

How gender information influences spontaneous speech in context

Amalia Spyromilio¹, Holly Jenkins², Elizabeth Wonnacott², Michael Ramscar¹

¹University of Tübingen

²University of Oxford

Abstract

Although noun class has been thought to lack function, it is suggested that gender-marked articles might help to enhance the predictability of upcoming nouns, much like prenominal adjectives. Supporting this idea, in a written paradigm Hoppe et al. (2025) showed that German speakers produced fewer prenominal adjectives than English speakers, but only when articles provided gender information (which English articles do not provide). Using a modified paradigm to examine this phenomenon in speech, we found that spoken prenominal adjectives were consistently produced across both informative and uninformative contexts. Instead, planned articulatory analyses reveal that in uninformative contexts, German speakers enhance, producing longer, more acoustically informative noun phrases at second mention than at first mention; in informative contexts, German speakers reduce, showing the opposite pattern, as do English speakers. We discuss these results in terms of how uncertainty about signals and the messages they communicate influence articulation in different languages and contexts.

Keywords: Language; speech production; information theory; grammatical gender.

Introduction

Many of the world's languages assign objects to seemingly arbitrary noun classes that apparently lack any obvious purpose (Corbett 1991). In keeping with the obscurity of its purpose, historically grammatical gender has often been treated as a pointless ornament by theorists (Kilarski, 2008). Gender was seen as a redundant aspect of language that lacked any functional utility, a perspective summarized by Bloomfield (1933, 280), who noted that “[t]here seems to be no practical criterion by which the gender of a noun in German, French, or Latin could be determined.”

However, from a predictive/information theoretic perspective, a potential function for noun class appears obvious: making nouns more predictable. Nouns (both common and proper) are by far the largest set of lexical types in most languages, and they dominate the long tails of low frequency types found in the distributions of words and other lexical combinations. These statistical facts guarantee that, on average, unless their prediction is supported systematically, nouns will inevitably be the least predictable part of speech in any communicative code (Ramscar, 2013; Dye et al., 2018). In morphologically richer languages like German, assigning nouns to classes conditioned on gender-marked articles provides a mechanism to redistribute the massive amount of information that would otherwise be associated with the set of all nouns: e.g., the neuter article “das” in German predicts that only a neuter noun can follow,

for example “Meer” (das (n.) Meer - the sea) but not “Sand” (der (m.) Sand - the sand) or “Sonnenschirm” (der (m.) Sonnenschirm - the parasol).

Notably, from this perspective, the proposed functional role of gendered articles can be seen to parallel the way that prenominal adjectives also help to reduce uncertainty about the subsets of nouns likely to follow them: in a language like English (which lacks grammatical gender) hearing the adjectives “nice cold” will make nouns like “beer”, “drink”, and “glass of water” more predictable while reducing the likelihood that “sweater”, “hamburger” or “cup of tea” will occur (see also, Ramscar & Futrell, 2011; Dye et al., 2018).

Consistent with this proposal, 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 than English articles (Futrell & Ramscar, 2011; Ramscar & Futrell, 2011; Dye et al., 2017, 2018).

To examine whether speakers do in fact use prenominal adjectives and gendered articles to help ‘smooth’ the information provided by nouns, Hoppe et al. (2025) conducted a written study of noun phrase production in English and German. Participants were asked to generate noun phrases to label target items (marked with a square in Figure 1), which were presented in arrays containing 3 other items. All items were different and had different colours. Crucially, in German, a language with three noun classes marked by different articles, an informative and uninformative condition was constructed in which the shape’s article was either unique in the array (informative condition) or shared the same article with the other shapes in the array (uninformative condition). Participants generating noun phrases could potentially be aided by the extra information provided by the unique gendered article in the informative condition (see Figure 1). Although both English and German participants could complete the shape naming task successfully by responding “the [shape]”, previous studies have shown that they were in fact far more likely to produce ‘redundant’ adjectives, responding “the [color] [shape],” (see Rubio-Fernandez, 2016; 2021; Rubio-Fernandez et al., 2021; Wu & Gibson, 2021).

However, Hoppe et al. (2025) found that when German speakers produced target noun phrases in the informative context (where the unique articles helped predict the target nouns), they used fewer prenominal colour adjectives than German speakers in the uninformative condition. Notably, German speakers in the uninformative condition produced

colour adjectives at the same rate as English speakers (where, since English lacks gender, articles were always uninformative).

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

Figure 1: Example trials in the three conditions tested in Hoppe et al (2025). In the German informative condition the target nouns have a different gender from the competitor nouns and thus have unique articles (here given in the expected accusative form: das for neuter nouns and den for masculine). Articles in the English condition were always uninformative.

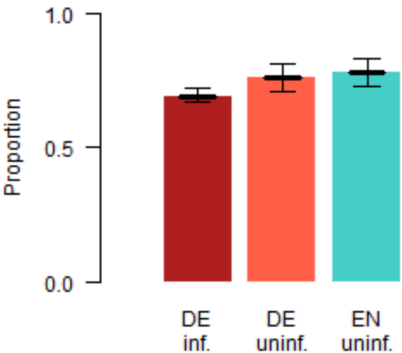


Figure 2: Mean proportions of prenominal color adjectives produced in writing by 900 participants in Hoppe et al (2025). As can be seen, adjectives were produced more often in the German uninformative and English conditions.

How does gender information influence spontaneous speech?

The following study was designed: **first** to see how the behavior observed by Hoppe et al. (2025) in writing transfers to spontaneous speech (which affords greater behavioral flexibility, e.g., reduction / enhancement); and **second** to explore a further factor that might affect the way in which participants produce prenominal color adjectives in the contexts tested by Hoppe et al.

First: Participants in Hoppe’s study responded by typing. It is important to note that writing and speaking are very

different behaviors. Although writers can abbreviate words (e.g., by writing *info* vs *information*, see Mahowald et al., 2013), their ability to do so is limited in scope, such that when it comes to the use of adjectives, their only option is either to write “the red ball” or “the ball”. By contrast, articulation in speech is very flexible. When signals are highly predictable/less informative, speakers are able to barely articulate them at all (because hearers are able to ‘perceive’ highly predictable speech sounds that are hardly, or even not actually present in the signal, Kohler, 1998; Tomaschek et al., 2020). By contrast, when messages are uncertain, speakers can enhance signal production (Tomaschek & Ramscar, 2022). On the basis of these considerations we did not expect the patterns of noun phrase production observed in writing to be exactly the same as those produced in speech.

Second: There is a potential problem in the way many current approaches apply information theory to language. Information theory (Hartley, 1928; Shannon, 1948) uses the probability of the occurrence of an event to quantify its information (entropy). In common with most studies in the field (e.g., Jaeger & Levy, 2006), Dye et al. (2018)’s methods determined the uncertainty associated with adjectives and nouns by using their corpus frequencies to quantify their occurrence probabilities. Because this approach measures nouns’ and adjectives’ likelihoods of occurrence as an average of each of their occurrences across a range of aggregated contexts, it implicitly assumes that words occur across texts in a uniform manner.

However, this is inconsistent with the fact that the occurrence of words in context is bursty (Katz, 1996). Most words, especially low frequency ones, are likely to occur multiple times in a few texts, but never in most others. Accordingly, after they first appear in a text, their chances of reappearance increases. This suggests that the likelihoods of word occurrences in context are dynamic. Consistent with this, it has been shown that reference phrases tend to become shorter with increased mentions of a noun in a discourse (Krauss & Weinheimer, 1964). These considerations suggest that estimates of word predictability made by averaging the occurrences of words across a corpus will not be accurate in many specific contexts.

To examine whether language users are sensitive to this, Kemper et al. (2024) conducted a spontaneous speech elicitation study to see whether English speakers used prenominal modifiers such as adjectives to provide information about upcoming nouns, and also whether they used fewer adjectives after the noun had already been mentioned. The logic behind this was as follows: if language users employ adjectives to reduce the uncertainty associated with nouns, and if they are also sensitive to burstiness, then when a noun occurs in context, their expectation that it will reoccur will increase. This means there will be less need for them to use an adjective to increase the predictability of the noun in subsequent mentions. Consistent with this hypothesis, Kemper et al. found that participants were more likely to use adjectives to

modify nouns at first mention, such that significantly fewer adjectives were used in subsequent mentions.

Accordingly, the following study was designed to examine not only whether subsequent mentions affect the amount of information in the noun phrases produced in speech, but also how these effects influence the spontaneous articulation of noun phrases in context.

Methods

It is important to note that this is a spontaneous speech elicitation study. Almost by definition, the study of spontaneous speech phenomena in a fully controlled environment is difficult, if not impossible to achieve. Unlike scripted speech, spontaneous speech cannot be pre-planned, and this means that while the behaviour of participants can be prompted, it cannot be controlled. For example, not only can we not control what our participants say in this study, we cannot control their identification of any of the target items they are required to name, nor the other items that provide context. Accordingly, while we believe that our paradigm succeeds in its basic aim – that of reliably eliciting noun phrases of a certain form – the results of pilot studies alerted us to the high degree of variability in the speech behaviour we could expect to elicit, and the importance of pre-establishing exclusion criteria prior to acquiring data.

To elicit spontaneously spoken noun phrases, we developed a paradigm that presented participants with a screen displaying a circle. Different objects (shapes) then appeared on screen, and then changed position to inside the circle or within the circle one after another. After each object’s movement, participants were asked to record a sentence that described the movement in a way that would allow it to be recreated by another, naive participant at a later time. (E.g. Along the lines of: “*The yellow heart has moved so it is now above the red arrow.*”)

The movement patterns were designed so that the target item would be mentioned two times, allowing us to analyse patterns of noun phrase production across mentions. Importantly, the items and their colours were always unique in the array (and as noted above, studies have shown that English and German speakers tend to produce these adjectives despite their apparent redundancy). All participants completed two trials (each comprising 2 target movements and 2 distractor movements): German participants completed one trial each of the informative and uninformative conditions. English participants completed two uninformative trials (there is no English informative condition).

70 native English speakers and 66 native German speakers were recruited via Prolific (<https://www.prolific.com>) to complete the task. The experiment lasted approximately 5 minutes, and participants were paid at a rate of £9 per hour. Participants were required to be L1 speakers, but no other demographic data was recorded. An automatic microphone check was completed, and any participants who did not complete the experiment within 60 minutes were automatically excluded.

18 English speakers and 11 German speakers were automatically excluded due to failed microphone tests or exceeding the time limit. Another 28 English speakers and 21 German speakers were excluded on the basis of the following predefined exclusion criteria: production of more than 55 or less than 6 words, inaudible audio files, clearly not being an L1 speaker, or using ungrammatical sentences.

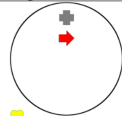
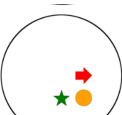
Condition	Image	Target	Distractors
German informative		der Pfeil, m.	das Kreuz, n. das Quadrat, n. das Herz, n.
German uninformative		der Blitz, m.	der Kreis, m. der Stern, m. der Pfeil, m.
English uninformative		the arrow	the donut (shape) the circle the star

Figure 3: The top 2 panels show the two different conditions German participants encountered in the study. In the informative condition (top), the target noun has a distinct noun class. In the uninformative condition (middle) all nouns share the same noun class. The bottom panel shows the English condition (English has no noun class, so articles are uninformative). The objects started outside the reference frame (the large circle) and then successively moved into it and around it. Participants were asked to describe the movements, which required them to mention each object at least twice.

For the articulatory speech analysis additional predefined exclusion criteria were applied in order to obtain comparable phrases. The noun phrases describing the target items had to be identical across mentions, a single movement description had to include no other repeated mentions of the target item, noun phrases had to include adjectives (almost all noun phrases did) and disfluencies should not interfere with key aspects of the noun phrase. After exclusion for the purpose of the speech analysis, 54.81% of the German target noun phrases and 57.14% of the English target noun phrases remained. Duration was used as a measure of reduced/enhanced speech and zip file size was used as a rough estimate of information (i.e. the size of the compressed audio file was taken to reflect the amount of information contained in the original speech signal).

Data Analysis

The data from the speech elicitation task were analysed using linear mixed effect models, as duration and zip file

size are continuous. To compare duration and zip file size for the German dataset alone, informativity (informative or uninformative) and mention (first and second), as well as mean intensity (in the duration model) were included as fixed effects, and participant was included as a random effect. For the comparison of duration and zip file size between the English and German uninformative conditions, language (English or German) and mention (first and second), as well as mean intensity (in the duration model) were included as fixed effects, and participant was included as a random effect. Across all models, all fixed factors were centred so that the intercept reflected the grandmean (and coefficients are main rather than simple effects).

Results

In comparison to the written responses collected by Hoppe et al. (2025), in their spontaneously produced spoken noun phrases our participants rarely dropped adjectives: only 16.3% of German target noun phrases and 7.6% of English target noun phrases were produced without an adjective. This means that no significant effect of information and mention was seen at a lexical level of analysis in these spontaneous speech responses.

However, articulatory speech analysis of the target noun phrases indicated that the articulatory behaviour of the speakers was sensitive to our contextual manipulation and changed systematically in response to it. Notably, however, the results we obtained were more complex than our initial hypotheses had predicted. We had anticipated that as compared to the German uninformative condition, in the German informative condition adjectives would be produced in a more reduced form, reflecting the extra information provided by the gendered article. In both conditions, we expected productions at second mention would be reduced as compared to at first mention.

At two planned levels of analysis - duration and zip file size - we did not find significant effects of informativity at the article and adjective level. Rather, we observed significant effects at the phrase and noun level¹: there were significant interactions between information and mention for both phrase duration ($\beta = 0.2$, $SE = 0.086$, $p = .02$; Figure 4)² and phrase zip file size ($\beta = 28.17$, $SE = 13.49$, $p = .04$; Figure 5) in German. More specifically, in the **German uninformative** condition, phrase and noun production was fastest, and least informative at first mention, and then slower, and more informative at second mention. By contrast, in the **German informative** condition, phrase production was slower, and more informative at first mention, and then faster, and less informative at second mention (see Figures 4 & 5).

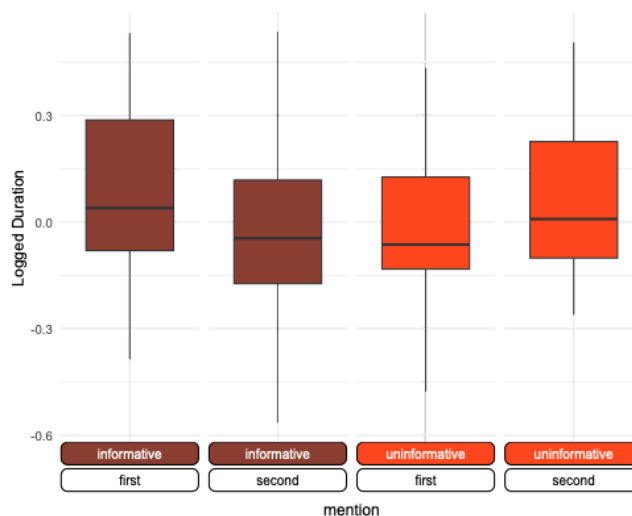


Figure 4: Log duration (y-axis) of German phrases split up over the four test conditions (x-axis). From left to right: informative first, informative second, uninformative first, and uninformative second.

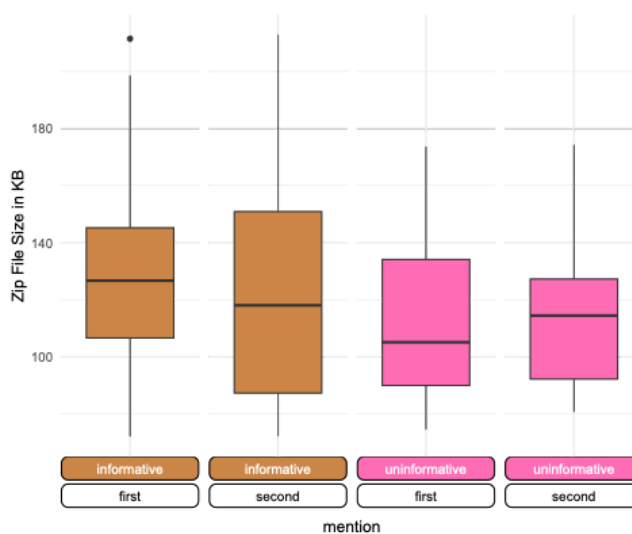


Figure 5: Zip file size in kibibytes (y-axis) of the German phrases split up over the four test conditions (x-axis). From left to right: informative first, informative second, uninformative first, and uninformative second.

Because English has no gendered articles, it was only possible to test English speakers in a condition with only one kind of article present. This meant that the articles in the English (uninformative) condition provided the same information as they did in the German uninformative condition, in that the same article was associated with all of the nouns in both. Models comparing first and second mention in the English condition and the German uninformative condition show a significant interaction

¹ For the purposes of brevity, we focus on the phrase level analyses here.

² For clarity, we have plotted these results with outliers removed. All statistical analyses were conducted on the full dataset.

between language and mention for duration ($\beta = -0.21$, $SE = 0.087$, $p = .02$; Figure 6) and for zip file size ($\beta = -29.81$, $SE = 14.37$, $p = .04$; Figure 7). That is, an effect of language was present despite the fact that the information provided by the articles in context in both languages appeared to be exactly the same.

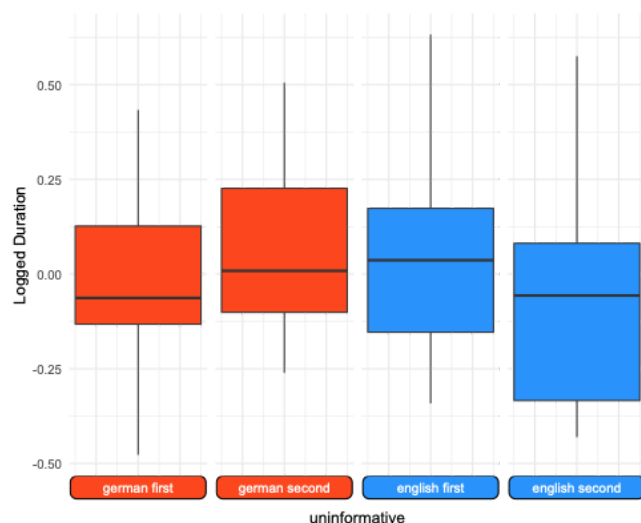


Figure 6: Log duration (y-axis) of German and English phrases split up over the first and second mention in the uninformative condition (x-axis).

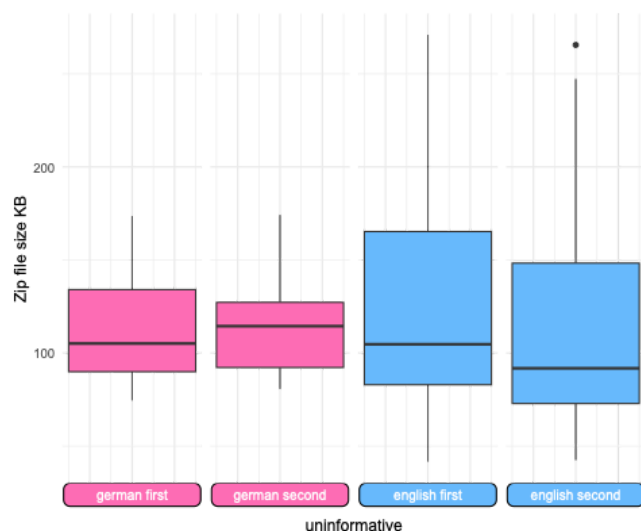


Figure 7: Zip file size in kibibytes (y-axis) of German and English phrases split up over the first and second mention in the uninformative condition (x-axis).

Interestingly, despite the lack of information provided by the articles in these conditions, the pattern of responses at first and second mention in the English uninformative

condition did not resemble the pattern observed for the German uninformative condition. Rather, the German informative condition (where the article provides information about the target noun) more closely resembled the English uninformative condition (see Figure 6 & 7).

Discussion

Unlike in previous findings from studies that elicited written responses, our participants spontaneously produced high numbers of adjectives in speech regardless of whether the articles were informative or not. By contrast, our results revealed very different patterns of spoken behavior at first and second mention by participants in the German informative and uninformative conditions: participants in the uninformative condition were faster at first mention than participants in the informative condition; however, while the phrase durations in the German informative condition decreased (as typically happens at second mention, and as was observed in the English condition, where articles were uninformative), in the German uninformative condition phrase durations increased at second mention.

A potential explanation for these effects has been proposed by Tomaschek and Ramsar (2022) who observe that a number of apparently contradictory findings relating to paradigmatic effects on the articulation of morphological markers can be easily reconciled by distinguishing between the effects of uncertainty on articulation that relate to signals and those that relate to messages. Specifically, whereas signal uncertainty (i.e., the amount of contextual support for the occurrence of a morphological marker such as an affix) modulates the degree to which articulation is *reduced*, increases in message uncertainty (i.e., the amount of contextual support for the message communicated by the affix) modulates the degree to which articulation is *enhanced*.

In the case of our German findings, we assume that in the spoken production of noun phrases, the effects of signal and message uncertainty occur on different timescales. From this perspective, signal uncertainty is associated with moment articulation as the signal (noun phrase) is produced in time. Meanwhile, message uncertainty depends on the ultimate ambiguity of the noun phrase, and hence can only be ascertained when the message (whole phrase) has been produced. Given that these two factors appear to modulate articulation in different ways, and necessarily seem to operate across different timescales in production, they can potentially account for the two different patterns of articulatory behavior we observe. At first mention, the greater support for the target noun phrase in the uninformative condition, where all of the items support the production of the same inflected article (signal) lead to it being articulated in a more reduced form than in the informative condition, where there is less evidence for the specific article (i.e., signal uncertainty will be greater, and hence the article less reduced).

The increased evidence for the article in the uninformative condition, and the support that inflected

articles provide for accessing inflected adjectives and nouns in German will then further reduce uncertainty about the whole noun phrase, leading to a noun-phrase that is less uncertain and hence more reduced in the uninformative condition than in the informative condition.

However, while there is more evidence for the signal as first mention proceeds in the uninformative condition, the noun phrases produced in this context will be more ambiguous in context than those in the uninformative condition, simply because they do not convey the extra information provided by the unique article. Moreover, if we also assume that the evolved structure of the German gender system tends to provide a degree of support for messages (as in, for example, the fact that the German words for knife, fork and spoon all take different gendered articles), then it follows that at the end of first mention, participants in the uninformative condition (which we have engineered to remove the support provided by gender) will have produced noun phrases that not only reduce message uncertainty to a lesser extent than participants in the informative, but also reduces it less than German noun phrases are typically wont to do. That is, this context causes them to produce noun phrases that are more ambiguous than German noun phrases tend to be at first mention, and this extra message uncertainty leads them to enhance their articulation at second mention.

This last suggestion is supported both by two further considerations. First, in the informative context, where the noun phrases contained the extra disambiguation information provided by the articles, the noun phrases produced were reduced rather than enhanced at second mention (because first mention reduced uncertainty about the message and the signal, see Kemper et al., 2024). Second, the same uninformative context that resulted in reduced noun phrases at first mention and enhanced noun phrases at second mention in our study prompted Hoppe et al.'s (2025) participants to enhance their written noun phrases (with more adjectives). This pattern of results can be seen to reflect the different time courses and pressures associated with speech versus writing, and this difference can be taken to suggest that the noun phrases in the uninformative context contained less disambiguating information than the noun phrases in the informative context. This then suggests that our participants failed to accommodate for the high message uncertainty associated with the uninformative condition at first mention, and thus had to deal with it at second mention, whereas Hoppe et al.'s participants made these accommodations in the course of writing their responses.

This interpretation is further supported by the fact that our English participants behaved like our German participants in the informative condition. Since our English participants did not expect articles to contribute much information to their noun phrases, and would hence be unperturbed by the fact that every phrase began with the same article, they would be insensitive to the “uninformative” nature of the context (a label which, except in contrast to German, makes little sense

in English). Since the degree of message uncertainty at the end of production in the English condition would be exactly what English participants are accustomed to expect, they could not experience the untoward increase in message uncertainty we engineered into our German uninformative condition, and thus they proceeded to reduce their noun phrases at second mention on the basis that first mention had made their signalling more predictable.

The information theoretic approach to noun phrase production examined here provides new insight into the principles that govern the structure of noun phrases and the way that they are produced in different contexts in speech. In particular, they highlight the critical role that uncertainty, both about the production of the signal, and the way that the signal serves to discriminate the message it communicates can both contribute to both what gets produced, and the way it is produced. In particular, our findings suggest that the uncertainty associated with the signal that is being transmitted exerts its influence on articulation at a different timescale than the uncertainty associated with the message that is being communicated. Signals are transmitted serially, by a series of incrementally produced forms, and thus signal uncertainty decreases and increases incrementally, as each form is produced. By contrast, since the uncertainty associated with the message can sometimes only be estimated once it has been fully communicated, it may be that the ambiguity associated with a message can only be appreciated after the whole signal has been transmitted in context. Further, whereas the uncertainty associated with signals can be somewhat objectively measured (since it is largely determined by the distribution of forms in a context and the prior distribution of forms a speaker/hearer has encountered), message uncertainty may rely more on less easily estimated factors. At first blush, the German uninformative condition and our English condition appeared to provide the same information to our participants. Yet their behavior suggests otherwise, indicating that their different linguistic experiences have led to them having priors that any simple comparison of the contexts will fail to capture.

Quine (1960) held that translation is indeterminate – both between languages, but also between individuals – because the propositions that many theories suppose messages comprise do not actually exist. While this idea has long been controversial, the behavior of our German and English participants when communicating in what appeared to be exactly the same contexts suggests that, at least when it comes to the production of speech in context, Quine might well have been right.

Acknowledgments

The contributions of A.S. & M.R. were supported by a grant from the Deutsche Forschungsgemeinschaft (DFG 547529231); E.W. & H.J. were supported by the Arts and Humanities Research Council (AHRC AH/Z507428/1). No AI was used in writing this submission.

References

- Bloomfield, L. (1933). *Language*. New York: Holt.
- Corbett, G.G. (1991). *Gender*. Melbourne: Cambridge University Press.
- 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.
- Futrell, R., & Ramscar, M. (2011). German grammatical gender manages nominal entropy. Presentation at *Information-Theoretic Approaches to Linguistics*, University of Colorado Boulder, July 16–17, 2011.
- Hartley, R. V. L. (1928). "Transmission of Information" (PDF). *Bell System Technical Journal*. 7 (3): 535–563. doi:10.1002/j.1538-7305.1928.tb01236.x.
- Hoppe, D., Gibson, E., van Rij, J., Hendriks, P., & Ramscar, M. (2025). How grammatical gender supports efficient communication. In *Proceedings of the 47th Annual Meeting of the Cognitive Science Society* (pp. 2500–2507).
- Jaeger, T., & Levy, R. (2006). Speakers optimize information density through syntactic reduction. *Advances in neural information processing systems*, 19.
- Katz, S. M. (1996). Distribution of content words and phrases in text and language modelling. *Natural language engineering*, 2(1), 15–59.
- Kemper, S. S., Jenkins, H. E., Wonnacott, E., & Ramscar, M. (2024). Rethinking Probabilities: Why Corpus Frequencies Cannot Capture Speakers' Dynamic Linguistic Behavior. In *Proceedings of 46th Annual Meeting of the Cognitive Science Society*.
- Kilarski, M. (2008). On grammatical gender as an arbitrary and redundant category. In *History of Linguistics: Selected papers from the Tenth International Conference on the History of the Language Sciences (ICHOLS X)*, Urbana-Champaign, Illinois (pp. 24–36). John Benjamins Publishing Company.
- Kohler, K. J. (1998). The disappearance of words in connected speech. *ZAS Papers in Linguistics*, 11, 21–33.
- Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science*, 1(1), 113–114.
- Mahowald, K., Fedorenko, E., Piantadosi, S. T., & Gibson, E. (2013). Info/information theory: Speakers choose shorter words in predictive contexts. *Cognition*, 126(2), 313–318.
- Quine, W. V. O. (1960). *Word and object*. MIT press. Cambridge MA.
- Ramscar, M. (2013). Suffixing, prefixing, and the functional order of regularities in meaningful strings. *Psihologija*, 46(4), 377–396.
- Ramscar, M., & Futrell, R. (2011). The predictive function of prenominal adjectives. In *Information-Theoretic Approaches to Linguistics*, University of Colorado Boulder, July 16–17, 2011.
- Rubio-Fernández, P. (2016). How redundant are redundant color adjectives? an efficiency-based analysis of color overspecification. *Frontiers 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>
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27(3), 379–423.
- Tomaschek, F., & Ramscar, M. (2022). Understanding the Phonetic Characteristics of Speech Under Uncertainty—Implications of the Representation of Linguistic Knowledge in Learning and Processing. *Frontiers in Psychology*, 1833.
- Tomaschek, F., Arnold, D., Sering, K., Tucker, B. V., van Rij, J., & Ramscar, M. (2021). Articulatory variability is reduced by repetition and predictability. *Language and speech*, 64(3), 654–680.
- 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.