual scene less broadly before the reveal of the target referent and expend less cognitive effort.

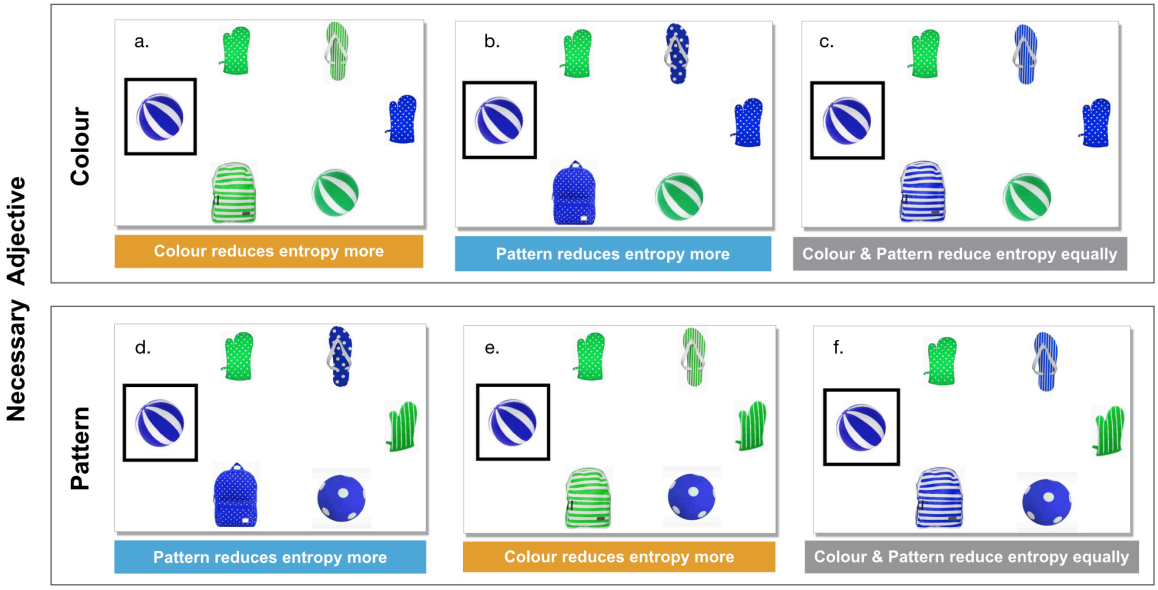

Figure1. Sample visual displays per condition. In the top panels, *colour* is necessary for specifying the target referent; in the bottom panels; *pattern* was the necessary feature. In the panels with the yellow label (a and e), *colour* is the more entropy reducing adjective, while in the panels with the blue label (b and d), *pattern* is the more entropy reducing adjective. In the panels with the grey label (c and f), both adjectives are equally entropy-reducing. The black frame indicates the target object and appears only on the Speaker's screen 2s picture onset. Figure is adapted from [8].

References

[1] Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: Speech acts (Vol. III, pp. 41–58*). New York: Academic.

[2] Davies C. & J. E. Arnold (2019). Reference and informativeness. In C. Cummins and N. Katsos (Eds.), *The Oxford Handbook of Experimental Pragmatics and Semantics*, pp. 29–88. Oxford, UK: Oxford University Press.

[3] Keysar, B., D. J. Barr & W. S. Horton (1998). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science*, 7(2), 46–49.

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

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

[6] Fukumura, K. & M. N. Carminati (2022). Overspecification and incremental referential processing: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 48(5), 680–701.

[7] Rubio-Fernandez, P., F. Mollica & J. Jara-Ettinger (2021). Speakers and listeners exploit word order for communicative efficiency: A cross-linguistic investigation. *Journal of Experimental Psychology: General*, 150(3), 583–594.

[8] Tourtouri, E. N., F. Delogu, L. Sikos & M. W. Crocker (2019). Rational over-specification in visually-situated comprehension and production. *Journal of Cultural Cognitive Science,* 3(2), 175–202.

[9] Kursat, L. & J. Degen (2021). Perceptual Difficulty Differences Predict Asymmetry in Redundant Modification With Color and Material Adjectives. *Proceedings of the Linguistic Society of America,* 6(1), 676–688.

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