UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
How grammatical gender supports efficient communication

Permalink
https://escholarship.org/uc/item/884148cc

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 47(0)

Authors
Hoppe, DorothŽe B.
Gibson, Edward
van Rij, Jacolien
et al.

Publication Date
2025

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

How grammatical gender supports efficient communication

Dorothée B. Hoppe (dorotheebdothoppe@gmail.com)

Center for Language and Cognition Groningen, University of Groningen

Edward Gibson (egibson@mit.edu)

Department of Brain & Cognitive Sciences, Massachusetts Institute of Technology

Jacolien van Rij (j.c.van.rij@rug.nll)

Department of Artificial Intelligence, University of Groningen

Petra Hendricks (p.hendriks@rug.nl)

Center for Language and Cognition Groningen, University of Groningen

Michael Ramscar (michael.ramscar@uni-tuebingen.de)

Department of Psychology & Cognitive Science Center, University of Tübingen

Abstract

The apparent redundancy of grammatical gender systems presents a puzzle to which information theory offers a solution: since nouns are the least predictable part of language, dividing nouns into semi-arbitrary classes reduces the uncertainty associated with them. Corpus studies show both how gendered articles make nouns more predictable, and how in languages that lack noun class, prenominal adjectives serve a similar function. This raises the question of whether language users are sensitive to this information and actually employ it in communication. In an elicitation study, we manipulated the contextual information provided by gendered articles to German speakers, and compared their behavior to English speakers (whose articles are always uninformative). When German articles were uninformative, German and English speakers produced prenominal adjectives at the same rates. However, when articles were informative, German prenominal adjective production decreased. These results illustrate how languages use both articles and prenominal adjectives to support communicative efficiency.

Keywords: Language; information theory; grammatical gender

Introduction

Although the use of language is a defining human characteristic, detailed explanations of how it actually serves to facilitate communication have proven elusive. Although the idea that linguistic messages comprise acoustic/semantic 'atoms' whose contributions are combined and modified by compositional rules to transfer meaning between interlocutors accords well with human intuitions;, historical attempts to operationalize these components – and to state what, exactly, the atoms and rules of language are supposed to be – have raised more questions than answers (see e.g., Pinker, 2015; Ramscar, 2023). Recently this has led researchers to examine other characterizations of human communication, such as framing its processes in terms of information theory (Gibson et al, 2019), and characterizing them as serving to reduce semantic uncertainty rather than literally transfer meaning (Ramscar, 2021).

The differences between these approaches can be easily explained by considering the question of why speakers use the specific words they do. Many psycholinguistic theories assume that the answer lies – at least partly – in efficiency: speakers seek to balance production costs against what is necessary to communicate messages to listeners (Gibson et al., 2019; Grice, 1975). This assumption raises two deep questions: 1., how do linguistic signals communicate messages, and 2., how can the cost of producing a signal be evaluated? Cooperative accounts of communication (e.g., Grice, 1975) assume that utterances and the words they contain communicate 'meanings' that specify referents such that an efficient noun phrase is one that contains only those words that are required to identify a given intended referent. Accordingly, in the context depicted in Figure 1 the 'minimal' expression "the dog" is sufficient and thus should be preferred to "the blue dog" (the adjective blue is superfluous since there is only one dog), yet empirically it has been shown that English speakers consistently produce noun phrases like "the blue dog" in these contexts (Deutsch & Pechmann, 1982; Engelhardt et al., 2006; Koolen et al., 2011; Rubio-Fernández, 2016; Tourtouri et al., 2019).

However, while this might suggest that speakers prefer inefficiency, it also might not, because traditional theories have encountered a range of seemingly insuperable problems when it comes to explaining exactly how words are supposed to specify their referents (Wittgenstein, 1953). Indeed, rather than serving to explain exactly what "blue" and "dog" contribute to the communication process, the results of research on word meanings appear to suggest that

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).

framing questions about communication in this way may in fact be wrong (Ramscar & Port, 2016; 2019).

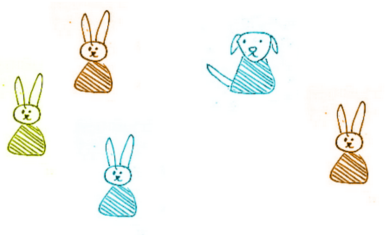


Figure 1. Which utterances are most efficient in this context: “the blue dog,” “the dog,” “the green rabbit,” or “the green one”?

By contrast, while traditional approaches have struggled to define what ‘atoms,’ rules and ‘meanings’ are, making the matter of ‘efficiency’ difficult to determine, information theory (Hartley, 1928; Shannon, 1948) has provided formal solutions to this problem in the domain of electronic signaling, presenting an explicit definition of information that enables it to be quantified in order to both define and optimize communicative efficiency. *Information* is defined as a function of the probability of the occurrence of a thing (given a distribution of things), such that if we take our things to be symbols in a code, as the probability of each symbol’s occurrence increases, the information it communicates decreases (Shannon, 1948). *Efficiency* can then be defined in terms of the average amount of information (entropy) communicated over time, which in turn enables the theory to show how by organizing encodings in systems and distributing information across sequences of symbols, codes that support communication close to a channel’s capacity can be defined. (All physical channels impose a limit – the channel capacity – on the amount of information that can be signaled across them at any given time.)

The first of these ideas can be best understood by considering the occurrence rates of the 20 most frequent English color adjectives in a corpus of contemporary American English in Table 1 (COCA, Davies, 2009; counts taken from Ramscar, 2021). According to the definition of ‘information’ above, *white* is the least informative color adjective (it occurs most often, so is most predictable), and *indigo* the most informative. More formally, Table 1 shows the Shannon information (defined as -log2(p(x))) of each adjective (in bits). This measure can be understood intuitively as follows: if an event has a 100% probability (p=1) of occurring, its information is -log2(p=1)=0 (i.e. its occurrence is uninformative). If it is 50% likely – e.g. flipping heads on a fair coin – its information is -log2(p=.5)=1. Shannon entropy (H which is defined as H(X) = -∑i P(X=i)* log2(X=i)) can then be used to specify the information (average number of bits) required to communicate the outcome of flips: since for each outcome, information = 1 and probability = .5, H=(-log2(p=.5)*.5)+ (-log2(p=.5)*.5) = 1 (i.e., outcomes encoded H=1, T=0 can be communicated using 1 bit).

	COCA Frequncy	Color Probaility	Information	P(X=i)* log2(X=i)
WHITE	256270	0.246	2.024	0.498
BLACK	216258	0.207	2.269	0.471
RED	112362	0.108	3.213	0.346
GREEN	87871	0.084	3.568	0.301
BROWN	80732	0.077	3.690	0.286
BLUE	74391	0.071	3.808	0.272
GREY/GRAY	39832	0.038	4.710	0.180
YELLOW	32288	0.031	5.013	0.155
SILVER	27668	0.027	5.235	0.139
ORANGE	27657	0.027	5.236	0.139
PINK	21385	0.021	5.607	0.115
NAVY	19172	0.018	5.765	0.106
OLIVE	15415	0.015	6.079	0.090
PURPLE	11237	0.011	6.535	0.070
LIME	6045	0.006	7.430	0.043
TAN	5267	0.005	7.628	0.039
VIOLET	4251	0.004	7.938	0.032
TURQUOISE	1847	0.002	9.140	0.016
MAROON	1174	0.001	9.794	0.011
INDIGO	1092	0.001	9.898	0.010
			Entropy =	**3.320**

Table 1: The frequencies of the 20 most frequent English color adjectives (Ramscar, 2021), their probabilities, the information communicated by each adjective, and the information entropy of this distribution (the amount of information communicated on average).

Which brings us back to color adjectives: as Table 1 shows, white has probability of (nearly) .25, so its information (approximately) =2. Following the logic of the fair coin above, if there were only 4 colors, each p=.25, the entropy of this distribution would be 2 (they could be encoded as 00, 10, 01 and 10). Yet, there are 20 colors in Table 1, and whereas their entropy would be 4.32 bits if all colors were equally frequent, empirically it is just 3.32 bits. This efficiency benefit can be understood if we consider that events can communicate different amounts of information: in the empirical distribution, *white* and *black* communicate only around 2 bits (and each occurs around 25% of the time), whereas *maroon* and *indigo* communicate around 10 bits. It follows that if the length of our encoding varies (fewer bits are used to communicate more likely events, and more for less likely events), the average number of bits needed to communicate adjectives based on the empirical distribution will be less – in this case, an order of magnitude less – than if all adjectives were equally likely, while the average amount of information communicated will be an order of magnitude closer to a channel’s capacity.

Notably, many of the properties just described have their counterparts in natural language: not only have the

probabilities of words in each language been shown to vary enormously, like the color adjectives in Table 1, they do so in seemingly systematic ways (Estoup, 1916; Zipf, 1936). Similarly, the forms of words that occur more often, and are more predictable, reliably tend to be shorter than more informative words (Piantadosi et al., 2021). Less informative words are also more likely to be phonetically reduced in production (Jurafsky et al., 2001; Bell et al., 2009), and in a similar vein, optional function words (such as 'that') are more likely to be omitted when the material they precede is more predictable (Jaeger, 2010).

Since the amount of information that can be transmitted by a single codeword will be constrained by the capacity of a channel, in order to extend the capacities of communication systems beyond this, as mentioned above, information theory utilizes combinatorics. By defining the information provided by a set of codewords conditioned on another codeword (its conditional entropy) as a function of the information provided by the conditioning codeword, the use of combinatoric techniques enables: the information in signals to be distributed across sequences of codewords; the statistical structure of data being signaled in sequences to be exploited in order to optimize the use of the channel; and 'smooth' codes that balance between the efficiency and the complexity of encoding and decoding to be defined.

As the example of color words suggests, natural languages exhibit many of the same properties as information theoretic codes: the probabilities of words vary enormously in systematic ways in all languages (Estoup, 1916, Zipf, 1936), and - as with the encoding above - words that occur more often reliably tend to be shorter (Piantadosi et al, 2021). Because of these and other similarities, many researchers have adopted information theory as a model for understanding human communication (Gibson et al, 2019). Further, taking an information theoretic approach to language can suggest functional roles for aspects of language that have seemed inexplicable using traditional, logical / compositional approaches, because highlighting the importance of the way lexical items are actually distributed in languages, serves to bring into focus the communicative problems raised by nouns.

Nouns comprise by far the largest number of lexical types in languages. This means that they dominate the long tails of linguistic distributions, and are the one part of speech (PoS) that language users never fully master, such that everyone continues to encounter previously unattested nouns throughout their lives (Ramscar et al, 2014). These statistical and distributional properties guarantee the information content of English nouns (both common and proper) is greater than that any of other English PoS, such as verbs (Ramscar et al, 2014), which means in turn that in theory, at the point that a name or noun phrase might be expected to occur in English, for example on hearing the sentence "Have you been to…" the uncertainty associated with the material following "to…" is orders of magnitude greater than it would be following "you…". In practice, however, this uncertainty is actually considerably lower: first because English uses articles to mark common nouns ("the University of Edinburgh") but not proper nouns ("Edinburgh University"), which means that only an article or a proper noun can follow "been to…"; second, because if an article does occur, the set of common nouns conditioned on it can be reduced either by the prior context (not all English common nouns plausibly follow "Have you been to…"), or by adding further context (for example, by adding an adjective: the set of nouns that can plausibly follow "Have you seen the…" will be far larger than those that can plausibly follow "Have you seen the blue…").

As in the examples above, in what follows we assume that communication is a process of incremental uncertainty reduction, in which speakers and listeners rely on a shared system of probabilistic relations between speech signals and intended messages (Ramscar et al, 2010; Ramscar, 2019, 2021), that is directly analogous to the way that codes are employed in artificial communication systems (Shannon, 1948). Crucially, from this perspective, words are assumed to incrementally guide (and restrict) expectations about not only messages, but also signals. Hence, while most accounts of reference production focus on content words, which are commonly conceptualized as direct pointers to specific (sets of) referents, this account blurs the traditional distinction between content and function words by considering the information that words provide about upcoming aspects of signal, and this in turn suggests a functional role for the existence of grammatical gender, which has often seemed puzzling (Futrell & Ramscar, 2011; Dye et al., 2018).

As has often been noted, many of the world's languages assign objects into seemingly arbitrary (and possibly sexist) noun classes that apparently lack any obvious purpose (Corbett 1991). In keeping with the obscurity of its purpose, theorists have historically treated grammatical gender as a pointless ornament (Kilarski, 1997), or at least as a redundant aspect of language that lacks any functional utility. This perspective is neatly summarized in comments by Bloomfield (1933) – "[t]here seems to be no practical criterion by which the gender of a noun in German, French, or Latin could be determined." – and Maratsos (1979): "The presence of [grammatical gender] systems in a human cognitive system constitutes by itself excellent testimony to the occasional nonsensibleness of the species."

By contrast, from an information theoretic perspective, the practice of assigning nouns to gender classes conditioned on gender marked articles (e.g. as in German) can be seen as a mechanism for redistributing the massive amount of information in the set of all nouns in much the same way as described earlier for English articles and prenominal adjectives: the neuter article "das" in German predicts that only a neuter noun can follow, for example "Meer" (sea) but not "Sand" (sand) or "Sonnenschirm" (parasol). This is no different from the way that a prenominal adjective can serve to predict a subset of nouns that are likely to follow it, such as the English "calm" can predict "sea", "person", or "music", while hardly anyone would expect nouns such as "sand", "parasol", or

"sunglasses" to follow this adjective (see also, Futrell & Ramscar, 2011; Dye et al, 2019).

Consistent with this analysis, studies of English and German corpora have shown that English prenominal adjectives convey more information about nouns than German prenominal adjectives, whereas German articles convey more information about nouns that English articles (Futrell & Ramscar, 2011; Ramscar & Futrell, 2011; Dye et al., 2018, 2019), which suggests that while German speakers have the option of using both articles and prenominal adjectives to reduce uncertainty about nouns, English speakers are largely constrained to prenominal adjectives. Building on these findings, here we examine whether the contrasting properties of German and English noun phrases can enable us to gain insight into when and why speakers use prenominal adjectives in noun phrases and, more generally, into some of the principles that govern communicative efficiency.

Corpus-based studies are essentially correlational, and thus provide only limited evidence about how people actually use words in context. Accordingly, we created an experimental reference production task, in which article informativity could be manipulated in context, to examine whether speakers actually do manage uncertainty reduction as they produce noun phrases, and whether they do so in the way that previous corpus analyses have suggested.

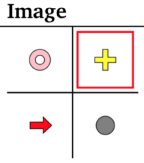
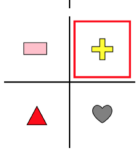
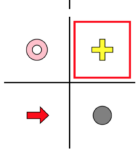
Condition	Image	Target	Competitors
German informative		**das** Kreuz, *n.*	**den** Ring, *m.* **den** Pfeil, *m.* **den** Kreis, *m.*
German uninformative		**das** Kreuz, *n.*	**das** Rechteck, *n.* **das** Dreieck, *n.* **das** Herz, *n.*
English uninformative		**the** cross	**the** ring **the** arrow **the** circle

Figure 2. Example trials in the three experimental conditions. In the German informative condition competitor nouns have a different gender from target nouns and therefore also take different articles (here given in the expected accusative form: "das" for neuter nouns and "den" for masculine nouns). Images used in the English condition were always the same as in the German informative condition.

To do so, in this preregistered study[1] we adapted a paradigm used by Wu and Gibson (2021; see also Rubio-Fernandez et al., 2021), in which participants are asked to produce a written reference for one of four different shapes that also differ in color (i.e., that are discriminable both by shape and by color). Although when given the prompt, "Please click on ...", English participants could complete this task successfully by responding "the [shape]", Wu and Gibson found that in doing so they were in fact far more likely to produce 'redundant' adjectives, responding "the [color] [shape]."

Because of the distribution of shapes across gender classes in German, we were able to create context in which the article of the target shape was informative (i.e. different from that of the other shapes in the array) or uninformative (i.e. the same as that of the other shapes in the array; see Figure 2). In the light of the foregoing, we predicted that when German articles are informative in context, German speakers would produce fewer prenominal color adjectives than the English speakers, for which articles are always (mostly) uninformative about the shape. By contrast, when German articles were uninformative in context, we expected German speakers to produce as many prenominal color adjectives as English speakers.

Method

Participants

A power simulation based on an a priori estimated effect of 70% prenominal adjectives produced in the German informative condition and 75% in the German uninformative and English condition (based on pilot data) suggested that 260 participants would be needed per condition in order to detect a difference of not more than 5% in prenominal color adjective proportion between conditions. Accordingly, anticipating that some participants might omit articles (thereby needing to be excluded), we tested 301 English-speaking and 602 German-speaking participants on Prolific, for a reward of £ 0.50 given an estimated completion time of 4 minutes (in reality, mean completion time was 154s; range = 45-1488 s). Of these, we excluded 8 English and 6 German participants who either consistently produced relative clauses (such as, e.g., "the symbol that looked like..." or gave totally unrelated answers, and 28 German and 54 English participants who did not produce articles in the majority of trials (i.e., in more than 3 trials). This left a total of 568 German and 239 English participants for analysis.

Procedure

Participants performed a written sentence completion task, completing five trials as depicted in Figure 3. The small number of trials per participant was chosen because participants have shown high consistency in the structure of produced utterances across trials in similar tasks (e.g., Wu & Gibson, 2021, Rubio-Fernandez et al. 2021, and our own preliminary experiments). Participants were asked to complete the sentence "Please click on ..." (or "Bitte klicke auf...", in German). Before the start of the experiment,

[1] The preregistration, materials, data and analysis are at http://osf.io/x2cmj (This paper reports Experiment 2).

participants were instructed to produce a grammatically correct description of the indicated shape, such that someone else could identify this shape in the same image without the marking. They were informed that the four shapes in the four images would always be different. The experiment was implemented in Qualtrics and could only be performed on a desktop computer or laptop.

Stimuli and Design

The stimuli comprised arrays of four geometrical shapes drawn from a list of nine shapes, which covered two of the German genders (as there are few feminine shape nouns, only masculine and neuter shapes were used). The shapes were colored in one of nine colors. The four shapes in an array were combined in a 2x2 grid, with a red box highlighting the target to be described (see Figure 3).

German participants were presented with one of two types of arrays: one in which the German gendered article would be informative about the target shape, and a second type in which German articles would not be informative about the target shape. These two types of shape arrays are directly aligned to the two German conditions, with participants in the German informative condition being presented with different-gender shape arrays, and participants in the German uninformative condition with same-gender shape arrays. Regarding the English condition it is important to note that since the English definite article is uninformative about noun categories, definite English articles would be uninformative in both shape array types. Accordingly, because we did not expect a difference in the answers between the two English conditions, we presented only one type of array to English participants. These were the different-gender shape arrays, so as to enable the responses in the English condition to be directly contrasted with those in the German informative condition.

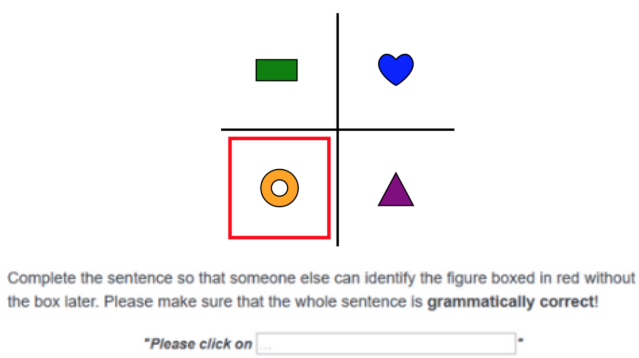


Figure 3. A sample trial with instructions.

In both array types, the position and color of the target as well as the colors of the competitors were kept constant, hence, only the form of the competitor shapes was manipulated (see Table 1). In different-gender shape arrays, the target was either commonly denoted by a neuter German noun and the competitors by masculine German nouns, or the reverse. In same-gender shape arrays, all four shapes were either commonly denoted by neuter German nouns or by masculine German nouns. Shape positions and colors were randomly determined with the constraint that each shape and each color appeared twice as a target in the stimulus set of each condition. Importantly, both color and shape were always unique in each array, meaning that both color adjectives and shape nouns would, independently of each other, be fully informative about the target shape. This yielded 18 stimuli for both different- and same-gender shape array types (in total 36 stimuli), with different-gender shape arrays being presented in the German informative and the English condition and same-gender shape arrays being presented in the German uninformative condition as described earlier. Each participant was presented with five experimental trials drawn randomly from the list of 18 stimuli created for each condition, with the restriction that every participant saw five different target shapes.

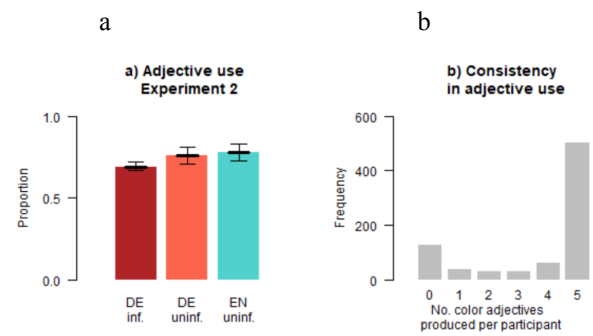


Figure 4. a) Mean proportions of prenominal color adjectives produced (black lines indicate the model estimates, CI ±1.96 SE); b) number of trials in which a prenominal color adjective was produced by a participant (shows their consistency). Data is shown excluding trials in which the article was omitted.

Analysis

To analyze the effect of condition on the production of prenominal color adjectives, we fitted a logistic regression model to color adjective productions and omissions aggregated by participant or by item (similar as in an F1 and F2 analysis, see, e.g., Clark, 1973). Note that a logistic mixed-effects regression model with random participant intercepts (as we had initially planned) overestimated mean adjective proportions for all conditions as close to 1 and explained almost all variation with the by-participant random intercepts - this was possible because of the high consistency of participants in producing or omitting adjectives (see Fig. 4b). To be able to compare adjective production behavior between conditions, we therefore did not include random by-participant intercepts and instead analyzed the data aggregated once by participant and once by item (we still report both analysis approaches in the supplementary materials under the preregistration link).

Results

All participants who omitted the article in more than 3 trials were excluded from these analyses. Figure 4a shows the mean proportion of prenominal color adjectives produced in the three conditions, which shows that both English speakers and German speakers in the uninformative condition produce more prenominal color adjectives than German speakers in the informative condition ($mean_{EN}$ = 78%, $mean_{DE:UNINF}$ = 76%, $mean_{DE:INF}$ = 69%). Note that most participants were quite consistent in their production of adjectives across the five trials as shown in Figure 4b.

The logistic regression models fitted to by-participant and by-item aggregated data were both constructed in a forward selection procedure comparing models with $\chi2$ tests and Akaike's Information Criterion. The predictor Condition significantly contributed to both the by-participant model ($\chi2(1) = 17.17$, $p < .001$, AIC difference = 15.17, as compared to the more simple predictor Language) and to the by-item model ($\chi2$ (1) = 17.17, $p < .001$, AIC difference = 10.63). Both models predicted fewer produced color adjectives in the German informative condition as compared to the two other conditions (log-odds Tukey-corrected estimates of by-participant model: $\beta_{DE:UNINF-DE:INF} = 0.354$, SE= 0.086, z-value = 4.130, $p < .001$; $\beta_{EN-DE:INF} = 0.461$, SE = 0.092, z-value = 5.009, $p < .001$), but they do not predict a difference between the German uninformative and the English condition. These results support our prediction that speakers will be more likely to produce prenominal color adjectives when articles are not informative about nouns — either due to context, as in the German uninformative condition, or due to the language, as in English

Discussion

This study adopted an information theoretic perspective to examine whether principles of communicative efficiency can explain speakers' choice of noun phrase in production. In particular, we aimed to provide a more principled account of when and why speakers produce 'overspecifying' adjectives, which remains something of a historical puzzle in research on reference production (e.g., Ford & Olson, 1975; Pechmann, 1989). Accordingly, we investigated noun phrase production in German and English, two languages in which articles and prenominal adjectives can both potentially contribute to incrementally reducing the amount of uncertainty in a signal, except to differing degrees. Consistent with previous corpus data (Dye et al., 2018; 2019) we found that English speakers produce more prenominal adjectives than German speakers. Further, going beyond this cross-linguistic difference, the results of our manipulation of German article informativity suggest that this difference arises because speakers react to local uncertainty in the signal, which is higher after an uninformative article than after an informative article. While English articles are generally less informative about upcoming nouns than German articles, German articles can be more or less informative, depending on the context. Accordingly, when German articles were uninformative about the target noun, German speakers produced color adjectives more often than when articles were informative. Crucially, prenominal adjective proportions in the German uninformative condition closely resembled the proportions produced by English speakers.

These results also contribute to a better understanding of the function of articles and prenominal adjectives. The traditional distinction between content and function words assumed by many linguistic theories implies that adjectives have a semantic function and articles a grammatical one. If, however, we consider the information that articles or adjectives provide about the signal or the non-linguistic environment, this functional distinction seems more nuanced. These findings indicate that speakers make up for uninformative articles by producing additional prenominal color adjectives, which suggests that, functionally, both articles and prenominal adjectives provide information that serves to smooth signals. This observation directly ties in with previous evidence that people use both word types to predict upcoming nouns (DeLange et al., 2005; Lau et al., 2016; Tribushinina & Mak, 2016; Wicha et al., 2004).

Notably, in our study, the prenominal color adjectives that speakers produced could not only be informative about the target noun (i.e., the signal) but also about the target shape (i.e., a state of the world). While we held the informativity about the target shape constant across conditions, there is evidence that the information an adjective can provide about, for example, a target in a set of competitor objects, is a crucial factor that influences prenominal adjective production: speakers have, for example, been found to produce more prenominal color adjectives in polychrome displays, in which color is informative about a specific referent, than in monochrome displays, where color is uninformative about a target referent (e.g., Koolen et al., 2013; Rubio-Fernández, 2016). In our study, this could have been the reason for the high baseline of produced prenominal color adjectives (roughly 70% in the German informative condition).

The discriminative, information theoretic approach to noun phrase production examined here thus provides new insight in the principles that govern the structure of noun phrases produced in different contexts. Taking account of the information that current parts of the signal provide about the upcoming signal in a specific context — aside from the information it provides about the non-linguistic environment — revealed that speakers' behavior is influenced by signal uncertainty when they produce noun phrases. Crucially, this suggests that we might need to redefine the concept of a message in natural communication: while the ultimate aim of a speaker may be to communicate 'content' – to help the listener discriminate between representations of possible states of the world – our findings suggest that the information communicated by (incrementally incoming) words in an utterance are not only aimed as influencing expectations about messages, but also about upcoming signals as well.

References

Akaike, H. (2011). Akaike's information criterion. International encyclope- dia of statistical science (pp. 25–25). Springer.

Bell, A., Brenier, J. M., Gregory, M., Girand, C., & Jurafsky, D. (2009). Predictability effects on durations of content and function words in conversational English. Journal of Memory and Language, 60(1), 92-111.

Bloomfield, L. (1933). Language. New York: Holt.

Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. Journal of Verbal Learning and Verbal behavior, 12(4), 335–359.

Corbett, G.G. (1991). Gender. Melbourne: Cambridge University Press.

De Lange, J., Vasic, N., & Avrutin, S. (2009). Reading between the (head) lines: A processing account of article omissions in newspaper head- lines and child speech. Lingua, 119(10), 1523–1540.

Deutsch, W., & Pechmann, T. (1982). Social interaction and the develop- ment of definite descriptions. Cognition, 11(2), 159–184.

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2017). A functional theory of gender paradigms. In F. Kiefer, J. Blevins, & H. Bartos (Eds.), Perspectives on morphological structure: Data and analyses (pp. 212– 239). Brill.

Dye, M., Milin, P., Futrell, R., & Ramscar, M. (2018). Alternative solutions to a language design problem: The role of adjectives and gender marking in efficient communication. Topics in cognitive science, 10(1), 209-224.

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

Estoup, J. B. (1916) Gammes Stenographiques. Institut Stenographique de France, Paris

Ford, W., & Olson, D. (1975). The elaboration of the noun phrase in children's description of objects. Journal of Experimental Child Psychology, 19(3), 371–382.

Futrell, R., & Ramscar, M. (2011). German grammatical gender manages nominal entropy. Presentation at Information-Theoretic Approaches to Linguistics 2011.

Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. Trends in Cognitive Sciences, 23(5), 389–407.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Speech acts (pp. 41–58). Brill.

Hartley, R. V. (1928). Transmission of information 1. Bell System Technical Journal, 7(3), 535–563.

Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. Cognitive Psychology, 61(1), 23–62.

Jurafsky, D., Bell, A., Gregory, M., & Raymond, W. D. (2001). Probabilistic relations between words: Evidence from reduction in lexical production. Typological studies in language, 45, 229-254

Kilarski. M. (2007). On grammatical gender as an arbitrary and redundant category. In Douglas Kilbee, editor, History of Linguistics 2005: Selected papers from the 10th International Conference on the History of Language Sciences (ICHOLS X), pages 24–36. John Benjamins, Amsterdam

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

Koolen, R., Goudbeek, M., & Krahmer, E. (2013). The effect of scene variation on the redundant use of color in definite reference. Cognitive

Lau, E. F., Namyst, A., Fogel, A., & Delgado, T. (2016). A direct comparison of N400 effects of predictability and incongruity in adjective-noun combination. Collabra, 2(1).

Lemke, R., Horch, E., & Reich, I. (2017). Optimal encoding!-Information theory constrains article omission in newspaper headlines. Proceed- ings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, 131–135.

Lew-Williams, C., & Fernald, A. (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. Journalof Memory and Language, 63(4), 447–464.

Maratsos, M.P. (1979). How to get from words to sentences. In D. Aaronson & R. Rieber (eds.), Perspectives in Psycholinguistics. Hillsdale, N.J.: Erlbaum.

Pechmann, T. (1989). Incremental speech production and referential over-specification. Liguistics, 27(1), 89–110.

Piantadosi, S. T., Tily, H., & Gibson, E. (2012). The communicative function of ambiguity in language. Cognition, 122(3), 280–291. Pickering, M. J., & Garrod, S. (2013). An integrated theory of language production and comprehension. Behavioral and Brain Sciences, 36(4), 329–347.

Pinker, S. (2015). Words and rules: The ingredients of language. Basic Books.

Ramscar, M. (2019). Source codes in human communication. arXiv preprint arXiv:1904.03991.

Ramscar, M. (2023). A discriminative account of the learning, representation and processing of inflection systems. Language, Cognition and Neuroscience, 38(4), 446-470.

Ramscar, M. (2021). How children learn to communicate discriminatively. Journal of Child Language, 48(5), 984-1022.

Ramscar, M., & Futrell, R. (2011), The predictive function of prenominal adjectives. In Workshop on information-theoretic approaches to linguistics.

Ramscar, M., & Port, R. (2019). Categorization (without categories). In E. Da̧browska & D. Divjak (Eds.), Cognitive linguistics - foundations of language (pp. 87–114). De Gruyter Mouton.

Ramscar, M., & Port, R. F. (2016). How spoken languages work in the absence of an inventory of discrete units. Language Sciences, 53, 58–74.

Ramscar, M., Hendrix, P., Shaoul, C., Milin, P., & Baayen, H. (2014). The myth of cognitive decline: Non-linear dynamics of lifelong learning. Topics in Cognitive Science, 6(1), 5–42.

Ramscar, M., Yarlett, D., Dye, M., Denny, K., & Thorpe, K. (2010). The effects of feature-label-order and their implications for symbolic learning. Cognitive Science, 34(6), 909–957.

Rohde, H., & Rubio-Fernandez, P. (2021). Color interpretation is guided by informativity expectations, not by world knowledge about colors. https://doi.org/10.31234/osf.io/jmd9w Rokers, B., Mercado III,

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

Rubio-Fernandez, P. (2021). Color discriminability makes over-specification efficient: Theoretical analysis and empirical evidence. Humanities and Social Sciences Communications, 8(1), 1–15.

Rubio-Fernandez, P., Mollica, F., & Jara-Ettinger, J. (2021). Speakers and listeners exploit word order for communicative efficiency: A cross- linguistic investigation. Journal of Experimental Psychology: General, 150(3), 583–594. https://doi.org/doi:10.1037/xge0000963

Rupp, L., & Tagliamonte, S. A. (2019). "They used to follow Ø river": The zero article in York English. Journal of English Linguistics, 47(4), 279–300.

Shannon, C. E. (1948). A mathematical theory of communication. Bell System Technical Journal, 27, 379–423.

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. Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: Evidence from eye-tracking. Journal of Child Language, 43(2), 425–441. https://doi.org/10. 1017/S0305000915000173

Tribushinina, E., & Mak, W. M. (2016). Three-year-olds can predict a noun based on an attributive adjective: evidence from eye-tracking. Journal of child language, 43(2), 425-441.

Wicha, N. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. Journal of Cognitive Neuroscience, 16(7), 1272– 1288.

Wittgenstein, L (1953) Philosophical Investigations, G. E. M. Anscombe, trans., Oxford: Basil Blackwell

Wu, S. A., & Gibson, E. (2021). Word order predicts cross-linguistic differ- ences in the production of redundant color and number modifiers. Cognitive Science, 45(1), e12934.

Zipf, G. K. (1949). Human behaviour and the principle of least effort. Cambridge, Mass.: Addison-Wesley Press.