Semantic contrast ahead: contrast guides pre-planning in complex noun-phrase production

Jens Roeser ⓘD, Mark Torrance and Thom Baguley

Department of Psychology, Nottingham Trent University, Nottingham, UK
Corresponding author: Jens Roeser; Email: jens.roeser@ntu.ac.uk

Abstract

Whether or not pre-planning extends beyond the initial noun in a noun phrase depends, in part, on the phrase's dependency structure. Dependency structure disambiguates, in many contexts, the noun phrase's reference. In the present experiment ($N = 64$), we demonstrate that advance planning is affected by the extent to which a dependency supports semantic disambiguation. Participants produced noun phrases in response to picture arrays. Syntax and lexemes were held constant, but semantic scope was manipulated by varying the contrastive functions of the first and the second noun. Evidence from eye-movement data revealed a stronger tendency for early planning in the extended scope condition. This is evidence that pre-planning requirements of structurally complex noun phrases are, in at least some contexts, determined by semantic functions.

Keywords: eye tracking; hierarchical planning; language production; planning scope; sentence processing; semantic structure

Before articulation can start, speakers must, to some extent, mentally prepare ('pre-plan') what they are about to say. There is debate, however, over the scope of this pre-planning. How far into the to-be-produced utterance (or text) that they are about to produce must planning reach before output can be initiated? For example, recent research on language production has explored the particular conditions under which pre-planning must extend beyond the initial determiner-noun pair (Gleitman et al., 2007; Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014; Kuchinsky, 2009; Kuchinsky et al., 2011). So, for a target utterance *A-B*, in some syntactic, semantic or lexical contexts, articulation can start as soon as *A* has been planned (e.g. Ferreira & Swets, 2002; Griffin, 2001; Griffin & Bock, 2000). In other contexts, *A-B* must be planned (Allum & Wheeldon, 2009; Konopka, 2012; Lee et al., 2013; Wheeldon et al., 2013). The research that we report in this article concerns how the language production system decides whether or not this more extensive pre-planning is required?

Mainstream contemporary theory of language production (often called syntax-based or hierarchical planning; e.g. Bock & Ferreira, 2014; Chang et al., 2006; Dell & O'Seaghdha, 1992; Ferreira & Dell, 2000; Garrett, 1975) assumes that minimal pre-planing scope – and particularly the fact that in some contexts this extends beyond the initial noun – is syntactically determined[1]: Speakers (or writers[2]) must pre-plan across a minimal syntactic unit before they can start outputting speech. One possibility is that before articulation can start, speakers must pre-plan the full verb-argument noun phrase (NP), at least for relatively short, simple sentences (e.g. Lee et al., 2013; Martin et al., 2010, 2014; Roeser et al., 2019; Wheeldon et al., 2013). This is evidenced by, for example, the finding that articulation of sentences with coordinated subject NPs (*The A and the B moved above the C*) is delayed compared with sentences starting with simple subject NPs *The A moved above the B and the C* (e.g. Martin et al., 2010; Roeser et al., 2019; Wagner et al., 2010; Wheeldon et al., 2013). For phrases with subordination (*The A above the B*), the evidence is conflicting, involving either less (Allum & Wheeldon, 2007) or more (Nottbusch et al., 2007) advance planning than coordinated NPs. Lee et al. (2013) found direct evidence that the syntactic organisation of the NP mediates advance planning even when holding the words that appear in the output constant (see also Lee, 2020).

These studies manipulated the syntactic structure of the target phrase – in terms of embedding (Allum & Wheeldon, 2007; Nottbusch et al., 2007), number of head nouns (Martin et al., 2010; Smith & Wheeldon, 1999; Wheeldon et al., 2013), and attachment (Lee et al., 2013). For example, Lee et al. (2013) investigated whether the advance planning of NPs is mediated by syntactic relations between phrase elements. In response to arrays of images, participants generated utterances in the form of instructions to the researcher, such as *Click on the fork of the king that is below the apple.* The image location of either the fork or the king in relation to the image of an apple changed the underlying syntactic structure of the utterance without affecting surface form (ignoring possible variation in prosody, which the authors did not study). Low attachment of the relative clause conveyed that the king is below the apple; high attachment conveyed that the fork was below the apple. The authors found longer onset latencies for high-attaching phrases and argued that this was due to the need for advance planning of the relative clause. This finding suggests that some or all of the dependencies within the NP had to be pre-planned: that is that the syntactic organisation of the initial NP determines whether or not planning extends beyond the first noun.

---

[1]This view does not claim that syntax – expressed in dependencies between phrase elements – is the only determinant of the planning scope. Many other factors are known to influence how much information is pre-planned such as language expertise (Frinsel & Hartsuiker, 2023), working-memory capacity (Klaus & Schriefers, 2018; Swets et al., 2014), context factors such as cognitive load (Ferreira & Swets, 2002; Klaus et al., 2017; Wagner et al., 2010), and production-speed requirements (Damian & Stadthagen-Gonzalez, 2009). These factors are seen as independent of and subordinate to a syntactically determined default scope (Wagner et al., 2010; Wheeldon et al., 2013).

[2]Most previous research has focused on spoken language production (but see Roeser et al., 2019), and for ease of expression in introducing our research we will continue to refer to language producers as speakers and not writers. However in the present context arguments extend to include written production. Our study included both spoken- and written-output conditions.

However, as with other studies involving the production of coordinated and subordinated NPs, Lee et al. (2013) manipulated not only grammatical form but also meaning. Thus, syntactic arrangement and semantic structure were confounded. Findings from these studies are therefore consistent with the hypothesis that the extent of advance planning is driven by syntactic processing. However, they are also consistent with the hypothesis that the extent of advance planning is semantically driven.

Semantic and syntactic processes are known to be closely intertwined during language production (Jackendoff, 1972; Konopka & Brown-Schmidt, 2014; Rizzi, 1997). In particular, contrastive focus – the assignment of prominence to particular elements of a message – influences not only conceptual organisation but often also syntactic structure and prosody. Thus, planning scope may be shaped by both syntactic dependencies and semantic (referential disambiguation)/pragmatic (information structure) factors. Our study builds on this study by examining whether semantic relations alone, in the absence of syntactic structural differences, can affect the scope of pre-utterance planning.

The present experiment removes the semantic/syntactic confound in previous studies by holding syntax and lexemes constant across conditions, but manipulating meaning. We report an image-description experiment in which we manipulated referential contrast (Jackendoff, 1972) to examine the influence of semantic scope of the phrase determiner when pre-planning NPs with nominal modification while keeping syntactic structure and lexical content constant. We specifically controlled those features of the syntax tree that were manipulated in previous research, such as attachment (Lee et al., 2013), number of nouns (Martin et al., 2010; Smith & Wheeldon, 1999; Wheeldon et al., 2013) and embedding (Allum & Wheeldon, 2007; Nottbusch et al., 2007). Our study elicited phrases of the form *The king's magnet*. This can express one of two different contrasts: The phrase could draw attention to one of several possessions of the king – his magnet and not, for example, his butter. The second noun in the phrase disambiguates the referential target from a competitor's possessions. Alternatively the referent could be a specific magnet – the magnet in the possession of the king – and not, for example, the magnet owned by the sailor: There are competing referent objects for the first noun which is required for disambiguation of the possessed object/the second noun. In the first case the semantic scope of the determiner extends over both nouns: *The* determines *magnet* modified by *king*. In the second case the semantic scope of the determiner extends just to the first noun: *The* determines just *king*. However, importantly, there is only one legal syntactic parse for this phrase. This is shown in Figure 1a. Here *The* determines *king*, and the possessing nominal modifier *king's* serves to determine *magnet*. The alternative shown in Figure 1b is not possible both because in this parse *king* lacks a
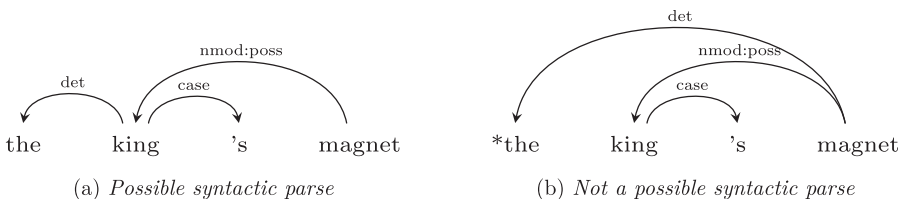

(a) *Possible syntactic parse*   (b) *Not a possible syntactic parse*

Figure 1. Syntactic representations of the phrase 'the king's magnet'.

determiner, which is required, and *magnet* is determined both by *The* and by *king's*, which is not legal. So, at least within the syntactic conventions assumed by the existing sentence production literature that we have cited, *The king's magnet* has only one possible syntactic parse and a single referent, but can communicate one of two distinct semantic contrasts.[3]

The present study tested the hypothesis that pre-utterance planning scope is influenced by semantic organisation of the message – in particular, the presence of contrastive referential relations – even in the absence of syntactic variation (under a specific, strict understanding of syntax as independent of semantics). Prior studies have shown that syntactic dependencies can constrain how much must be planned before utterance onset (e.g. Lee et al., 2013; Wheeldon et al., 2013). However, because syntactic and semantic factors are often confounded, it remains unclear whether semantic structure can, in and of itself, influence planning scope. By manipulating contrastive focus while holding syntactic structure and lexical content constant, we examined whether relational semantics alone can modulate the extent of pre-planning.

We operationalised 'pre-planning' as the deployment of visual attention to referents prior to utterance onset, as indexed by gaze patterns. While this does not necessarily reflect full lexical or syntactic encoding, it plausibly reflects conceptual-level activation of referents (Griffin, 2004). Specifically, we recorded eye movements on stimulus arrays depicting possession with two different levels of contrast, as illustrated in the example above. Eye movement on the stimulus array provides an effective indicator of rapid processes such as message planning (Griffin, 2004; Griffin & Bock, 2000; Meyer & Lethaus, 2004; Swets et al., 2014). Data from eye tracking were necessary in the context of our experiment to understand what information provided by the visual array was used and when it became relevant (Griffin, 2004; Meyer & Lethaus, 2004); eye movements have been used as dependent variable of previous studies that manipulated contrast information (Brown-Schmidt & Hanna, 2011; Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006). We predicted earlier looks to the referent image of the second noun (*magnet*) in wide scope conditions (*The king's magnet* contrasted with *The king's butter*) compared to narrow scope conditions (*The king's magnet* contrasted with *The sailor's magnet*). We also measured time from stimulus display to response-onset. It is possible that differences in onset latency follow the same pattern, with longer onset latencies in the wide scope condition. However – at least in some contexts – pre-syntactic planning has been found not to directly influence planning duration (e.g. Gleitman et al., 2007; Konopka & Meyer, 2014).

1. Method

In a within-subjects experiment, we used image arrays (Figure 2) to elicit noun-plus-modifier phrases[4] in which we manipulated Noun Contrast of the first and the second

---

[3]Our study also elicited postnominal responses of the form *The magnet with the king*. This meets the same conditions of offering only a single syntactic parse but having the potential to communicate two different contrasts.

[4]Most research we cited in the introduction has focused on full sentence production. In contrast, the present study focuses on NPs. This allowed us to simplify pre-planning requirements on a conceptual level

(a) *"The king's magnet" (N1 contrast) / "The magnet with the king"*
(N2 contrast) [and not the sailor's magnet / the magnet with the sailor]

(b) *"The king's magnet" (N2 contrast) / "The magnet with the king"*
(N1 contrast) [and not the king's butter / the butter with the king]

Figure 2. Example stimulus arrays. Participants produced either prenominal or postnominal phrases as indicated, uniquely identifying the circled image. This was preceded by a preview of a similar task with a specific contrast indicated in squared brackets (see Figure 3).

(because of the absence of e.g. verb-argument structure). We are particularly interested in planning variation within the NP because the NP was frequently described as the "default unit" of syntactic pre-planning (see Martin et al., 2010).

noun in the target utterance (henceforth, N1 and N2, respectively), keeping lexical form and syntactic structure constant[5]. We recorded eye movements to the image array. Earlier eye movements towards a contrasting but non-phrase initial image would be evidence that semantic scope affects utterance pre-planning independently of syntactic factors.

To eliminate alternative explanations, we incorporated two additional factors. We elicited NPs with modifiers that either preceded the head noun (a *prenominal* modifier; *The king's magnet*) or followed the head noun (a *postnominal* modifier; e.g., *The magnet with the king*). By allowing participants to freely produce NPs with modifier location, we were able to eliminate alternative explanations related to head position (Allum & Wheeldon, 2007) and stimulus array: Head-initial phrases may allow planning of the modifier to be postponed until after production onset. A phrase-initial modifier, on the other hand, may require pre-planning across the head-noun. This was important because there is evidence that pre-planning takes into account the structural function of the phrase-initial noun (Brown-Schmidt & Konopka, 2015 but see Allum & Wheeldon, 2007).

Participants completed separate blocks of trials in which they responded by speaking – the typical output modality in previous research – and in writing (by typing on a computer keyboard). This was for the following reason: Almost all existing research on sentence production has focused on spoken utterances.[6] Comparing typed and spoken responses therefore allows us to generalise beyond one output modality: Some previous studies with spoken output might have wrongly interpreted effects related to, for example, fluency requirements as evidence for fundamental (modality-independent) pre-planning requirements (see Roeser et al., 2019). This is because in spoken production there is a pressure for fluency, whereas hesitations during written output have no communicational implications (e.g. Clark & Fox Tree, 2002).
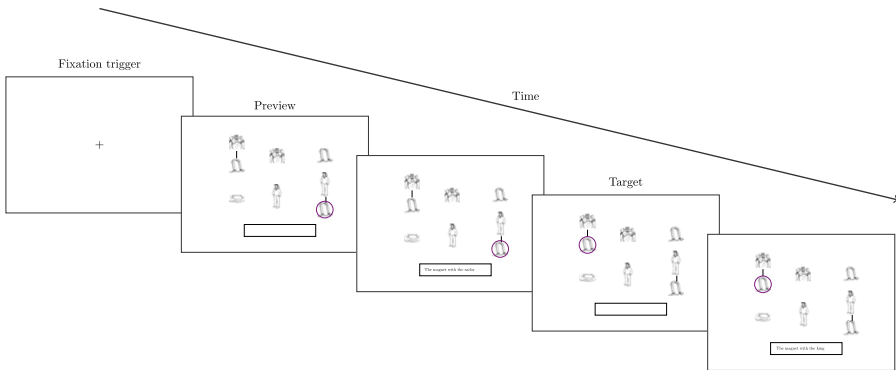
1.1. Participants

Sample size was set a priori at 64 participants. Psychology students (57 female, 7 male, mean age = 20.4 years, SD = 5.4, range: 18–50) participated as part of a research-reward scheme. All participants were self-reported native speakers of British English, free of linguistic impairments, and had normal or corrected-to-normal vision. Five participants were replaced because of failure to follow task instructions or failure to provide well calibrated eye data.

1.2. Design and materials

Participants were presented with arrays containing two image pairs and four other images; see Figure 2. In each trial, participants were presented with an array with six cells. Image pairs were linked by a vertical line with the target image always appearing below its modifier. Participants were asked to produce a phrase that identified the highlighted

---

[5]Specifically, we ensured that the dependency structure between the determiner, head noun, and modifier remained identical across conditions, so the surface syntactic form was held constant. However, we acknowledge that contrastive focus operates at the syntax-semantics interface and influences other aspects of language planning, such as prosody and information structure.

[6]But see Damian and Stadthagen-Gonzalez (2009), Roeser et al. (2019), Nottbusch et al. (2007), Nottbusch (2010) and Torrance and Nottbusch (2012).

Figure 3. Example trial. Participants responded to a preview screen prior to producing an utterance for the target screen to increase their awareness of the semantic contrast. The example trial illustrates a N2 contrast condition if the participant used a postnominal phrase ('The magnet with the king' [and not the magnet with the sailor]) and a N1 contrast condition when the participant used a prenominal phrase ('The king's magnet' [and not the sailor's magnet]). For an example array with N1 contrast in prenominal phrases and N2 contrast in postnominal phrases, see Figure 2.

image. Their response acted as instruction to the experimenter, who then clicked the phrase referent. The participant then saw the image that the experimenter had clicked. Participants were free to produce phrases with either prenominal or postnominal modification. Examples at the start of the experiment were provided for each of these options. As can be seen from Figure 2, contrast was dependent on choice of phrase structure.

To increase the salience of contrast and modifier, each target trial was preceded by a preview. This is visualised in Figure 3. In both the preview and the target trial, participants saw the same array. In the preview, participants produced an NP identifying one object (*The sailor's magnet/The magnet with the sailor* in Figure 3). In the target trial, they were required to identify a different image. The first (N1) and second noun (N2) in the target utterance served the same contrastive functions in both preview and target and so NPs required to express preview and target were semantically isomorphic (and syntactically identical). This preview-target setup increased the participants' awareness of the semantic contrast in the target referent. This is because in the preview participants had to encode the image pair (sailor-magnet in Figure 2a; king-butter in Figure 2b) that contained the semantic contrast for target phrase (i.e. king-magnet). For analysis, we used only preview-target pairs with the same NP structure (see Results section).

The remaining four cells of the stimulus arrays shown in Figure 2 were filled with four images that appeared in the target and the preview pair. This ensured that unique identification of the target object always required a specific noun-plus-modifier phrase.

Each participant completed a written and a spoken session. These manipulations rendered a N1 contrast/N2 contrast × prenominal/postnominal × written/spoken design.

Thirty-two arrays were created[7]. Black and white drawings were taken from the database of the International Picture Naming Project (Bates et al., 2003; Székely et al.,

---

[7]See OSF repository: https://osf.io/wst3a

2003, 2004, 2005). Items were counterbalanced over four Latin Square lists (Contrast, Modality Type [speech, writing]). Order of modality session was counterbalanced between participants; half of the participants started with a written session and the other half with a spoken session. The location of preview and target pairs was randomised within each Latin square list. Two filler lists, each of 16 trials, were created. One filler list was used in the first session and the other in the second session. Fillers elicited structures that were different from those referred to in target trials to avoid strategic use of descriptions and anticipation of the target pair. For example, some fillers prompted the top image of image pairs, and colour (in combination with modifiers) was used for disambiguation (e.g. *the cat's green ball*). Filler lists and stimuli were presented in randomised order within sessions. In total, each participant saw 64 arrays, 32 per modality (50% fillers).

To reduce the lexical variability in the response, we used pictures that were easy to name according to existing naming norms provided by the International Picture Naming Project. Summary statistics of picture naming norms are given in Appendix A (Table A1).

We recorded eye gaze to the stimulus array. To assess the time course of planning the second noun in the utterance relative to production onset and the likelihood of pre-onset planning of both noun, we calculated two variables: (1) time relative to production onset when gaze shifted from the image of N1 to N2; (2) proportion of trials for which this gaze shift happened before production onset. Gaze shift from the N1 referent to N2 was defined as the first fixation on the image of N2 after the gaze left N1 (fixation required a dwell time of at least 100 ms). These measures estimate the attention shift from the referent of the first noun to the preparation of the second noun (see e.g. Griffin, 2004; Griffin & Bock, 2000; Meyer & Lethaus, 2004; Roeser et al., 2019). AOIs were circular with a diameter equal to the height of the images.

1.3. Apparatus

Eye movements were recorded using an SR Research EyeLink 1000 desk-mounted eye tracker used in remote mode. Eye data were sampled at 500 Hz from just the right eye. The participant was seated 55 to 60 cm away from the lens. The experiment was built in SR Research Experiment Builder. Stimuli were displayed on a 19' ViewSonic Graphic Series (G90fB) CRT monitor with a screen resolution of 1,280 × 1,024 pixels and 85 Hz refresh rate using an Intel Core 2 PC. A Logitec headset and ASIO audio driver supported by a Creative SB X-Fi sound card were used for recording spoken output. An additional display and a mouse were used to allow experimenter response.

1.4. Procedure

Every experiment started with camera set-up, and 9-point calibration and validation. Participant and experimenter were seated at different screens unable to see each other's screen. Each preview/target trial pair was initiated by fixating an ellipse at the centre of the screen (21 × 21 pixels) for 200 ms. Recalibration was performed if this fixation trigger failed. The stimulus array then appeared, with a violet circle highlighting the image that was the target of the preview trial. The participants spoke or wrote NPs intended to uniquely identify the image, and then pressed enter. This activated mouse input for the experimenter who then clicked on the image identified

by the participant's utterance. For ambiguous utterances, the experimenter clicked a plausible referent image. Mouse clicks on the correct image prompted a green circle as feedback. Clicks on an incorrect image prompted a red circle. Feedback circles were shown for 250 ms to encourage participants to use unambiguous descriptions. The second image (the target) was highlighted immediately after the feedback circle disappeared. The target trial then proceeded in the same way as the preview trial. In the written condition participants were instructed to position their fingers over the keyboard at the start of the trial and to respond as quickly and accurately as soon as possible after appearance of the stimulus array. Participants were familiarised with the experimental task in 10 practice trials. The images, size 150 × 150 pixels, appeared around the centre of the screen. In the written condition participants gave their response in a text box (896 × 50 pixels) that appeared in the centre of the screen. Participants could pause, if they wished, between trial pairs. The duration of the experiment was approximately 45 min.

1.5. Data analysis

We used Bayesian mixed-effects models (Gelman et al., 2014) for statistical inference.[8] All models were fitted with maximal random effects structure, that is random intercepts for participants and items and by-participant and by-item slope adjustments for all predictors (Baayen et al., 2008; Barr et al., 2013; Bates et al., 2015). Predictors were sum-coded (Brehm & Alday, 2022; Schad et al., 2020) and reported as posterior estimates with 95% probability interval (95% PI[9]).

To assess the strength of support for the effects of interest over the null hypothesis, Bayes Factors (BF) were calculated using the Savage-Dickey method (see Dickey et al., 1970; Nicenboim & Vasishth, 2016; Wagenmakers et al., 2010). BF = 2, for example, indicates that the alternative hypothesis is twice as likely compared to the null hypothesis given the data. BFs >3 indicate weak evidence, >8 moderate evidence, and >32 strong evidence for a statistically meaningful effect (see e.g. Baguley, 2012; Jeffreys, 1961; Lee & Wagenmakers, 2014).

2. Results

Prior to analysis, we removed trials in which the participant paused for > 2.5 secs at any point after production onset, trials where they self-corrected, and trials where the response did not unambiguously reference the target image (9.9% of trials in total). Trials in which participants gave different modifier positions in response to preview and target (5.6%) were also excluded, as were trials with onset latencies > 10 s (0.3%). Trials in which participants produced an utterance that contained neither prenominal nor postnominal modification were removed (speech: 0.9%, writing: 3.4%). For

---

[8]Models were fitted in rstanarm (Goodrich et al., 2020) with weakly informative priors (McElreath, 2016; Sorensen et al., 2016) and run with 2,000 iterations on 4 chains with a warm-up of 1,000 iterations and no thinning. Model convergence was confirmed by the Rubin-Gelman statistic ($\hat{R} = 1$; Gelman & Rubin, 1992) and inspection of the Markov chain Monte Carlo chains. R-scripts and data are available on OSF: https://osf.io/wst3a

[9]Probability intervals are sometimes referred to as credible intervals. Unlike confidence intervals they have a straightforward interpretation as the range that contains the true parameter value with a 95% probability (Kruschke, 2014; Morey et al., 2016).

Table 1. Modifier choice

Contrast	Writing		Speech	
	Mean with SE	Est. with 95% PI	Mean with SE	Est. with 95% PI
N1	.55 (.02)	.37 [.07, .79]	.69 (.02)	.96 [.70, 1.00]
N2	.58 (.02)	.44 [.08, .85]	.71 (.02)	.97 [.74, 1.00]

Note. Proportion of target trials in which participants produced a postnominal modifier ('The king with the magnet') instead of a prenominal modifier ('The king's magnet') represented as model estimate with 95% PIs in brackets and empirical mean with standard error (SE).

analysis of eye data, we removed a further 2.4% of the trials which had more than 75% of the eye fixations away from any of the images in the array, before production onset.

As can be seen from Table 1, participant choice of modifier-position (prenominal or postnominal) was independent of Contrast: A full-factorial Bayesian logistic mixed-effects model found negligible evidence for a Contrast effect on Modifier-position choice (BF = 1.3) or any by-Contrast interactions (BFs < 0.6). Participants were roughly equally likely to choose prenominal and postnominal constructions in the written modality, but showed a greater tendency towards postnominals when speaking (BF > 100 for the main effect of modality). Model coefficients are summarised in Appendix B (Table B1).
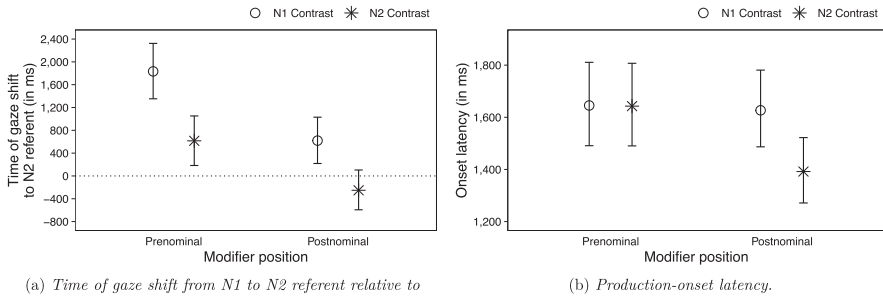
We established the effect of Contrast on each of three dependent variables. From the eye data, we calculated the time relative to production onset when gaze shifted from the image representing the first noun to the image representing the second noun. We interpreted this as an indication of attention shift from the first to the second referent. We then evaluated the probability that this shift occurred prior to the participant starting to output the phrase. For completeness, we provide the posterior of the fixation time-course for the eye data in Appendix C for the time window prior to production onset. Last, we calculated onset latency – time from appearance of the circle prompt to output onset.

Gaze shift and onset latency were log-transformed; probability of gaze shift before output onset was analysed with a logistic model. All models were fitted with main effects and all interactions of Modality (written, spoken), preview/target, Modifier position (prenominal, postnominal), and Contrast (N1, N2). We then conducted a series of eight planned comparisons, based on posteriors from these models, to test for evidence of advance planning in the N2 contrast condition, relative to N1 contrast within each cell of the Modality by Modifier position by preview/target interaction.

We present findings from the overall model first. Then, for clarity, we report effects in just target trials, first in the written modality, then in the spoken modality. An overview of all fixed effects coefficients can be found in Appendix D (Table D1).[10]

Findings from the main model were as follows: We found strong evidence for a main effect of Contrast (BF > 100) showing earlier gaze divergence for N2 contrast. The evidence for all by-Contrast interactions was at best very weak (BF < 2). The analysis of the probability that gaze shift to the N2 referent occurred before production onset provided moderate evidence of a Contrast by Modifier-position by Modality interaction (BF = 8.5) and strong evidence of a Contrast by preview/target interaction (BF > 100). These effects are unpacked below. In response onset latency,

---

[10]Effects of the preview are summarised in Appendix E.

(a) *Time of gaze shift from N1 to N2 referent relative to production onset (dotted line). Negative values are prior to onset.*

(b) *Production-onset latency.*

Figure 4. Written production results. Estimates from full-factorial Modality × preview/target × Modifier position × Contrast Bayesian mixed-effects models. Error bars represent 95% PIs.

we found strong evidence for a main effect of Contrast (BF = 68), and substantial support for a Modifier-position by Contrast interaction (BF = 3.5). There was little or no evidence for other effects (BF < 2) across any of the three dependent variables.

2.1. Written output modality

Estimated mean time of gaze shift from the N1-referent to the N2-referent in the written condition, in Figure 4a, shows earlier gaze on N2 under N2 contrast for both modifier types (evidence for effect of contrast: prenominal BF = 15; postnominal BF = 23). In fact, for postnominal modifiers gaze shift in the N2 contrast condition tended to occur before production onset but not in the N1 contrast condition. For prenominal modifiers, gaze shift was observed after production onset but earlier for the N2 contrast condition.
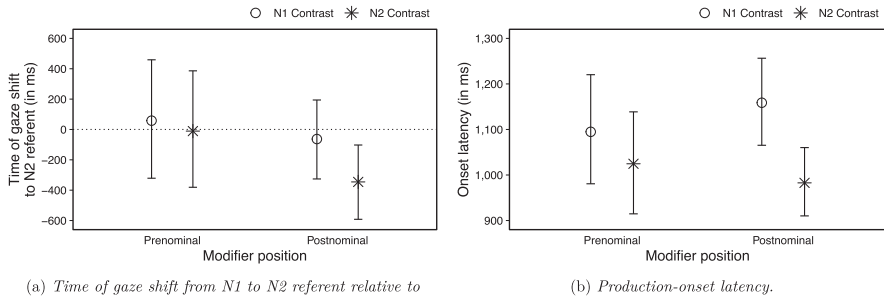
The probability of gaze shift to occur before production onset followed the same pattern. Estimated probability was greater in the N2 contrast condition for both modifier types (prenominal: N1-contrast = .19, 95% PI [.13, .26], N2-contrast = .39, 95% PI [.31, .5], BF = 95; postnominal: N1-contrast = .27, 95% PI [.19, .35], N2-contrast = .55, 95% PI [.45, .65], BF > 100).

The inferred cell means for the onset latency are visualised in Figure 4b. The results show that production onset latency was shorter under N2 contrast for postnominal modifiers (BF = 12) but not for prenominal modifiers (BF < 0.05).

2.2. Spoken output modality

Estimated time of gaze shift from the N1 to the N2 referent in the spoken condition is shown in Figure 5a. Although parameter estimates suggest earlier gaze-shift for N2 contrast for postnominal modifiers, we found evidence for no effect in either modifier position (BF < 0.08).

We found strong evidence indicating that the probability of pre-onset gaze-shift from N1 to N2 referent was greater in the N2-contrast condition in postnominals (N1-contrast: .51, 95% PI [.42, .6], N2-contrast: .81, 95% PI [.74, .87], BF > 100) but not in prenominals (N1 contrast: .4, 95% PI [.29, .52], N2 contrast: .5, 95% PI [.38, .63], BF < 0.8). These findings therefore reproduce Contrast effects found in written production for phrases with postnominal modifiers but not for prenominals.

(a) *Time of gaze shift from N1 to N2 referent relative to production onset (dotted line). Negative values are prior to onset.*

(b) *Production-onset latency.*

Figure 5. Spoken production results. Estimates from full-factorial Modality × preview/target × Modifier position × Contrast mixed-effects models. Error bars represent 95% PIs.

As can be seen from Figure 5b, we found strong evidence that production onset occurred sooner under N2 contrast when participants produced phrases with post-nominal modifiers (BF = 26) but not when modifiers were prenominal (BF = 0.11). This is the same pattern as was found for written production.

3. Discussion

Our results support the hypothesis that conceptual-level semantic structure, particularly contrastive focus, affects scope of planning prior to utterance onset even in the absence of morpho-lexical and syntactic variation. Participants in our study directed attention to the second noun earlier when it served a contrastive function in utterances/texts in which surface syntax and word-forms were identical. Our results therefore demonstrate that contrastive focus – typically viewed as operating at the syntax-semantics interface – can impact pre-utterance planning independently of syntactic differences. These findings extend prior study showing syntactic effects on planning (e.g. Lee et al., 2013) by demonstrating that semantic relations can drive early planning even in the absence of structural variation. Our findings clearly are not evidence against syntactic structure playing a role in shaping planning scope. They are, however, evidence against planning scope determination being exclusively driven by syntax. They also point to the possibility that the previous studies in which researchers have claimed syntax as the driver for variation in planning scope might in fact be observing effects of semantic structure.

Before we discuss the implications of these findings, we want to address three potential caveats. First, we did not find evidence that response latency was longer in N2 contrast conditions. Although this is perhaps surprising, it is consistent with findings of some previous studies (Griffin, 2004; Swets et al., 2014; but see Ferreira & Engelhardt, 2006; Fodor, 1983; Kempen & Hoenkamp, 1987). In the postnominal condition, response latency was in fact reliably *shorter* where N2 was contrastive. This is likely to have been because the lexical name of the phrase-initial noun was pre-activated in the N2 contrast condition (in contrast to pre-activation of the N2 referent for N1 Contrast), allowing for quick response onsets. That we observed no such speed-up effect in the prenominal condition is likely because for prenominals, both nouns had to be lexically retrieved before production onset. Indeed, the time-course data in Figure C1 suggest that in the prenominal condition but not in the postnominal

condition, participants attended to both phrase referents before production onset. Therefore, as both nouns in the prenominal condition had to be pre-planned, there was no timing advantage to be gained from a pre-activation of either noun. In contrast, that we observe the speed-up effect for onset latencies in postnominal phrases suggests that earlier looks to the N2 referent did not cascade into lexical retrieval, as lexical retrieval is known to be time-consuming (Griffin, 2001; Konopka & Bock, 2009; Wheeldon et al., 2013). Instead, it is possible that the postnominal modifier noun was retrieved after production onset, in parallel to production (Griffin, 2003).

Second, reviewers of a previous version of this article suggested that priming from the preview task may explain the contrast effect. This is plausible, although in principle there is no reason why pre-activation of the first noun should lead to pre-planning of the second noun. Some research suggests that pre-activation leads to faster speech onsets but not necessarily more advance planning (Hardy et al., 2020; Smith & Wheeldon, 1999; Wheeldon et al., 2011, 2013). Our preview task was identical in form to the experimental task. Therefore, a direct test of whether priming explains our eye movement findings would be to establish whether or not the same effects are present during preview. This analysis can be found in Appendix E: We again found higher probability of pre-onset gaze shifts under N2 Contrast in phrases with postnominal modification (N1 Contrast = .65, 95% PI [.56, .74], N2 Contrast = .78, 95% PI [.71, .85], BF = 6.78) although not in phrases with prenominal modification (N1 Contrast = .45, 95% PI [.33, .58], N2 Contrast = .5, 95% PI [.36, .63], BF < 0.4). This effect cannot be explained on the basis of lexical pre-activation.

Third, although throughout this article we have talked in terms of planning in advance of production onset, gaze shift to N2 in the written condition tended, in both contrast conditions, to be after the participant's first keypress of the determiner. In the spoken condition, gaze shift was earlier, and tended to be at times equal to or earlier than determiner onset. Note that in all cases, the determiner was the same word (*The*) and therefore entirely predictable. Written output is both slower and unlike speech is not associated with expectation for output fluency (e.g. Clark & Fox Tree, 2002). This means both that there is more time available for planning in parallel with output (Damian & Stadthagen-Gonzalez, 2009; Roeser et al., 2025), and that there is minimal pressure for participants to avoid hesitation between the determiner and N1. In contrast, speakers must balance between mental buffering of pre-planned items and intra-utterance pausing (Levelt & Meyer, 2000; Meyer, 1997). To reduce pre-planning demands, and to maintain output fluency, speakers may assess the possibility of lexical retrieval in parallel with articulation (Griffin, 2003), and then delay output until the determiner and N1 can be produced without an intervening hesitation. This is particularly likely to be the case in the present study, given that the person to whom they were communicating was physically present.[11]

---

[11]This account is consistent with participants' preference for postnominal phrase, when speaking (Allum & Wheeldon, 2007) but not when writing. To reduce pre-planning demands participants may have preferred postnominal modification in speech. Indeed further analysis suggested that, for speech, modifier choice (pre vs. postnominal) was indeed reliably larger than what we would expect if either structure was equally likely (95% PI[.7, .9], BF = 43.6). No such preference was found in writing (95% PI[.1, .8], BF = 0.9). In contrast, for colour adjectives Gagné et al. (2020) found a preference for prenominal modifiers (*the blue dog* vs. *the dog that is blue*).

Our findings, therefore, present a relatively complex picture of the mechanisms underlying phrase planning in writing and in speech. However, taken together, we believe they provide clear evidence that, at least in some contexts, semantic dependency alone determines the minimum obligatory unit of pre-planning: Meaning-level representation of the referent impacts the size of the minimum linguistic unit that must be prepared prior to output-onset. This conclusion is non-trivial. Under a lexical-incremental understanding of pre-planning (for review see Bock, 1987; Bock & Ferreira, 2014; Wheeldon & Konopka, 2023), in order to start language output, only the initial referent needs to be identified and lexically retrieved. Any additional planning introduced by the contrast manipulation could in principle be postponed until after production onset. Under a hierarchical view (Dell & O'Seaghdha, 1992; Ferreira & Dell, 2000; Garrett, 1975), the order of planning is syntactically determined. If planning has a phrasal scope that determines the order in which nouns are planned (but not necessarily the timing), we would expect no difference across contrast conditions. This is, however, not what we found. Contrast modulated attention allocation to a non-sentence initial noun referent before production onset. This is evidence that the pre-planning demands obligated by the language production system can be determined by semantic features of the message. This is in line with Konopka and Meyer (2014), who proposed a horse-race process in which, on the semantic level, both a relational or non-relational route are available to guide planning.

The convergence of results across spoken and written production reinforces the view that semantic structure can independently shape early stages of language production. Our central findings – that semantic contrast influences the scope of pre-utterance planning – were replicated across both modalities. We believe, therefore, that the observed effects reflect a modality-general planning mechanism, operating at the level of message formulation or conceptual encoding, rather than modality-specific articulation or motor execution. While writing introduces variability in typing fluency and error correction, these differences are unlikely to affect the conceptual planning processes that drive gaze allocation and referent access prior to output (at least in the context of our experiment, but for an alternative view, see Rønneberg et al., 2022).

Our overall conclusion is consistent with three different ways of understanding the role of syntax in determining obligatory planning scope in phrase (and presumably also sentence) production. In, we think, all previous studies that have argued that planning must scope over the initial verb-argument NP, dependencies within the semantics of the intended message provide a potential alternative explanation. Minimally, we have demonstrated that semantic relationships contribute to planning scope even when syntactic structure – understood in terms of Universal Dependency Grammar or similar formalisms that have been assumed in previous research was held constant. Under Dependency Grammar, syntactic structure was constant across the contrast conditions in our study. Under a Generative Grammar analysis and for other syntacto-semantic formalisms that explicitly encode information that is interpreted semantically, that would not be the case. An alternative framing of our findings, therefore, is that the structures that determine planning scope are not captured by a strongly meaning-independent parse of the target utterance structure.

Following from this, there are two stronger interpretations that our findings are consistent with, though do not provide direct support for; both of which represent worthwhile targets for future research. First, it is possible that the language

production system makes use of meaning-independent syntactic representations but that these are not implicated in determining planning scope. Second, and most strongly, it might be that representations of syntax (as distinct from syntacto-semantics) play no role in planning spoken or written production. These positions are consistent with recent arguments against abstract representations of syntax playing a role in language production (Nielsen & Christiansen, 2025)[12]. Also, these interpretations are consistent with various forms of construction grammar which assume that speakers (and writers) build mental representations that do not separate semantic and syntax representations (see Croft & Cruse, 2004; Fried & Nikiforidou, 2025) and large language models that do not represent syntax at all (see Mahowald et al., 2024; Piantadosi, 2024; Ramscar, 2021).

Finally, we acknowledge two potential confounding factors raised by reviewers of a previous version of this article. First, our contrast manipulation is associated with prosodic variations. In English, contrast can be marked prosodically (Selkirk, 1995)[13]. Although, in the written modality, it was not possible to use prosody to convey referential contrast, participants might still have implicitly prepared the prosody of the utterance. Therefore, our findings could in principle be the result of a need to encode surface-level phonological representations (Fuchs et al., 2013; Fuchs & Krivokapić, 2016). We believe that it is relatively unlikely that the effects of semantic dependency – similar to the syntactic planning-scope effects we cited in the introduction[14] – were entirely mediated by prosody. To rule out a phonological interpretation, future research would need to assess to what extent planning-scope effects can be reproduced in contexts where phonological/prosodical planning was experimentally controlled; for example, by using articulatory suppression (repeatedly saying, the, the, e.g. Richardson & Baddeley, 1975) while preparing and executing utterances in writing.

Second, our task conditions minimised syntactic variability, including the use of repeated phrase structures and a focus on phrase planning rather than full sentence production. Specifically, our design used a preview/target structure in which the syntactic frame of the target phrase was highly predictable, and in the vast majority of trials (94.4%), participants reused the same syntactic form as used in the preview trial. Structural repetition is likely to have minimised the need for active syntactic planning, and may not reflect naturalistic language production. This was, however, intentional: by reducing syntactic variability, we were able to examine whether semantic structure, in the absence of syntactic uncertainty, influences planning scope. Specifically, we tested whether semantics is sufficient to drive pre-planning when syntax is controlled.

In summary, we report evidence that demonstrates, for the first time, variation in whether or not a written or spoken phrase is pre-planned beyond the first noun just as a result of semantic dependency, in a context where all other features of the target production (with the exception of prosody in the spoken modality) are held constant. We found that the semantic organisation of the phrase affects pre-planning beyond

---

[12]The idea that syntactic representations are not involve in determining planning scope is not new to the psycholinguistic literature; for a discussion see section 3 of F. Ferreira and Engelhardt (2006).

[13]Also particles such as *only*, *even* and *too* that indicate contrastivity of a linguistic element with contextually available alternatives within the particle's scope (Paterson et al., 2003; Rooth, 1992).

[14]Different syntactic structures (as in e.g. Allum & Wheeldon, 2007; Lee et al., 2013; Smith & Wheeldon, 1999; Wheeldon et al., 2013) are inherently associated with different prosodic contours.

the first noun. While prior study proposed that syntactic factors constrain the scope of pre-planning (e.g. Lee et al., 2013; Wagner et al., 2010; Wheeldon et al., 2013), our findings suggest that semantic factors such as contrastive focus can drive early referent access, even when syntactic and lexical structure are controlled.

Open Practices Statement. The data and materials for all experiments are available at https://osf.io/wst3a (Roeser et al., 2022). The experiment was not preregistered.

References

Allum, P. H., & Wheeldon, L. R. (2007). Planning scope in spoken sentence production: The role of grammatical units. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(4), 791–810.

Allum, P. H., & Wheeldon, L. R. (2009). Scope of lexical access in spoken sentence production: Implications for the conceptual-syntactic interface. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(5), 1240–1255.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4), 390–412.

Baguley, T. (2012). *Serious stats: A guide to advanced statistics for the behavioral sciences*. Palgrave Macmillan.

Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59(4), 457–474.

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

Bates, E., D'Amico, S., Jacobsen, T., Székely, A., Andonova, E., Devescovi, A., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., *et al.* (2003). Timed picture naming in seven languages. *Psychonomic Bulletin & Review*, 10(2), 344–380.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. Preprint, arXiv: 1506.04967.

Bock, J. K. (1987). Coordinating words and syntax in speech plans. In A. Ellis (Ed.), *Progress in the psychology of language* (Vol. 3, pp. 337–390). Erlbaum.

Bock, J. K., & Ferreira, V. S. (2014). Syntactically speaking. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 21–46). Oxford University Press.

Brehm, L., & Alday, P. M. (2022). Contrast coding choices in a decade of mixed models. *Journal of Memory and Language*, 125, 104334.

Brown-Schmidt, S., & Hanna, J. E. (2011). Talking in another person's shoes: Incremental perspective-taking in language processing. *Dialogue & Discourse*, 2(1), 11–33.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, 109(2), 274–280.

Brown-Schmidt, S., & Konopka, A. E. (2015). Processes of incremental message planning during conversation. *Psychonomic Bulletin & Review*, 22(3), 833–843.

Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54(4), 592–609.

Chang, F., Dell, G. S., & Bock, J. K. (2006). Becoming syntactic. *Psychological Review*, 113(2), 234–272.

Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition*, 84, 73–111.

Croft, W., & Cruse, D. A. (2004). An overview of construction grammars. In *Cognitive linguistics* (pp. 257–290). Cambridge University Press.

Damian, M. F., & Stadthagen-Gonzalez, H. (2009). Advance planning of form properties in the written production of single and multiple words. *Language and Cognitive Processes*, 24(4), 555–579.

Dell, G. S., & O'Seaghdha, P. G. (1992). Stages of lexical access in language production. *Cognition*, 42, 287–314.

Dickey, J. M., Lientz, B. P., *et al.* (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov chain. *The Annals of Mathematical Statistics*, 41(1), 214–226.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development*, 59(5) i–185.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive Psychology*, 40(4), 296–340.

Ferreira, F., & Engelhardt, P. E. (2006). Syntax and production. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (Vol. 2, pp. 61–91). Elsevier. https://doi.org/10.1016/B978-0-12-369374-7.X5000-7.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57–84.

Fodor, J. A. (1983). *The modularity of mind*. MIT Press.

Fried, M., & Nikiforidou, K. (2025). *The Cambridge handbook of construction grammar*. Cambridge University Press.

Frinsel, F. F., & Hartsuiker, R. (2023). Planning scope in second language sentence production: Do bilingual speakers plan ahead more in L1 than in L2? *Bilingualism*. *Language and Cognition*, 26(4), 684–694. https://doi.org/10.1017/S1366728923000068.

Fuchs, S., & Krivokapić, J. (2016). Prosodic boundaries in writing: Evidence from a keystroke analysis. *Frontiers in Psychology*, 7, 1678.

Fuchs, S., Petrone, C., Krivokapić, J., & Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics*, 41(1), 29–47.

Gagné, C. L., Spalding, T. L., Burry, J. C., & Adams, J. T. (2020). Production of multiword referential phrases: Inclusion of over-specifying information and a preference for modifier-noun phrases. In S. S. Walde & E. Smolka (Eds.), *The role of constituents in multiword expressions: An interdisciplinary, cross-lingual perspective* (Vol. 4, pp. 155–178). Language Science Press.

Garrett, M. F. (1975). Levels of processing in sentence production. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 9, pp. 133–177). Academic Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis*. (3rd ed.). Chapman; Hall/CRC.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4), 457–472.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57(4), 544–569.

Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). Rstanarm: Bayesian applied regression modeling via Stan. [R package version 2.19.3]. https://mc-stan.org/rstanarm

Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition*, 82(1), B1–B14.

Griffin, Z. M. (2003). A reversed word length effect in coordinating the preparation and articulation of words in speaking. *Psychonomic Bulletin & Review*, 10(3), 603–609.

Griffin, Z. M. (2004). Why look? Reasons for eye movements related to language production. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 213–247). Psychology Press.

Griffin, Z. M., & Bock, J. K. (2000). What the eyes say about speaking. *Psychological Science*, 11(4), 274–279.

Hardy, S. M., Segaert, K., & Wheeldon, L. (2020). Healthy aging and sentence production: Disrupted lexical access in the context of intact syntactic planning. *Frontiers in Psychology*, 11, 257.

Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. MIT Press.

Jeffreys, H. (1961). *The theory of probability* (Vol. 3). Oxford University Press, Clarendon Press.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11(2), 201–258.

Klaus, J., Mädebach, A., Oppermann, F., & Jescheniak, J. D. (2017). Planning sentences while doing other things at the same time: Effects of concurrent verbal and visuospatial working memory load. *Quarterly Journal of Experimental Psychology*, 70(4), 811–831. https://doi.org/10.1080/17470218.2016.1167926.

Klaus, J., & Schriefers, H. (2018). An investigation of the role of working memory capacity and naming speed in phonological advance planning in language production. *The Mental Lexicon*, 13(2), 159–185. https://doi.org/10.1075/ml.18009.kla.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language*, 66, 143–162.

Konopka, A. E., & Bock, J. K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58(1), 68–101.

Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. In M. Goldrick, V. S. Ferreira, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 3–20). Oxford University Press.

Konopka, A. E., & Kuchinsky, S. E. (2015). How message similarity shapes the timecourse of sentence formulation. *Journal of Memory and Language*, 84, 1–23.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, 73, 1–40.

Kruschke, J. K. (2014). *Doing bayesian data analysis: A tutorial with R, JAGS, and Stan*. (2nd ed.). Academic Press.

Kuchinsky, S. E. (2009). *From seeing to saying: Perceiving, planning, producing* [Doctoral dissertation, University of Illinois].

Kuchinsky, S. E., Bock, J. K., & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(3), 748–756.

Lee, E.-K. (2020). Look-ahead planning of lexical heads in adjunct-head structures. *Lingua*, 244, 1–10.

Lee, E.-K., Brown-Schmidt, S., & Watson, D. G. (2013). Ways of looking ahead: Hierarchical planning in language production. *Cognition*, 129(3), 544–562.

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Levelt, W. J. M., & Meyer, A. S. (2000). Word for word: Multiple lexical access in speech production. *European Journal of Cognitive Psychology*, 12(4), 433–452.

Mahowald, K., Ivanova, A. A., Blank, I. A., Kanwisher, N., Tenenbaum, J. B., & Fedorenko, E. (2024). Dissociating language and thought in large language models. *Trends in Cognitive Sciences*, 28(6), 517–540.

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., II, & Yang, C.-L. (2010). Planning in sentence production: Evidence for the phrase as a default planning scope. *Cognition*, 116(2), 177–192.

Martin, R. C., Yan, H., & Schnur, T. T. (2014). Working memory and planning during sentence production. *Acta Psychologica*, 152, 120–132.

McElreath, R. (2016). *Statistical rethinking: A Bayesian course with examples in R and Stan*. CRC Press.

Meyer, A. S. (1997). Conceptual influences on grammatical planning units. *Language and Cognitive Processes*, 12(5–6), 859–864.

Meyer, A. S., & Lethaus, F. (2004). The use of eye tracking in studies of sentence generation. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 191–211). Psychology Press.

Mirman, D., Dixon, J. A., & Magnuson, J. S. (2008). Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory and Language*, 59(4), 475–494.

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, 23(1), 103–123.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas—Part II. *Language and Linguistics Compass*, 10(11), 591–613.

Nielsen, Y. A., & Christiansen, M. H. (2025). Context, not grammar, is key to structural priming. *Trends in Cognitive Sciences*. https://www.cell.com/trends/cognitivesciences/abstract/S1364-6613(25)00143-3

Nottbusch, G. (2010). Grammatical planning, execution, and control in written sentence production. *Reading and Writing*, 23(7), 777–801.

Nottbusch, G., Weingarten, R., & Sahel, S. (2007). From written word to written sentence production. In M. Torrance, L. van Waes, & D. W. Galbraith (Eds.), *Writing and cognition: Research and applications* (Vol. 20, pp. 31–53). Elsevier.

Paterson, K. B., Liversedge, S. P., Rowland, C., & Filik, R. (2003). Children's comprehension of sentences with focus particles. *Cognition*, 89(3), 263–294.

Piantadosi, S. T. (2024). Modern language models refute Chomsky–s approach to language. In E. Gibson and M. Poliak (Eds.), *From fieldwork to linguistic theory: A tribute to Dan Everett* (Empirically Oriented Theoretical Morphology and Syntax, 15). Berlin: Language Science Press.

Ramscar, M. (2021). A discriminative account of the learning, representation and processing of inflection systems. *Language, Cognition and Neuroscience*, 38(4), 446–470.

Richardson, J., & Baddeley, A. (1975). The effect of articulatory suppression in free recall. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 623–629.

Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar* (pp. 281–337). Kluwer.

Roeser, J., Conijn, R., Chukharev, E., Ofstad, G. H., & Torrance, M. (2025). Typing in tandem: Language planning in multisentence text production is fundamentally parallel. *Journal of Experimental Psychology: General*, 154(7), 1824–1854. https://doi.org/10.1037/xge0001759.

Roeser, J., Torrance, M., & Baguley, T. (2019). Advance planning in written and spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 45(11), 1983–2009. https://doi.org/10.1037/xlm0000685.

Roeser, J., Torrance, M., & Baguley, T. (2022). Planning contrast ahead: Contrast structure guides pre-planning in complex noun phrases [OSF dataset]. https://doi.org/10.17605/OSF.IO/WST3A

Rønneberg, V., Torrance, M., Uppstad, P. H., & Johansson, C. (2022). The process-disruption hypothesis: How spelling and typing skill affects written composition process and product. *Psychological Research*, 86(7), 2239–2255.

Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1), 75–116.

Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038.

Selkirk, E. O. (1995). Sentence prosody: Intonation, stress, and phrasing. In J. Godsmith (Ed.), *The handbook of phonological theory* (Vol. 1, pp. 550–569). Blackwell.

Smith, M., & Wheeldon, L. R. (1999). High level processing scope in spoken sentence production. *Cognition*, 73, 205–246.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 174–215.

Sorensen, T., Hohenstein, S., & Vasishth, S. (2016). Bayesian linear mixed models using stan: A tutorial for psychologists, linguists, and cognitive scientists. *Quantitative Methods for Psychology*, 12(3), 175–200.

Swets, B., Jacovina, M. E., & Gerrig, R. J. (2014). Individual differences in the scope of speech planning: Evidence from eye-movements. *Language and Cognition*, 6(1), 12–44.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., Anal'a, L. A., Vargha, A., & Bates, E. (2005). Timed action and object naming. *Cortex*, 41(1), 7–25.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., Jacobsen, T., & Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, 35(4), 621–633.

Székely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pléh, C., Wicha, N., *et al.* (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, 51(2), 247–250.

Torrance, M., & Nottbusch, G. (2012). Written production of single words and simple sentences. In V. Berninger (Ed.), *Past, present, and future contributions of cognitive writing research to cognitive psychology* (pp. 403–422). Taylor Francis.

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., & Grasman, R. (2010). Bayesian hypothesis testing for psychologists: A tutorial on the savage-Dickey method. *Cognitive Psychology*, 60(3), 158–189.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the flexibility of grammatical advance planning during sentence production: Effects of cognitive load on multiple lexical access. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2), 423–440.

Wheeldon, L. R., & Konopka, A. (2023). *Grammatical encoding for speech production*. Cambridge University Press. https://doi.org/10.1017/9781009264518.

Wheeldon, L. R., Ohlson, N., Ashby, A., & Gator, S. (2013). Lexical availability and grammatical encoding scope during spoken sentence production. *The Quarterly Journal of Experimental Psychology*, 66(8), 1653–1673.

Wheeldon, L. R., Smith, M. C., & Apperly, I. A. (2011). Repeating words in sentences: Effects of sentence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5), 1051–1064.

Appendix A

Picture-naming data

Picture naming norms were taken from the International Picture Naming Project (e.g. Bates et al., 2003).

Table A1. Picture-naming data

Preview/ Target	Phrase	Contrast	Name agreement	Syllables	Characters	AoA	Frequency
Preview	head	head	0.2 (0.2)	1.4 (0.6)	4.9 (1.5)	2.2 (0.9)	8.4 (0.6)
Preview	head	modifier	0.3 (0.3)	1.6 (0.8)	5.2 (1.6)	2.2 (0.9)	8.4 (0.5)
Preview	modifier		0.4 (0.3)	1.6 (0.6)	5.5 (1.2)	2.3 (0.9)	8.2 (0.5)
Target	head	head	0.2 (0.2)	1.4 (0.6)	4.9 (1.5)	2.2 (0.9)	8.4 (0.6)
Target	head	modifier	0.3 (0.3)	1.6 (0.6)	5.6 (1.7)	2.3 (0.9)	8.3 (0.7)
Target	modifier		0.3 (0.3)	1.6 (0.8)	5.1 (1.8)	2.1 (0.9)	8.5 (0.5)

Note. AoA = Age of acquisition based on MacArthur Communicative Development Inventories (Fenson et al., 1994); Name agreement = 0 indicates perfect agreement; values $>$ 0 indicate lower agreement (see Snodgrass & Vanderwart, 1980); Frequency = log frequency based on British National Corpus. Means and standard deviations (in parentheses) by 'Phrase' indicating the location of the picture in the phrase, and 'Contrast' indicating the phrase location of the contrasting picture.

Appendix B

Fixed-effects summary of modifier-choice data

Table B1. Modifier choice

| Predictor	Est. with 95% PI	BF
Main effects		
Preview	0.1 [−2.2, 2.34]	0.5
Modality	18.75 [10.5, 28.65]	> 100
Contrast	−1.92 [−6.82, 2.95]	1.3
Two-way interactions		
Modality: Contrast	−0.2 [−3.17, 2.58]	0.6
Preview: Modality	−0.07 [−2.38, 2.29]	0.5
Contrast: Preview	0.02 [−2.3, 2.22]	0.5
Three-way interaction		
Contrast: Preview: Modality	−0.01 [−2.42, 2.32]	0.5

Note. Fixed effects for proportion of trials in which participants produced a postnominal modifier ('The king with the magnet') instead of a prenominal modifier ('The king's magnet'). Shown are main effects and interactions on logit scale. 95% PIs are shown in brackets.
 Abbreviations: BF = evidence in favour of the alternative hypothesis over the null hypothesis (Bayes Factor); PI = probability interval; ':' = interaction.

Appendix C

Eye data of target trials

For an overview of the eye-movement data, we modelled the time-course of eye-movement. The model included gaze data to the four AOIs of the preview (for the preview trials see Appendix E) and target trials that correspond to the two nouns in each target utterance. Data from pictures that were not part of either the preview or target utterance were not included in the analysis as only a negligible proportion of eye samples were on either of these four AOIs (all means = .02, *SEs* = .02).

The proportion of eye samples was converted to empirical logits and modelled in Bayesian mixed-effects models with three orthogonal polynomials for the time-course predictor (see Barr, 2008; Mirman et al., 2008)
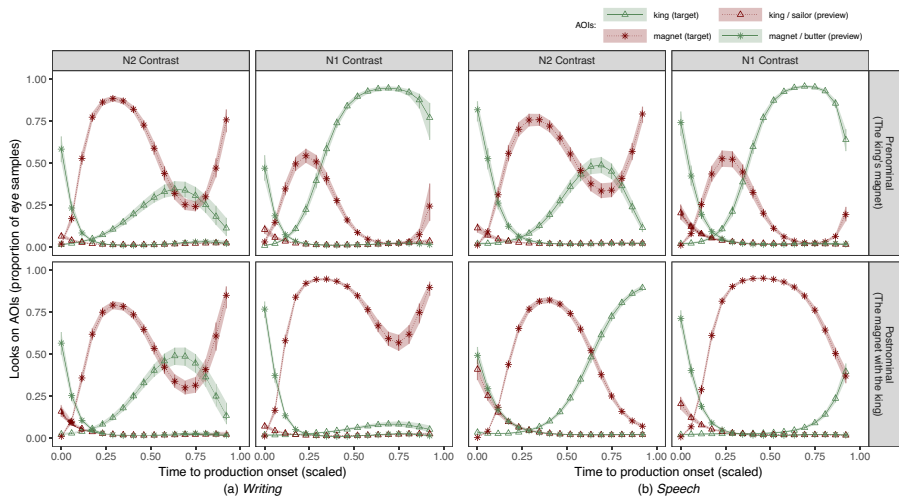
Figure C1. Gaze time-course of target trial. Displayed are the proportion of eye samples to the AOIs corresponding to noun referents in the target and preview trials (as indicated in legend keys) before production onset. Lines represent population estimate and error bars show 95% PIs.

and main effects and interactions of AOIs, preview/target, Modifier position and Contrast. The time axis was scaled between 0 (stimulus onset) and 1 (response onset) to account for by-trial differences in the time-to-onset duration.

The modelled time-course data can be found in Figure C1. As can be seen, looks to the referents of the preview utterance were rare (referent of the modifier noun, e.g. 'king'/'sailor': mean = .06, SE = .03; head-noun referent, e.g. 'magnet'/'butter': mean = .1, SE = .04) and mainly preceded the gaze switch to the referents of the target utterance. Thus, these looks are mainly an artifact of attention re-allocation from preview to target referents.

Appendix D

Fixed-effects summary of eye data and onset latency

Table D1. Fixed effects for eye and response-latency data

Predictor	Time-to-gaze divergence		Prob. of gaze divergence		Onset latency	
	Est. with 95% PI	BF	Est. with 95% PI	BF	Est. with 95% PI	BF
Main effects						
Modality	−0.27 [−0.46, −0.1]	7.6	5.63 [3.26, 8.03]	> 100	−3.2 [−3.74, −2.66]	> 100
Preview	−0.54 [−0.7, −0.4]	> 100	3.28 [1.7, 4.91]	> 100	2.59 [2.27, 2.93]	> 100
Modifier	0.34 [0.14, 0.55]	64.5	−5.39 [−8.25, −2.68]	> 100	0.38 [−0.22, 0.94]	0.7
Contrast	0.27 [0.15, 0.39]	> 100	−5.19 [−6.7, −3.75]	> 100	0.53 [0.25, 0.81]	67.8
Two-way interactions						
Modality: preview	0.33 [0.18, 0.47]	> 100	−1.93 [−3.61, −0.34]	16.1	−0.06 [−0.35, 0.21]	0.2
Modality: modifier	−0.14 [−0.31, 0.04]	0.3	−2.8 [−5.16, −0.31]	15.2	−0.86 [−1.34, −0.36]	41.6
Modality: contrast	−0.15 [−0.27, −0.03]	1.4	−0.22 [−1.89, 1.47]	0.8	0.09 [−0.17, 0.35]	0.2
Preview: modifier	−0.13 [−0.29, 0.02]	0.3	0.55 [−1.16, 2.3]	1.1	0.06 [−0.23, 0.34]	0.2
Preview: contrast	−0.19 [−0.33, −0.04]	1.9	3.05 [1.48, 4.65]	> 100	−0.25 [−0.49, 0.01]	0.8
Modifier: contrast	0.05 [−0.08, 0.17]	0.1	1.29 [−0.33, 2.81]	2.9	−0.36 [−0.64, −0.08]	3.5

(Continued)

Table D1. (*Continued*)

Predictor	Time-to-gaze divergence		Prob. of gaze divergence		Onset latency	
	Est. with 95% PI	BF	Est. with 95% PI	BF	Est. with 95% PI	BF
Three-way interactions						
Modality: preview: modifier	0.16 [0.02, 0.29]	0.8	−0.99 [−2.6, 0.63]	1.7	−0.47 [−0.76, −0.19]	62.9
Modality: preview: contrast	0.17 [0.04, 0.31]	1.5	−0.98 [−2.6, 0.61]	1.7	−0.05 [−0.31, 0.2]	0.1
Preview: contrast: Modifier	0.04 [−0.09, 0.18]	0.1	−1.11 [−2.66, 0.41]	2.4	0.14 [−0.15, 0.45]	0.2
Modality: contrast: modifier	−0.1 [−0.23, 0.02]	0.2	1.73 [0.15, 3.35]	8.5	0.13 [−0.12, 0.39]	0.2
Four-way interaction						
Modality: preview: modifier: contrast	−0.01 [−0.14, 0.13]	0.1	0.07 [−1.45, 1.65]	0.8	0.02 [−0.23, 0.28]	0.1

Note. Estimates are shown for all main effects and interactions. Time-to-gaze-divergence and onset latency are shown on the log scale and the probability of pre-onset gaze divergence on the logit scale. 95% PIs are shown in brackets. Abbreviations: BF = evidence in favour of the alternative hypothesis over the null hypothesis (Bayes Factor); PI = probability interval; ':' = interaction.

Appendix E

Results of preview trials

This section summarises the results for preview trials extracted from the models reported in the main text. Figure E1a shows the estimates of the posterior parameter values for the time of gaze shift from N1 to N2 in the written session. Evidence for scope effects was negligible (BF < 0.06). Similarly, the probability of gaze shift to occur before production onset showed negligible evidence for a contrast effect (BF < 0.6). The estimated probability of gaze-shift was similar in the N2 contrast condition both modifier types (prenominal: N1 contrast = .44, 95% PI [.34, .54], N2 contrast = .52, 95% PI [.41, .62]; postnominal: N1 contrast = .52, 95% PI [.42, .63], N2 contrast = .5, 95% PI [.4, .61]). The inferred parameter values for the response-onset latency can be found in Figure E1b. Evidence for contrast effects was negligible (BF < 0.2).

For the spoken responses, the posterior parameter values for the time of gaze shift from N1 to N2 relative to the production onset can be found in Figure E2a. Evidence for noun contrast effects was negligible (BF < 0.08). Although we found negligible evidence for a noun contrast effect in the gaze-shift timing data, we observed evidence for a larger probability for gaze shift to happen before production onset under N2 contrast in phrases with postnominal modification but not for phrases with prenominal modification (prenominal: N1 contrast = .45, 95% PI [.33, .58], N2 contrast = .5, 95% PI [.36, .63], BF < 0.4; postnominal: N1
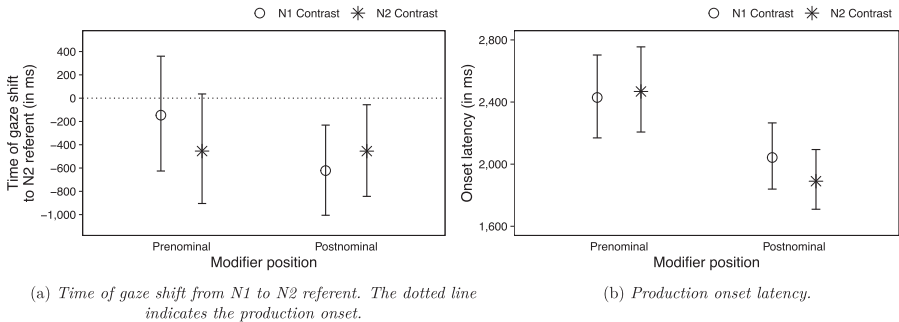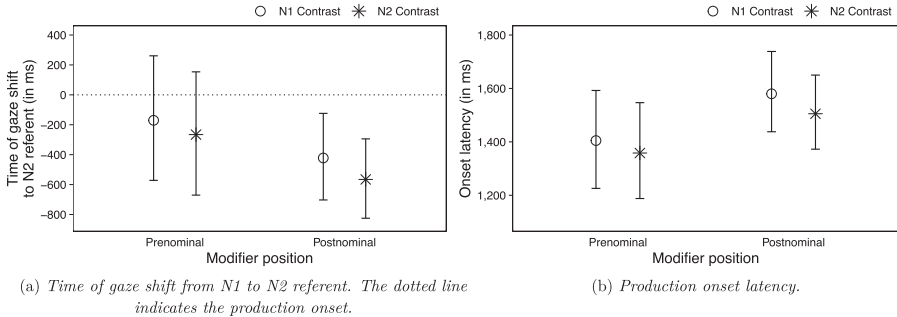

(a) *Time of gaze shift from N1 to N2 referent. The dotted line indicates the production onset.*

(b) *Production onset latency.*

Figure E1. Written production. Estimates from full-factorial Modality × preview/target × Modifier position × Contrast linear mixed-effects models. Error bars represent 95% PIs.

(a) *Time of gaze shift from N1 to N2 referent. The dotted line indicates the production onset.*

(b) *Production onset latency.*

Figure E2. Spoken production. Mean gaze shift and production onset latencies. Error bars represent 95% PIs. Parameter estimates from full-factorial Modality × preview/target × Modifier-position × Contrast linear mixed-effects models.
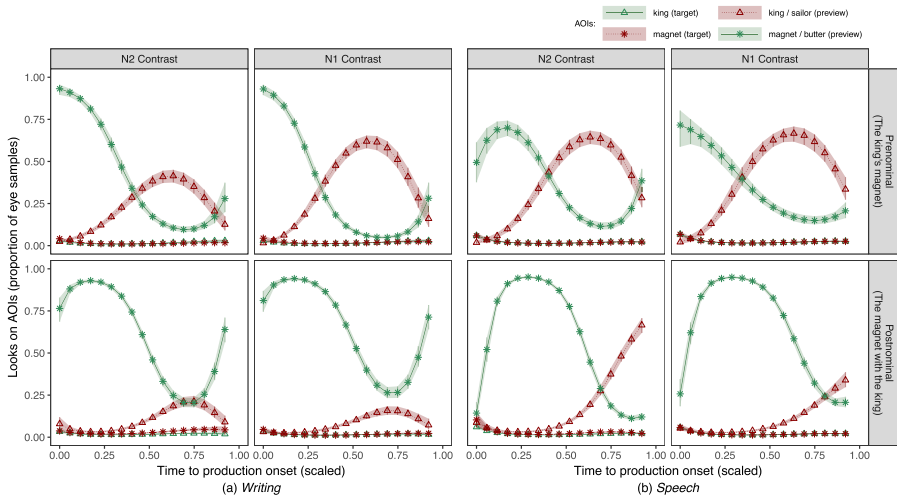
Figure E3. Gaze time-course of preview trial. Displayed are the proportion of eye samples to the AOIs corresponding to noun referents in the target and preview trials (as indicated in legend keys) before production onset. Lines represent population estimate and error bars show 95% PIs.

contrast = .65, 95% PI [.56, .74], N2 contrast = .78, 95% PI [.71, .85], BF = 6.78). The inferred parameter values for the onset latency can be found in Figure E2b. Evidence for noun-contrast effects was negligible (BF < 0.09).

The time-course data, modelled as described in Appendix C, are shown in Figure E3. As can be seen, looks to the referents of target-trial utterances that followed the preview were rare (referent of the modifier noun, e.g. 'king': mean = .03, SE = .02; head-noun referent, e.g. 'magnet': mean = .04, SE = .03). These were non-different from looks to any of the other four AOIs (all means = .02, SEs = .02).