COGNITIVE NEUROPSYCHOLOGY
https://doi.org/10.1080/02643294.2025.2515831

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Speakers preplan lexical and phonological representations in semantically constraining linguistic contexts

Irina Chupina [a], Britta U. Westner[a,b], Ardi Roelofs[a] and Vitória Piai[a]

[a]Radboud University, Donders Centre for Cognition, Nijmegen, The Netherlands; [b]Department of Medical Neuroscience, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

ABSTRACT

Rapid turn-taking in conversation suggests that speakers plan part of their turn in advance, but evidence for this is scarce. Using context-driven picture naming, we examined whether (a) speakers preplan lexical-semantic and phonological information at the word level in constraining sentential contexts, and (b) phonological preplanning encompasses the whole word. Analysis of naming response times (RTs) showed that constraining contexts enable preplanning of both lexical-semantic and phonological representations (Experiment 1). Using a picture-word interference version of the same task (Experiment 2), we found that speakers preplan the phonological form of the whole word, however, only the subset of constraining trials with the shortest RTs indicated preplanning. The results confirm previous findings from turn-taking, which suggested that speakers can complete later stages of lexical access in advance, but also highlight that the presence of preplanning varies from trial to trial.

ARTICLE HISTORY
Received 18 November 2024
Revised 28 May 2025
Accepted 30 May 2025

KEYWORDS
Prediction; word production; lexical frequency; phonological facilitation; word length; Bayesian estimation; ex-Gaussian

Introduction

In conversational contexts, speaking relies on the interaction of language comprehension and production. The speaker has to comprehend what has been said and then, taking turn, first plan relevant conceptual information (the meaning), then select a lemma (the lexical-syntactic representation) and finally generate the word form (Figure 1). According to models of spoken word production (e.g., Dell, 1986; Levelt et al., 1999), word form generation requires the encoding of the morphological, phonological and phonetic information of the target word. These stages represent the retrieval of morphemes, the generation of a phonological representation, and of the articulatory programme, respectively. Going through these planning stages takes at least 600 ms when naming a depicted object (Indefrey & Levelt, 2004) and at least 500 ms in speeded naming (e.g., Lampe et al., 2023). Starting your speaking turn in a conversation, however, has been suggested to be as fast as 200 ms (Heldner & Edlund, 2010; but see Corps et al., 2022). This is close to the minimal response speed observed when repeating syllables in a simple reaction task without stimulus discrimination or choice, as first demonstrated by Frans Donders in the mid-nineteenth century (Donders, 1968; see Meyer, 2023; Roelofs, 2018, for discussion). The reaction time (RT) limits of picture naming and response generation tasks pose an interesting question regarding conversations: How can turn-taking happen so quickly if simple reaction tasks take equally long?

It has been argued that the fast timing of turn-taking in conversation is enabled by early planning of the response while still listening to the interlocutor (Levinson & Torreira, 2015). To initiate articulation promptly, multiple stages of utterance planning need to be completed in advance, with representations stored in working memory ready to be executed (Levinson, 2016). Although advance planning in context has been hypothesized to proceed as far as articulatory preparation (e.g., measured as tongue movement using ultrasound, Bögels & Levinson, 2023), empirical research that directly investigates the processing stages completed before speech initiation is scarce. Consequently, we do not fully understand the extent of response preplanning before articulation in different types of communicative scenarios where language comprehension and production interact. To further complicate the picture, the extent and even presence of advance utterance

CONTACT Irina Chupina irina.chupina@donders.ru.nl Radboud University, Donders Centre for Cognition, Nijmegen, Thomas van Aquinostraat 4, 6525 GD, The Netherlands

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02643294.2025.2515831.

© 2025 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

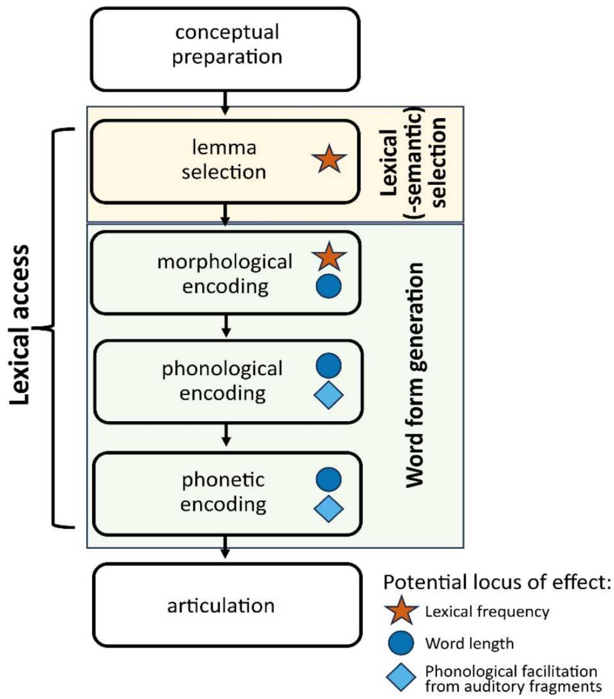

Figure 1. Word (pre)planning stages with tentative loci for the lexical frequency, word length and phonological facilitation effects. Based on Levelt el al., 1999. Reproduced with permission from the authors from doi:10.6084/m9.figshare.27645696.

planning might depend on the situational demands (e.g., is there time to plan or is a quick response necessary?) and task demands (e.g., is a yes/no response sufficient or does a bigger utterance need to be planned?) (Schriefers, 1999). Finally, preplanning can also vary from trial to trial within speakers (Kilbourn-Ceron & Goldrick, 2021).

One comprehension-production paradigm that can elicit word preplanning is the context-driven object naming task (Griffin & Bock, 1998; Piai et al., 2014). In this task, participants read or listen to sentences that are missing the final word. After a short interval, they get presented with a picture, which the participants have to name to complete the sentence. The lead-in sentences are either semantically constraining (*The cat hunts a* … [mouse]) or unconstraining (*The stain looks like a* … [mouse]). Naming in the unconstrained sentences is similar to bare picture naming (Chupina et al., 2022; Roos et al., 2024) since the context is neutral and word planning initiates after picture presentation. A high constraint in the sentence context, however, enables participants to start response preparation before seeing the picture. The picture presentation can be interpreted as the response initiation cue. Shorter picture naming RTs have been consistently observed in the constrained compared to unconstrained condition, suggesting that the target word is preplanned (Griffin & Bock, 1998; Piai et al., 2014; Roos et al., 2024). Although, at first glance, this type of sentence completion task may seem different from more naturalistic paradigms used in turn-taking studies, such as question-answer, it likely induces a sequence of highly similar cognitive events (for the discussion of the differences, see General Discussion): The listener retrieves a concept that is activated by an externally-generated input and plans a one-word utterance (here, as a completion) that can be executed after the cue for the end of the turn is perceived (here, the start of the pre-picture gap).

Previous context-driven naming literature suggests that the constrained condition encourages at least lexical-semantic preplanning. Piai et al. (2014) and Gastaldon et al. (2020) reported that the classical *lexical frequency effect* (i.e., words with higher lexical frequency are named faster than words with lower lexical frequency) was present in the RTs in the unconstrained but not constrained condition. The authors interpreted the absence of the effect in the constrained condition (where word preplanning is likely to occur) as an indication that the lexical-semantic information had already been retrieved by the time the picture appeared, rendering the frequency effect not detectable in the RTs measured from picture onset. Another piece of evidence suggesting that the pre-picture processes might be lexical-semantic in nature comes from an electrophysiological study by Piai et al. (2020). The authors demonstrated that the electrophysiological signatures associated with word-retrieval processes become delayed when a semantically related distractor is presented in the pre-picture interval but not when an unrelated distractor is presented.

Phonological preplanning has been reported in other types of production paradigms. One of the few studies that directly tapped into advance phonological encoding is a behavioural study by Barthel and Levinson (2020). In this study, participants were presented with four objects shown simultaneously on a screen. They were then asked to answer a subsequent question by naming the matching object out of the four. When preparing the response, participants had to switch from the naming task to an intervening lexical decision task on a quarter of the trials prompted by a change in the display. The lexical decision word following the four pictures was either phonologically related or unrelated to the target picture. Participants made more errors and showed an interference effect in the RTs in the related compared to the unrelated condition, which the authors interpreted as evidence for the word form of the original target picture name having already been retrieved by the time of the task switch. Another study by Corps and Pickering (2024) investigated whether speakers plan beyond lexical-semantic information when giving responses of different length (single- or multi-word), using a question-answer paradigm. They found that, concurrently with listening to speech, speakers would preplan the content but not the form, retrieving the latter only when they were ready to initiate articulation. Phonological preplanning has also been investigated using electrophysiology. Participants in Bögels et al. (2015) answered questions which contained key words allowing response planning either in the middle or at the end of the sentence. The authors suggested that phonological preplanning might have occurred based on the post-hoc observation that the brain regions involved in these conditions are associated with phonological and syllabification processes. Jongman and colleagues (2020), who focused on the phonological hypothesis more directly using a pseudo-word naming paradigm, however, did not find electrophysiological signatures associated with early phonological encoding. The authors argued that either those processes could not be detected with EEG or that in tasks like Bögels and colleagues' (2015), preplanning could stop at the earlier stages without reaching phonological encoding. Thus, while there is some evidence for word form preplanning in various contexts, most findings remain indirect or inconclusive.

When talking about phonological planning, another open question is how far it proceeds before the speaker can start articulation. On the one hand, it has been proposed that articulation can initiate only after the whole phonological form has been encoded (Damian & Dumay, 2007; Meyer et al., 2003; Meyer & Schriefers, 1991; Roelofs, 1997). On the other hand, the size of the phonological planning unit has been suggested to vary depending on the speaking demands (Kilbourn-Ceron & Goldrick, 2021; Schriefers, 1999). In particular, when planning utterances are longer than one word, articulation could already start after the first syllable has been planned (Schriefers, 1999; but see Meyer et al., 2003). This debate is also relevant for advance phonological encoding, since speakers could require less time and resources (e.g., buffering) for preplanning if articulation started already after phonological encoding of the first syllable of the preplanned word compared to after planning the whole-word phonological representation.

Present study

Here, we investigated preplanning at the word level during two stages of lexical access, namely lexical (lemma) selection and the following phonological encoding stage (see Figure 1), with the focus on the latter. Using context-driven naming, we asked whether preplanning in the constrained contexts proceeds as far as lexical-semantic and phonological encoding. Furthermore, if phonological information is indeed preplanned, can speakers proceed to prepare the articulatory code with only an initial fragment (e.g., phoneme or syllable) encoded or do they prepare the whole word?

To answer the question about the extent of word preplanning, in Experiment 1, we analysed RTs previously collected during the context-driven naming task. Using the reasoning similar to Piai and colleagues (2014), we used the *lexical frequency effect* and *word length in phonemes* as proxies for lexical-semantic and phonological processes, respectively. It should be noted that the psycholinguistic locus of these effects has not been unambiguously established. The lexical frequency effect could potentially arise at either or both lexical selection and word form encoding stages (Almeida et al., 2007; Corps & Meyer, 2023; Jescheniak & Levelt, 1994; Kittredge et al., 2008), while word length could impact both phonological and phonetic encoding and even, to a lesser extent, morphological encoding (Levelt et al., 1999; Meyer et al., 2003). Since differentiating these stages is not within the scope of the current study, broadly speaking, we use the term *lexical(-semantic)* to refer to the earlier processes of the "holistic" retrieval of the word representations from memory and the term *phonological* to the later-occurring combinatorial word-generation processes. Assuming that both processes are completed in the interval preceding picture presentation but after sentence offset (i.e., pre-picture interval), we expected to see no lexical frequency and word length effects in the RTs in the constrained condition. By contrast, we expected to see both effects in the RTs in the unconstrained condition since there, word planning can start only upon picture presentation. In other words, we predicted that participants would name shorter words and words with higher frequency faster, but only in the unconstrained condition.

In Experiment 2, we investigated the presence and extent of phonological preplanning using a different manipulation of phonological encoding. By adding auditory fragments (distractors) to a subset of the stimuli from Experiment 1, we effectively created a picture-word interference (PWI) type of task (e.g., Shao & Rommers, 2020). The fragments were presented simultaneously with the picture. Each auditory fragment was either phonologically related or unrelated to the target word. Compared to unrelated distractors, related words speed up naming, known as the *phonological facilitation effect* (e.g., Lupker, 1982). Since the fragments we used carry no lexical or morphological information, the facilitation is likely to arise at the phonological encoding stage (although phonetic encoding impact cannot be ruled out, see *General discussion*). In Experiment 2A, we presented fragments overlapping with the initial phonemes of the target word (e.g., target *mouse*, related: /MAU/, unrelated: /TREI/). In Experiment 2B, we presented fragments overlapping with the ending of the target word (related: /AUS/, unrelated: /EIN/). In both experiments, we expected phonological facilitation in the unconstrained condition given that this condition tends to elicit naming behaviour similar to bare picture naming (Chupina et al., 2022; Roos et al., 2024). By contrast, in the constrained condition, we assumed that if phonological encoding has been completed before the picture onset, there will be no phonological facilitation from distractors. If the final phonemes of the target word do not facilitate naming (Experiment 2B), this would indicate that the final phonemes, i.e., the whole phonological word, have been preplanned. However, seeing no facilitation exclusively with the initial fragments (Experiment 2A) would suggest that the preplanning did not proceed beyond the initial phoneme(s).

Experiment 1

Method

Participants. We analysed RT data from 58 university students (15 male; for n = 48,[1] mean age = 23.3, SD = 6.1, range 18–48) who performed the context-driven object naming task (Piai et al., 2014). The data were collected at the Donders Centre for Cognition (Radboud University) in Nijmegen, the Netherlands as part of two electrophysiological experiments (Chupina et al., in prep; Westner et al., in prep). The study was approved by the Ethics Committee of the Faculty of Social Sciences at Radboud University, following the Declaration of Helsinki. Participants gave written informed consent and received monetary compensation or course credits for participation.

Materials. The materials used in the two experiments included 106 and 80 target nouns, with 35 items overlapping between experiments. When combined for the current study, this resulted in 151 unique targets, each paired with a constrained and unconstrained lead-in sentence in Dutch (see Figure 2B). The average cloze probability (Taylor, 1953) was 0.88 (range: 0.57–1) in the constrained and 0.03 (range: 0–0.39) in the unconstrained sentences. These cloze probabilities have been collected in our group over the years in multiple studies (each with N ≥ 16). The sentences were recorded by a female native speaker and normalized using Audacity® (Audacity Team, 2014). The average sentence duration was 2400 ms (1390–4590 ms). The pictures were selected from the MULTIPIC database (Duñabeitia et al., 2018). The 23 pictures not available in the database were found online. The list of the stimuli can be found at https://osf.io/5k2qj/.

Design and procedure. We operationalized lexical-semantic and phonological processes as the lexical frequency and word length effects, respectively. As the lexical frequency measure, we used *Zipf frequency* from the SUBTLEX-NL database (Keuleers et al., 2010). The Zipf frequency is a scaled lexical frequency, obtained by taking the logarithm base 10 of the word frequency per billion words (for more detail, see Brysbaert et al., 2018). For the picture names from the present experiment, the mean Zipf frequency was 4.2 (SD = 0.62, median = 4.1, range 2.5–5.7). For reference, the values of 1–3 are considered to represent low-frequency words and the values of 4–6 high-frequency words, with frequencies above 6 mainly belonging to function words such as articles or prepositions (Brysbaert et al., 2018). To measure the word length effect, we used word length in phonemes. The target word lengths in phonemes were obtained using the CLEARPOND database (Marian et al., 2012), with values for several missing words added manually. The target word length in the sample ranged from 1 (*ui*, onion) to 9 (*postzegel,* postal stamp), with an average of 4.4 phonemes (SD = 1.35, median = 4).

The experiment was programmed in Presentation® software (version 23.1, Neurobehavioral Systems, Inc. (2023), Berkeley, CA, www.neurobs.com). The presentation of stimuli was pseudorandomized using Mix (van Casteren & Davis, 2006), with a maximum of three successive repetitions of the same constraint in a row and a minimum distance of 20 trials between the same target pictures. Each participant received a unique pseudorandomized list. Participants performed the task in the EEG lab. The responses were recorded with a microphone for offline analysis. Before the experiment, participants were familiarized with the target pictures: The procedure is known to increase naming accuracy while not impacting response speed (e.g., Rabovsky et al., 2016). For the experiment proper, participants were instructed to listen to the sentences and complete them by naming the picture presented at the end. The experiment started with a short practice session. Each trial started with the fixation cross which stayed onscreen throughout the duration of the auditory lead-in sentence and the following 800-ms pre-picture interval (see Figure 2A). The choice of the interval in the original studies was determined by the requirements of electrophysiological analyses. Given that the extent of

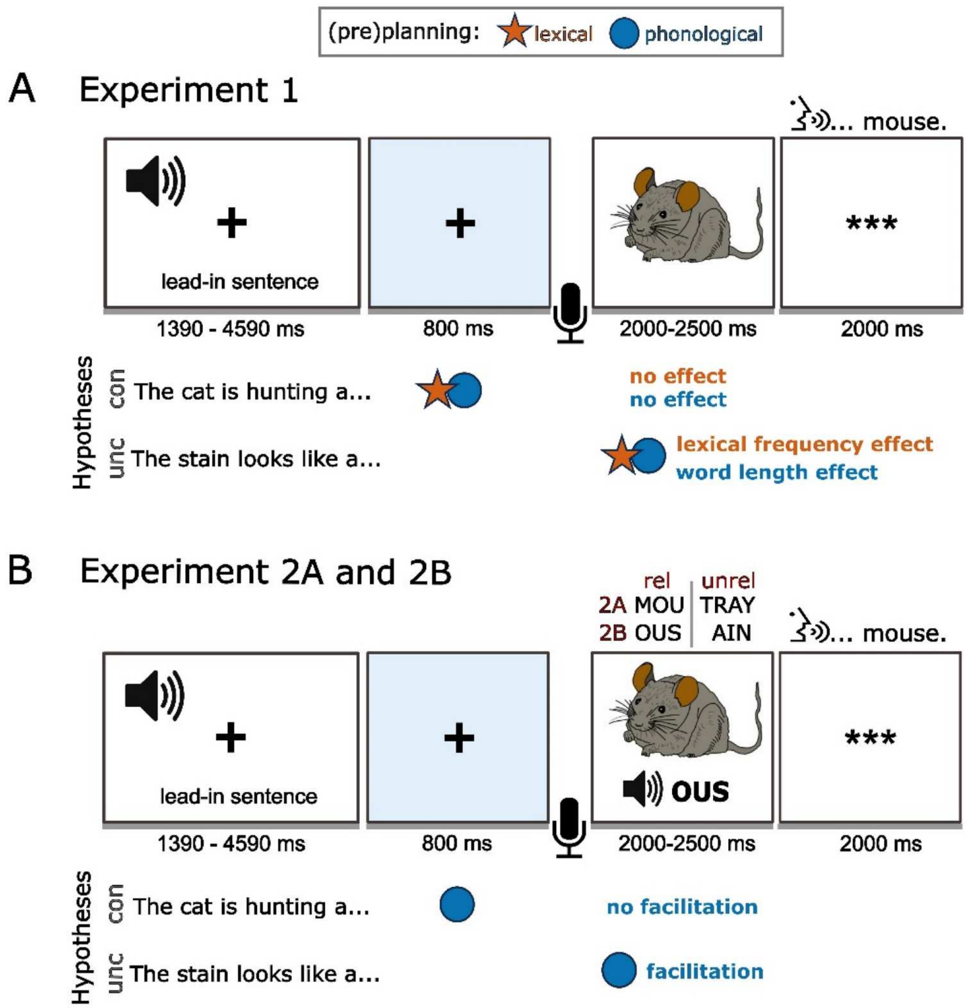

Figure 2. Trial design and hypotheses for word (pre)planning effects in Experiments 1, 2A and 2B. Reproduced with permission from the authors from doi:10.6084/m9.figshare.27645741. **A.** In Experiment 1, participants listened to lead-in sentences, followed by an 800-ms window (highlighted in light-blue) and the target picture (presented for a random interval between 2000 and 2500 ms). The orange star and blue circle indicate the hypothesized timing of (pre)planning associated with lexical and phonological information retrieval, respectively (cf. Figure 1). These processes were expected to occur in the pre-picture interval in the constrained condition (con) and after picture presentation in the unconstrained condition (unc). Accordingly, the corresponding lexical frequency and word length effects were expected to be present in the RTs in the unconstrained condition but absent in the constrained condition, indicating preplanning of information. **B.** In Experiments 2A and 2B, participants performed the same task, but phonologically related (rel) and unrelated (unrel) auditory fragments were presented simultaneously with the picture. The fragments overlapped either with the initial phonemes of the target word (2A) or with the final phonemes of the target word (2B). The hypotheses for phonological preplanning inferred based on the phonological facilitation effect follow the same logic as in Experiment 1.

preplanning can be impacted by the time constraints (see General Discussion), we kept it consistent throughout the present study (also in Experiment 2) for comparability with the prior literature. After the interval, the picture was presented for 2000–2500 ms, followed by a 2000-ms screen with asterisks. Next, the new trial began. The microphone started response recording simultaneously with picture presentation.

Data analysis. All analyses were conducted in R Statistical Software (v4.1.3; R Core Team, 2022). The RT data and R code are available at https://osf.io/5k2qj/. RTs were calculated manually using PRAAT (Boersma & Weenink, 2013) with experimenters blind to the experimental condition. Responses were coded as correct only when the target name was produced. For instance, for the target *boontjes* (beans), the otherwise acceptable naming alternatives such as *sperziebonen* (green beans) or *bonen* (beans) were excluded from the RT analyses. The exact target name was required since the analyses relied on the lexical frequency and word length measures that would be different for alternative versions. Three hundred eighty-one incorrect (3.5% out of 10,892 total responses) and 217 non-target trials (2% out of total responses) were excluded. Furthermore, some responses in the unconstrained condition were uncharacteristically fast. This was observed mainly when the unconstrained sentence was presented after the constrained one, suggesting priming effects from the overlapping semantic information in the sentences. Thus, we removed an additional 55 unconstrained responses (out of 10,294, 0.5%) faster than 470 ms based on the cut-off of 2.5 SD from the mean RT (609 ms) obtained in a speeded naming task (Lampe et al., 2023).

First, we performed an RT distributional analysis. The raw RT distributions are typically non-normally distributed. Distributional analysis using ex-Gaussian decomposition (Balota & Yap, 2011) allows to separate the normally distributed bulk of the RTs (represented by the parameters *mu* and *sigma* that correspond to the mean and SD of the normal distribution, respectively) and the exponentially distributed tail of the RTs (parameter *tau*, the combined mean and SD). This analysis aimed to confirm that, overall, word planning in the constrained condition starts earlier than in the unconstrained condition. This would be indicated, visually, by the RT distributions in both conditions having a similar shape but the whole constrained condition distribution being shifted towards the left, that is, faster responses. Statistically, the difference between conditions should come from the bulk of the distribution and not (only) from the tail, i.e., a significant difference between the *mu* but not (only) *tau* parameters. The ex-Gaussian parameters were estimated per individual participant on data quantiles within condition using the Quantile Maximum Likelihood Estimation (Brown & Heathcote, 2003). To compare the parameters across conditions, we used dependent t-tests.

In addition, we calculated the effect sizes (ES) for the *mean*, *mu* and *tau* as *unbiased Cohen's d*, using the following repeated design equation (for detail see Cumming, 2013):

$$unbiased\ d = \frac{mean_{unc} - mean_{con}}{\sqrt{\frac{SD_{unc} + SD_{con}}{2}}} * \left(1 - \frac{3}{4*(n-1)-1}\right)$$

with $mean_{unc}$ and $mean_{con}$ being the mean of the unconstrained and constrained condition, respectively, and SD_{unc} and SD_{con} being the standard deviations of the respective conditions. n refers to the number of participants. The mean per condition was calculated first within participant across trials, and then averaged across participants, with standard deviations calculated over the individual participant means. Based on the *unbiased d*, the 95% confidence intervals (CIs) for the ES were calculated using the R package MBESS (Kelley et al., 2018).

Finally, we performed Bayesian modelling using the *brms* package (Bürkner, 2017) in R. We chose Bayesian modelling over the frequentist linear mixed models (including those using link functions, e.g., Lo & Andrews, 2015) because this approach allows for a flexible choice of the likelihood function to model the non-normally distributed RTs. Furthermore, Bayesian modelling promotes a different approach to parameter estimation based on the notion of result credibility: Unlike confidence intervals, providing a range of values which would not be rejected given a certain significance level, the Bayesian credible intervals (CrIs, also called uncertainty intervals) provide the probability distribution of the estimates falling within a certain range of parameter values given the data (Kruschke & Liddell, 2018; for the discussion of the other advantages, see Kruschke, 2010; Wagenmakers et al., 2018). In the present experiment, if the probability (posterior) distribution of the effect did not overlap with zero and/or did not contain estimates with the opposite sign, we interpreted the effect to have more credibility.

We fitted a hierarchical linear regression model using the ex-Gaussian likelihood function with identity link (4 chains, 20,000 iterations, 4000 warm-up). The dependent variable was the raw picture-naming time in milliseconds. *Condition* (constrained, unconstrained) and interactions between *condition / Zipf frequency* and *condition / word length* entered the model as fixed effects. By-subject and by-item intercepts entered the model as random effects. We standardized the continuous *Zipf frequency* and *word length* variables and used treatment contrast coding for the categorical variable *condition*, with the constrained condition as the baseline. We ran the model with a default (flat) prior. To confirm that the prior choice did not impact our model estimates (Kruschke & Liddell, 2018), we recomputed the model with a narrow (informed) and a broad (uninformed) prior (see Supplement A). We focused on estimation interpretation using uncertainty (Cumming, 2014; Kruschke & Liddell, 2018) since hypothesis testing based on Bayes factor (BF) remains debated in literature (Gelman et al., 2013; Kruschke & Liddell, 2018), among other things, due to its dependence on the frequently arbitrary prior specifications. To demonstrate that our interpretation of the findings based on uncertainty intervals would still be consistent with the hypothesis testing outcomes, we report BFs for the effects of interest calculated from the models with narrow and broad priors in the supplement Table S1.

Results

The descriptive, ex-Gaussian and ES statistics can be found in Table 1. Speakers were on average 223 ms

Table 1. Descriptive statistics, effects sizes and ex-Gaussian parameter t-test results for Experiment 1.

Descriptive and Ex-Gaussian statistics (in ms)				
Condition	Mean (SD)	Mu	*Sigma*	*Tau*
Constrained	596 (108)	452	97	142
Unconstrained	820 (84)	639	59	167
Dependent-samples t-test: unconstrained vs constrained				
		t(57) = 15.3, *p*<.0001, CI [163, 212]	t(57) = 3.9, *p*<.001, CI [18, 56]	t(57) = 2.2, *p* = 0.03, CI [3, 49]
Effect sizes, 95% CI				
Unbiased d	2.28, [1.81, 2.75]	1.84, [1.4, 2.27]		0.36, [−0.01, 0.72]

Note: CI = confidence interval, SD = standard deviation, d = Cohen's d.

faster at naming the picture after constraining sentences as compared to the unconstraining condition. In the constraining condition, there was a weak correlation (r = −0.11, *p*<0.001) between RT and cloze probability, meaning that participants named pictures only slightly faster when the cloze was higher. Ex-Gaussian estimations converged on average within 12 iterations (range 2–77). *Mu*, *sigma* and *tau* parameters were significantly different across conditions. The ES calculated as *unbiased d* was 2.28 for the mean and 1.84 for the *mu* parameter, both demonstrating a large effect. The tail parameter *tau* showed a small to medium ES of 0.36, which was unstable as indicated by the 95% CI (−0.01–0.72). Taken together with the visual inspection of the cumulative distribution plot (Figure 3A), these results indicate that both condition distributions have a similar shape, the difference between conditions is large and it results predominantly from distributional shift. Altogether, this pattern strongly supports the claim that planning stages are completed earlier in the constrained relative to the unconstrained condition.

The Bayesian modelling results are presented in Table 2 and Figure 3B and C. The MCM chains converged as indicated by R-hat values of 1. The ESS values indicate stable parameter estimates for the effect of condition (>10,000 as recommended by Kruschke, 2015, when using 95% highest density intervals [HDIs] as CrIs) while the estimates were less stable for the interaction effects. There was a reliable effect of condition (i.e., credible interval not including zero), with participants naming pictures considerably faster in the constrained condition (posterior mean = 231 ms). The regression slopes for the interactions (Figure 3B and C, left panels**)** show that in the unconstrained condition participants named higher frequency and shorter words faster. Statistically speaking, however, since the 95% uncertainty intervals do include zero (Figure 3B and C, right panels), we cannot decisively conclude that these

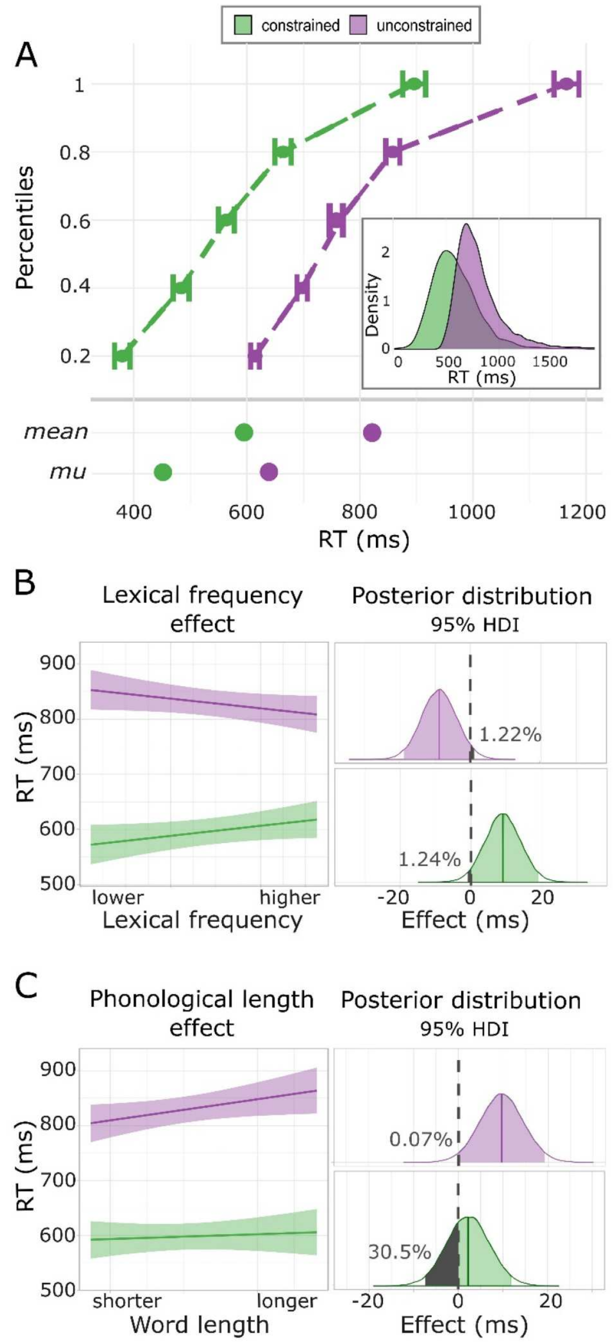

Figure 3. Cumulative distributions of RTs per condition and Bayesian modelling results for Experiment 1. Reproduced with permission from the authors from doi:10.6084/m9.figshare.27645750. **A**. Cumulative RT (in quantiles) and density distributions for constrained (green) and unconstrained (purple) conditions with their respective mean and mu values. **B, C**. The lexical frequency and word length effects within condition. *Left panels*. The regression slopes from the main model with default (flat) priors. *Right panels*. Posterior distributions showing the effects with 95% HDI (Highest density interval) as credible intervals. The areas shaded in dark grey show the percentage of the distribution falling below/above zero, indicating the probability of the absence or the opposite direction of the expected effect (see main text for details).

Table 2. Bayesian hierarchical regression estimates (in ms) for Experiment 1.

Model estimates		
Predictor	*Mean*	*95% HDI*
Intercept	597	[574, 620]
Effect of condition	231	[225, 237]
Constrained: Lexical frequency effect	9	[−1, 19]
Unconstrained: Lexical frequency effect	−9	[−19, 1]
Constrained: Word length effect	2	[−8, 12]
Unconstrained: Word length effect	10	[0, 19]

The intercept is the constrained condition. The highest density credible interval (HDI) is 95%.

effects are present, nor can we completely rule out their absence (lexical frequency: posterior probability of observing values ≥0, i.e., participants being slower when naming higher frequency items, is 1.22%; word length: posterior probability of observing values ≤0, i.e., participants being faster when naming longer words, is 0.07%). In the constrained condition, participants unexpectedly named higher frequency items slower than lower frequency items, while the regression line for the word length effect looks relatively flat. Statistically speaking, the unexpected "flipped" lexical frequency effect might also not be generalizable (the 95% CrI includes zero, with posterior probability of observing values ≤0 being 1.24%). Finally, the word length effect is not credibly present, with the bulk of the posterior probability distribution more symmetrically centred around zero (mean = 2.2 ms, with 30.5% of the values ≤0).

Discussion

We analysed existing RT data from the context-driven object naming task in order to investigate whether speakers start response preplanning early when the linguistic context is semantically constraining. Furthermore, we asked whether speakers can preplan both lexical-semantic and phonological information.

In line with prior literature on context-driven picture naming (e.g., Chupina et al., 2022; Gastaldon et al., 2020; Griffin & Bock, 1998; Piai et al., 2014; Roos et al., 2024), we found a large and reliable effect of condition, with participants naming pictures considerably faster when the preceding context is semantically constraining compared to unconstraining. Ex-gaussian decomposition of RTs within condition demonstrated that the difference between conditions comes from the bulk of the RTs, with the constrained condition distribution shifted to the left, that is, responses being faster throughout the whole distribution. Together with the equally large effect sizes obtained from the mean and the mu parameters, the evidence supports the idea that speakers initiate word planning earlier in the constraining contexts and do not simply complete planning faster after the picture presentation.

Previous studies on context-driven naming using analogous analyses of the lexical frequency effect (Gastaldon et al., 2020; Piai et al., 2014) as well as semantically related distractors in the pre-picture interval (Piai et al., 2020) suggested that this early planning comprises at least the lexical-semantic stages. Following these studies in using the lexical frequency effect as a proxy for lexical processes (Jescheniak & Levelt, 1994), we found that participants produced the higher frequency words faster than lower frequency words in the unconstrained condition. In the constrained condition, surprisingly, participants produced higher frequency items slower. Post-hoc correlational and clustering analyses excluded the cloze probability of the constrained sentences as a potential confound. In other words, sentences with lower cloze probability were not disproportionately associated with higher lexical frequency of the target that could have led to the observed effect, suggesting that this reversed frequency effect might be a spurious finding. Importantly, while the results are consistent with the predictions descriptively, the effects are uncertain statistically: Since small proportions of the posterior distributions' tails in both conditions include zero, we cannot exclude with confidence the possibility of the absence of the lexical frequency effect in the unconstrained and its presence in the constrained condition.

Regarding phonological preplanning, no previous research has used the context-driven picture naming to investigate whether word preplanning in constrained contexts proceeds to phonological encoding. We assumed that the speed of phonological processing will be associated with the length of the target words in phonemes, with shorter words being produced faster. Descriptively, the predicted word length effects were detected in the RTs after picture presentation in the unconstrained condition but were absent in the constrained condition. Statistically, the lack of the effect in the constrained condition is most certain, while the presence of the effect in the unconstrained condition, despite less than one percent of the posterior distribution overlapping with zero, should be considered uncertain.

In sum, while the interaction effects descriptively suggest preplanning at both levels, the modelling results indicate uncertainty regarding these effects. Specifically, while we can talk about the absence of the word length effect in the constrained condition with most certainty, we cannot confidently exclude the absence of both effects in the unconstrained condition. Moreover, all observed effects were rather small, even compared to similar psycholinguistic studies (e.g., word length effect of 20 ms, Meyer et al., 2003; lexical frequency effect of 40 ms, Piai et al., 2014). In the light of the inconclusive picture from the statistics, the small effects, and the experiment being an analysis of the data that were not originally collected to answer the research question, no strong conclusions can be drawn. To further investigate the lack of phonological effects observed in the constrained sentences, we conducted two experiments (Experiments 2A and 2B), measuring phonological preplanning and its extent via the phonological facilitation effect.

Experiment 2

In Experiment 2, participants performed the picture-word-interference (PWI)-like context-driven picture naming task with phonologically related and unrelated auditory fragments (see Figure 2B). The related fragments overlapped with the beginning (Experiment 2A) or with the end (Experiment 2B) of the target words. The related fragments were hypothesized to speed up naming in the unconstrained condition, that is, show the classical phonological facilitation effect. In the constrained condition, we expected no facilitation, indicating that phonological planning was completed earlier. No facilitation from the initial phonemes of the targets (Experiment 2A) would suggest advance planning of at least the initial phonemes, whereas no facilitation from the final fragments (Experiment 2B) would suggest preplanning of the whole phonological representation.

Method

Participants. Twenty-five healthy participants took part in each experiment. One participant aborted the experiment early (2A) and one participant was excluded due to procedural errors (2B). This resulted in the final sample of 24 participants for each experiment (2A: six male, mean age = 22, SD = 6; range 18-30; 2B: five male, mean age = 23.3, SD = 2.9; range 18–43). For details on ethical approval, see Method 2.1.

Materials. First, we chose a subset of target words out of the 151 items used in Experiment 1. Since the auditory fragment needed to overlap precisely with the target, we removed all targets to which participants frequently gave alternative names, such as *rugtas* or *tas* (bag) for the target *rugzak* (backpack). Next, we removed one- and two-phoneme words such as *ui* (onion) or *koe* (cow), since the auditory fragment for these items would be the target word itself. This yielded a subset of 130 items. The list of the stimuli with matched fragments can be found at https://osf.io/5k2qj/.

Experiment 2A: We created a list of fragments identical to the beginning of the target words. The fragments had to start with a consonant, be maximum three phonemes in length and have no lexical meaning. Furthermore, the fragments did not have to follow syllabification rules. This allowed us to avoid the loss of monosyllabic stimuli where the fragment did not completely overlap with the target (e.g., bril – /BRI/, eyeglasses). Such distractors that do not coincide with syllable boundaries are still known to produce phonological facilitation similarly to full syllables (Nardo et al., 2017; Roelofs, 2021). This yielded 123 items in the final list, with 67 monosyllabic (54.4% out of total number), 54 bisyllabic (44% out of total number) and 2 trisyllabic (1.6% out of total number) targets. The initial fragments formed either a consonant–vowel (CV, *maan* – /MA:/, moon) or CCV (*pleister* – /PLEI/, plaster) sequence (see Table 3). Using fragments as opposed to phonologically related whole words ensured that the fragments did not carry any lexical-semantic information (for similar use of fragments see Schriefers, 1999; Starreveld, 2000). Moreover, we chose to present the fragments in the auditory modality in order to avoid orthographic confounds also known to produce facilitation effects (Bi et al., 2009).

A female Dutch speaker recorded the fragments as isolated sound combinations to avoid coarticulation effects (e.g., /MA:/ silence, /PLEI/ silence, etc.). Fragments were normalized using Audacity® (Audacity Team, 2014) and the duration of the fragments longer than 250 ms was shortened using PRAAT (Boersma & Weenink, 2013), resulting in a mean fragment duration of 246 ms (range 174–253 ms). The fragments were checked by two native speakers of Dutch to ensure that the manipulation did not introduce any phoneme distortion.

The fragments overlapping with the target words formed the *related* distractor condition. For all monosyllabic targets, the related distractors did not follow the syllabification rules (48 CV and 19 CCV items). For the bisyllabic targets, 38 distractors coincided with the first syllable (5 CV, 33 CCV) and 16 distractors did not (12 CV, 4 CCV). For the trisyllabic targets, one coincided with the syllable boundary and one did not (both CV). To create the *unrelated* fragment pairings, we recombined the fragments from the related set with other target words. CV fragments were paired with CV-initial target words, and CCV with CCV-initial ones. For the fragment to be unrelated to the target, none of its phonemes should overlap with any phoneme of the target word. Moreover, the unrelated fragment could not be a part of a semantically related word, for example, the fragment /HO/ from the word *hond* (dog) could not be used as an unrelated distractor for the word *kat* (cat).[2]

Experiment 2B: We recorded the final fragments for a subset of target pictures and created the unrelated fragment pairings analogously to Experiment 2A. When defining the final fragment of the target word, we aimed to select the VC combination where possible (*bed* /ET/, bed; see Table 3). The next most common combination was CVC, where we included an additional consonant either because the vowel was the neutral schwa sound (*boter* /TəR/, butter) or because the VC fragment formed a real content or functional word (*banaan* /NA:N/, banana, *aan* is preposition "on"). Some words were removed because the overlap with real words could not be avoided (*boom* /O:M/, tree – *oom* /O:M/, uncle) or because the final fragment overlapped phonologically with the beginning of the word (e.g., /B/ is a phoneme shared by the beginning and the end fragments of *bamboe,* bamboo). This led to the inclusion of 105 items in the final set, with 55 monosyllabic (52.4% out of total number), 48 bisyllabic (45.7% out of total number) and 2 trisyllabic (1.9% out of total number) targets. After the duration manipulation shortening the fragments longer than 330 ms, the mean duration of the fragments was 313 ms (range 296–330 ms). In the related condition, none of the distractors used with monosyllabic targets followed syllabification rules (41 VC, 12 VCC, 2 CVC). For the bisyllabic targets, 30 distractors coincided with the second syllable (23 CVC, 5 CV, 1 CVCC, 1 VC) and 18 distractors did not (17 VC, 1 VCC). For the trisyllabic targets, one coincided with the last syllable (CVC) and one did not (VC). When combining fragments with unrelated words, eleven fragments were impossible to match with zero overlap to the target, however, the only overlapping phoneme in these target fragment pairs was the schwa sound (e.g., *bliksem*, lightning, related: /SəM/, unrelated: /TəR/).

Design and procedure. In each experiment, we used a 2×2 design, with the within-participant categorical variables *constraint* (constrained, unconstrained) and *distractor* (related, unrelated). Each participant saw each target picture in the constrained and unconstrained condition, however, whether the pictures were presented with a related or an unrelated fragment was counterbalanced across two lists. Specifically, if a certain target was paired with the related fragment in list 1, it was paired with the unrelated fragment in list 2. Furthermore, the two lists were separated into sublists A and B, where the constraint condition in which the participant encountered the target first was counterbalanced. This design resulted in four stimuli lists (1A, 1B, 2A, 2B). The presentation of the stimuli within each list was pseudorandomized using Mix (van Casteren & Davis, 2006), with a maximum of three successive repetitions of the same constraint and the same distractor type as well as a minimum distance of 20 between the same target pictures. Each participant received a unique list.

The experiment was programmed in Python and presented using Psychopy (version 2023.2.1; Peirce et al., 2019). Participants performed the task on a computer in the behavioural lab, using the same procedure as in Experiment 1. Trial structure was also kept identical, the only difference being the presentation of the auditory fragment simultaneously with the picture (Figure 2A). The additional task instructions informed the participants that they would hear a syllable which they should ignore. Furthermore, it was emphasized that the picture could already be named while the sound was playing to discourage participants from using the end of the sound rather than the presentation of the picture as the cue to start speaking. In Experiment 2A, 246 trials (61 or 62 per constraint by

Table 3. Types of auditory fragments in Experiments 2A and 2B with examples.

Target	*Related fragment*	*Related fragment type*	*Total number of fragments*	*Unrelated fragment with target of origin*
Experiment 2A: Initial fragments				
hond [dog]	/HO/	CV	67	/LA:/ laars [boot]
tafel [table]	/TA:/			/KO:/ konijn [rabbit]
slak [snail]	/SLA/	CCV	56	/BRY/ brug [bridge]
vlinder [butterfly]	/VLI/			/STA:/ staart [tail]
			Total: 123	
Experiment 2B: Final fragments				
boek [book]	/UK/	VC	60	/IL/ bril [glasses]
skelet [skeleton]	/ET/			/IR/ rivier [river]
gitaar [guitar]	/TA:R/	CVC	26	/ZEM/ bezem [broom]
helm [helmet]	/ELM/	VCC	13	/ANT/ krant [newspaper]
auto [car]	/TO:/	CV	5	/E:L/ kasteel [castle]
koffers [suitcases]	/FəRS/	CVCC	1	/ELM/ helm [helmet]
			Total: 105	

Translations from Dutch into English are indicated in square brackets.

distractor type) were presented in six blocks with self-paced rest periods in-between. In Experiment 2B, 210 trials (52 or 53 per constraint by distractor type) were presented in five blocks with self-paced rest periods. Each trial was recorded in its entirety using the microphone.

Data analysis. All analyses were conducted in R Statistical Software (v4.1.3, R Core Team 2022). The RT data and R code are available at https://osf.io/5k2qj/. RTs were calculated manually using PRAAT (Boersma & Weenink, 2013). In experiment 2A, the responses were marked as correct when the target word or a morphological derivate of the target word (e.g., diminutive suffix, *bonen* instead of *boontjes*) was produced. In experiment 2B, only the target responses were coded as correct since suffixes and alternative names (e.g., *sperziebonen* [green beans] instead of *boontjes* [beans]) would change the relation between the target and the fragment.

First, we conducted ex-Gaussian distributional analyses following Experiment 1. Then, we fitted a Bayesian hierarchical model per experiment using the same parameters as in Experiment 1. The dependent variable was the raw picture-naming RTs, while categorical independent variables of *constraint* (constrained, unconstrained), *distractor* (related, unrelated) and their *interaction* entered the model as fixed effects. By-subject and by-item intercepts entered the model as random effects. We used treatment coding with constrained condition and related fragment as the baseline. As in Experiment 1, we fit the models using the default prior (for the estimates produced with the narrow vs broad prior and the associated BFs, see Supplement S2A and S2B). The estimates for the level contrasts of interest were obtained from the model using R package *emmeans* (Lenth, 2023).

Results

Experiment 2A: The descriptive, ex-Gaussian and ES statistics can be found in Table 4. The overall effect of context was 182 ms, with a mean RT of 525 ms (SD = 204 ms) in the constrained and 707 ms (SD = 219 ms) in the unconstrained condition. Participants named pictures on average 167 ms faster in the related compared to the unrelated condition in unconstrained contexts, which serve as a baseline condition without preplanning. Participants also demonstrated a smaller but substantial phonological facilitation effect of 81 ms in the constrained condition.

Ex-Gaussian estimations converged on average within 16 iterations (range 2–102). In the unconstrained condition, the *mu* and *tau* parameters, describing the centrality of the bulk and tail distributions, respectively, were significantly different across distractor conditions. The *sigma* parameters, describing the spread of the bulk, were not significantly different. The ES was large

Table 4. Descriptive, ex-Gaussian and ES statistics as well as Bayesian model estimates for Experiment 2A with initial fragments.

		Unconstrained Condition		
Distractor type	*Mean (SD)*	*Mu*	*Sigma*	*Tau*
Related	626 (129)	538	60	81
Unrelated	793 (124)	585	63	201
Facilitation effect, unbiased d	*Mean (95% CI)* 1.3, [0.6, 1.9]	*Mu (95% CI)* 0.4, [−0.2, 1]		*Tau (95% CI)* 1.5, [0.9, 2.2]
	Bayesian estimates for phonological facilitation effect		*Median* 103	*95% HDI* [92, 114]
		Constrained Condition		
Distractor type	*Mean (SD)*	*Mu*	*Sigma*	*Tau*
Related	485 (124)	403	121	77
Unrelated	566 (158)	415	124	147
Facilitation effect, unbiased *d*	*Mean (95% CI)* 0.6, [0, 1.1]	*Mu (95% CI)* 0.07, [−0.5, 0.6]		*Tau (95% CI)* 0.8, [0.2, 1.4]
	Bayesian estimates for phonological facilitation effect		*Median* 38	*95% HDI* [26, 48]
		Dependent-samples t-test: related vs unrelated distractor		
Constraint		*Mu*	*Sigma*	*Tau*
Constrained		t(23)= 0.63, p = 0.54, CI [−28, 51]	t(23)= 0.23, p = 0.82, CI [−27, 34]	t(23) = 3.5, p <.01, CI [29, 111]
Unconstrained		t(23)= 3.7, p <.01, CI [21, 74]	t(23)= 0.28, p = 0.78, CI [−20, 27]	t(23) = 5.7, p <.001, CI [76, 163]

Note: CI = confidence interval, SD = standard deviation.
All values are in milliseconds, apart from *p* values and *unbiased d*, the Cohen's *d* estimates. Estimates of the phonological facilitation effect within condition with 95% highest density credible intervals (HDIs) were obtained with *emmeans* package.

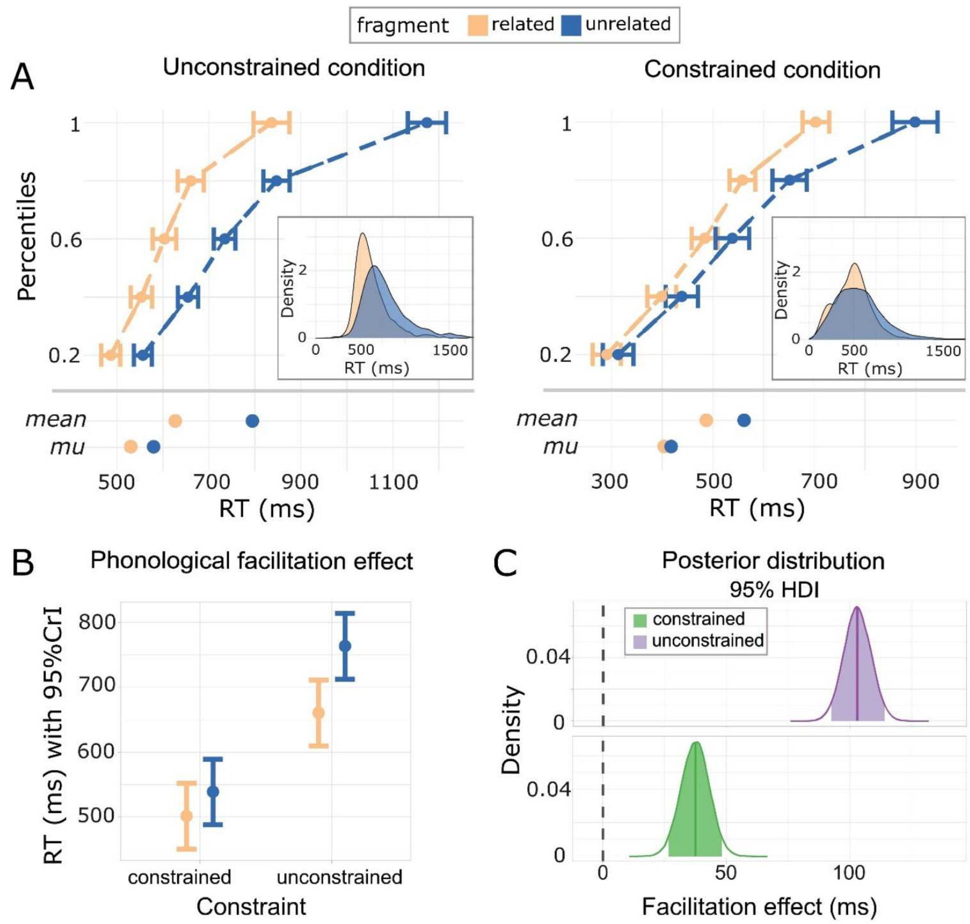

Figure 4. Cumulative distributions of RTs per condition and Bayesian modelling results for Experiment 2A. Reproduced with permission from the authors from doi:10.6084/m9.figshare.27645768. **A**. Cumulative (in quintiles) and density RT distributions for related and unrelated fragments with the mean and mu values within condition. **B**. Bayesian parameter estimates (mean with 95% CrI) for the interaction effect of fragment and condition. **C**. Posterior distributions for the phonological facilitation effect within condition, showing the median with 95% highest density credible intervals (HDI).

for the tail part of the distribution (d = 1.54), as well as for phonological facilitation calculated on the mean (d = 1.27) and *mu* (d = 0.4) indicating moderate to large effects (but note that the CI for *mu* contains zero). In the constrained condition, *mu* and *sigma* parameters describing the bulk of the related compared to unrelated distributions were not significantly different, whereas the tail *tau* parameter estimations were. The tail effect was large (d = 0.83), but, compared to the moderate ES on the mean (d = 0.55), the ES on the *mu* was 0.07, indicating that the effect disappears when only the bulk of the RT distribution is considered.

Taken together with the inspection of the cumulative distribution plot (Figure 4A), these results indicate that phonological facilitation effects observed in raw RTs were driven both by the bulk and the tail of the distributions in the unconstrained condition but were observed largely due to the tail effects in the constrained condition. The cumulative distribution plot of RTs demonstrated that facilitation was least prominent in the 20% fastest responses in both constraints (Figure 4A), with the smallest difference between the means in the constrained condition (mean facilitation effect of 30 ms and overlapping 95% confidence intervals, compared to 80 ms in the unconstrained condition).

In Bayesian modelling, MCM chains converged (R-hat values = 1) and the ESS values indicated stable parameter estimates (over > 10,000 as recommended by Kruschke, 2015). Following the mean-based tendencies of the distributions observed in the ex-Gaussian and ES analyses, the results (Table 4, Figure 4B and C) showed that phonological facilitation effects were credibly present in both conditions, with zeros not included into credible intervals of the difference between related and unrelated conditions in the unconstrained (95% HDI [92, 113.8]) as well as constrained (95% HDI [26, 48.4]) conditions.

Experiment 2B: The descriptive, ex-Gaussian, and ES statistics can be found in Table 5. The effect of context was 204 ms, with mean RT of 483 ms (SD = 206 ms) in the constrained and 687 ms (SD = 211 ms) in the

Table 5. Descriptive, ex-Gaussian, and ES statistics as well as Bayesian model estimates for Experiment 2B with final fragments.

		Unconstrained Condition		
Distractor type	*Mean (SD)*	*Mu*	*Sigma*	*Tau*
Related	622 (82)	538	66	77
Unrelated	755 (121)	588	86	156
Facilitation effect, unbiased d	*Mean (95% CI)* 1.25, [0.6, 2]	*Mu (95% CI)* 0.46, [−0.1 1]		*Tau (95% CI)* 1.15, [0.5, 1.8]
Bayesian estimates for phonological facilitation effect			*Median* 78	*95% HDI* [67, 89]
		Constrained Condition		
Distractor type	*Mean (SD)*	*Mu*	*Sigma*	*Tau*
Related	458 (107)	341	95	119
Unrelated	507 (135)	325	92	174
Facilitation effect, unbiased d	*Mean (95% CI)* 0.4, [−0.2, 1]	*Mu (95% CI)* −0.1, [−0.7, 0.5]		*Tau (95% CI)* 0.6, [0.1, 1.2]
Bayesian estimates for phonological facilitation effect			*Median* 11	*95% HDI* [−0.6, 22]
		Dependent-samples t-test: related vs unrelated distractor		
Constraint		*Mu*	*Sigma*	*Tau*
Constrained		t(23)= 1.1, *p* = 0.28, CI [−13, 44]	t(23)= 0.2, *p* = 0.84, CI [−31, 40]	t(23) = 2.5, *p* =.02, CI [9, 102]
Unconstrained		t(23)= 3.1, *p* <.01, CI [17, 84]	t(23)= 1.13, *p* = 0.27, CI [−16, 56]	t(23) = 3.9, *p* <.001, CI [37, 121]

Note: CI = confidence interval, SD = standard deviation.
All values are in millisecond, apart from *p* values and unbiased d, the Cohen's d estimates. Estimates of the phonological facilitation effect within condition with 95% highest density credible intervals (HDIs) were obtained with emmeans package.

unconstrained condition. Participants named pictures on average 133 ms faster in the related compared to unrelated condition in unconstrained contexts, indicating quite a substantial phonological facilitation effect. In the constrained condition, participants also demonstrated phonological facilitation, but the average effect was smaller at 49 ms.

Ex-Gaussian estimations converged on average within 14 iterations (range 3–84). The findings followed the patterns observed in Experiment 2A, with the phonological facilitation effect driven by both the bulk and tail of the RT distributions in the unconstrained condition, but mainly by the tail in the constrained condition, i.e., the trials without preplanning. In the constrained condition, the differences between related and unrelated distractors on the bulk of RTs suggested lack of phonological facilitation (*mean* difference = 49 ms, *mu* difference = −16 ms). The ES for phonological facilitation was moderate when calculated on the mean (d = 0.4, but CI contains zero) and disappeared when calculated on the *mu* (d = −0.1). These patterns are reflected in the cumulative distribution plot (Figure 5A, left panel), where the facilitation in the constrained condition is absent in the majority of the responses (except for the 40% slowest). Thus, the presence of facilitation at the aggregated group level was driven by the slower responses where – based on the RTs comparable to the unconstrained condition – participants likely did not preplan the target word and, hence, were facilitated by the related auditory fragment upon seeing the picture. In the unconstrained condition, facilitation can be seen starting already with the 20% fastest responses (Figure 5A, right panel).

In Bayesian modelling, MCM chains converged (R-hat = 1) and the ESS values indicated stable parameter estimates (over > 10,000). Following the findings of the ex-Gaussian and ES analyses, the modelling (Table 5, Figure 5B and C) showed that there was a credible phonological facilitation effect in the unconstrained condition (median 77.7 ms, 95% HPD [66.5, 88.8]). Likewise, there was phonological facilitation in the constrained condition, however, the effect was smaller and the credible interval included zero (95% HPD [−0.6, 22.1], posterior probability of values ≤ 0, i.e., of not observing facilitation, is 0.05%).

Discussion

The goal of Experiment 2 was to acquire further evidence for phonological preplanning in constrained contexts as well as to establish whether speakers plan beyond the initial phonemes of the target words. In both experiments, we observed a large phonological facilitation effect in the unconstrained condition. In this condition, since the preceding context is not semantically constraining, speakers can start planning the target word only upon picture presentation, similar to bare picture naming (Chupina et al., 2022; Roos et al., 2024). Hence, analogously to classic

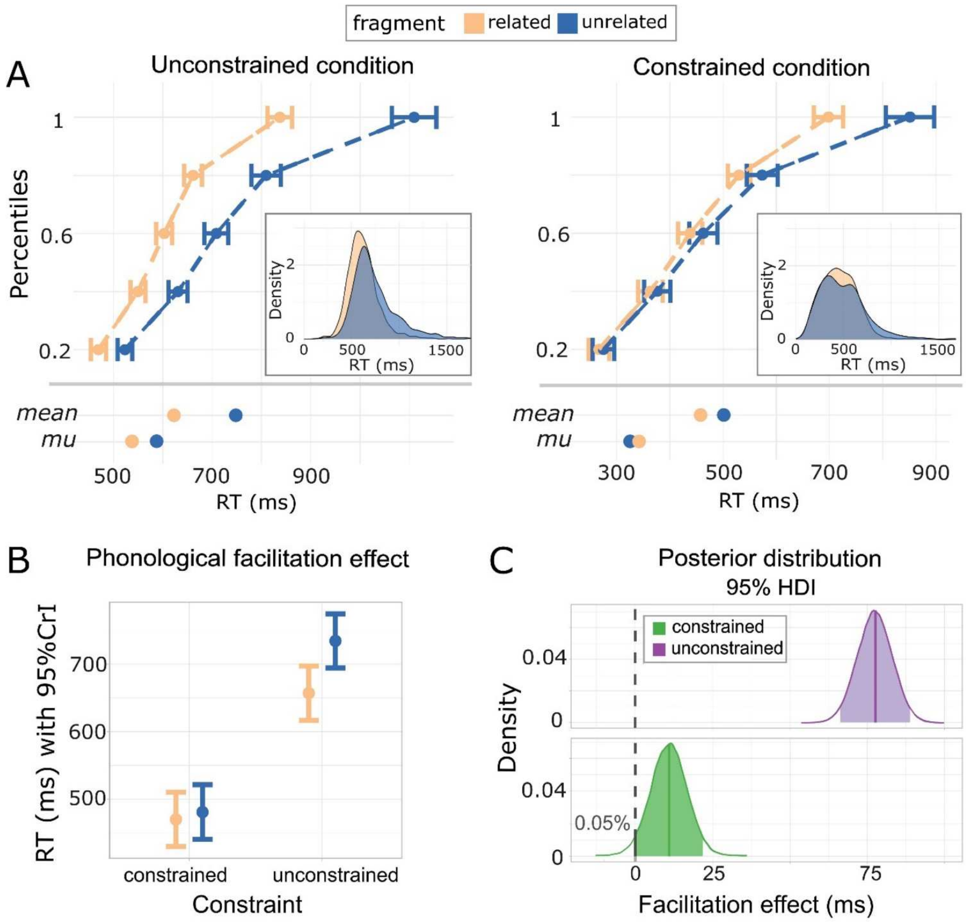

Figure 5. Distributions of RTs and Bayesian analyses results for Experiment 2B. Reproduced with permission from the authors from doi:10.6084/m9.figshare.27645777. **A**. Cumulative (in quintiles) and density RT distributions for constrained and unconstrained conditions with the mean and mu values. **B**. Bayesian model estimates (mean with 95% CrI) for the interaction effect of fragment and condition. **C**. Posterior distributions for the phonological facilitation effect within condition, showing the median with 95% highest density credible intervals (HDI). The area shaded in red shows the percentage of the distribution falling at and below zero, indicating the probability of the absence or the opposite direction of the effect.

PWI naming (e.g., Lupker, 1982), hearing the initial (Experiment 2A) or final (Experiment 2B) phonemes of the target word sped up its planning and led to shorter RTs. In the constrained contexts, assuming that speakers preplanned phonological representations, we expected phonological facilitation to be absent from RTs. While we still observed substantial phonological facilitation from the initial fragments in both experiments at the aggregated RT level, when considering the cumulative RT distribution, the shortest RTs showed little differences between related and unrelated fragments. Based on the cumulative RT distributions, there was little-to-no facilitation for around 20% of the responses in experiment 2A and at least 50% in experiment 2B. According to the ex-Gaussian and ES analyses in both experiments, the phonological facilitation effect was observed in the unconstrained condition throughout the whole distribution, i.e., from fastest to slowest responses. In the constrained condition, however, the presence of the effect was largely driven by the right tail of the distribution, i.e., the slowest responses. We interpret the responses in the constrained condition that are slower than or equally slow to the responses in the unconstrained condition as indicating lack of preplanning.

The facilitation effect was the smallest in the final-fragment experiment, with statistics showing uncertainty regarding the presence of the effect since the credible interval contained zero. These patterns indicate that speakers likely preplanned phonological information, namely at least the initial phonemes (2A) or the whole form (2B; Meyer et al., 2003), but only on a subset of trials. Such results are expected given the behavioural variability frequently demonstrated by speakers from trial to trial (e.g., Kilbourn-

Ceron & Goldrick, 2021). This variability often stems from the tail and not the central part of the RT distribution and can be attributed to a range of factors, including attentional fluctuations (Wundt, 1918, as cited in Roelofs & Piai, 2017), practice effects (Balota et al., 2008) or the ease of access to conceptual and syntactic representations due to familiarity from previous exposure (e.g., Konopka, 2012). If shorter naming latencies are a result of preplanning, trials with slower responses that fall within the range of RTs typical for bare picture naming (i.e., around or larger than 600 ms) are likely not prepared in advance. The interpretation that participants did not preplan words on multiple constrained trials, especially in experiment 2A, is supported by the fact that the RTs from the bulk of the constrained condition distribution (500 ms and above) were comparable to the fastest RTs in the unconstrained condition with unrelated fragments, where neither preplanning nor phonological facilitation occurred.

Another factor that could introduce variability to phonological planning is the length of the response (e.g., Levelt et al., 1999; see Experiment 1). In turn taking, for instance, speakers are known to show slower RTs when preplanning multi-word utterances (Corps & Pickering, 2024). Although we had both mono- and bisyllabic target words, since the word length mainly impacts the speed of the response and the speakers had at least 800 ms before the picture to retrieve information (which is sufficient for word-level preplanning, see General Discussion), the word length had no systematic impact on the RTs (see Supplement C, figure S3). Interestingly, speakers showed a larger facilitation effect (i.e., no phonological preplanning) from the initial rather than final fragments: Compared to the final-fragment experiment, a smaller percentage of the faster trials in the initial-fragment experiment showed no differences between related and unrelated fragments. This pattern indicates that participants likely engaged in preplanning on fewer trials during the initial-fragment experiment. While such differences can always stem from the inherent behavioural performance variability (see *General discussion*), speculatively, they can also be explained by speakers engaging in different word planning strategies driven by the experimental design. Despite the task and instructions being identical across the two experiments, the two participant samples were different, the materials did not fully overlap, and the initial-phoneme auditory fragments were on average 60 ms shorter. Furthermore, hearing the initial phonemes of the target word as opposed to the rhyming final phonemes might have encouraged participants to choose a different behavioural strategy. For instance, more participants might have not preplanned the word since hearing the initial sound later made planning considerably easier upon picture presentation. This speculation is consistent with participants being overall slower at naming pictures in the initial-phoneme experiment, especially in the constrained condition, with the *mu* estimates larger by 60–80 ms compared to the final-phoneme experiment.

To sum up, Experiment 2 provided further evidence for phonological preplanning of both initial and final phonemes of the target word. However, phonological information was preplanned only on a subset of semantically constraining trials, with more trials preplanned in the final-fragment experiment.

General discussion

During conversation, predictable contexts enable speakers to preplan responses before their speaking turn starts (Levinson & Torreira, 2015). While this advance planning in language interaction has been suggested to reach the word form stage of lexical access, that is, the generation of a phonological form and an articulatory motor programme (Bögels & Levinson, 2023; for review see Meyer, 2023), the exact extent of word preplanning in different communicative scenarios is not well understood. We conducted three behavioural experiments using the context-driven object naming paradigm in order to elucidate whether speakers preplan as far as phonological word form when they name pictures to complete semantically constrained sentences.

In Experiment 1, the analysis of existing RT data demonstrated that, in constrained contexts, response planning can start before participants see the target picture. Based on the tentative absence of the lexical frequency (Jescheniak & Levelt, 1994) and word length (Meyer et al., 2003) effects in the RTs measured from picture onset in constrained contexts, both lexical-semantic and phonological processes that are indexed by them seem to occur in general before picture presentation. In other words, compared to unconstrained contexts or bare picture naming, higher lexical frequency and shorter word length do not speed up word planning since speakers retrieve lexical-semantic and phonological information in advance, prior to picture presentation. This said, whereas our statistical modelling could not confidently confirm the presence of the effects in the unconstrained contexts, more certainty could be allocated to the absence of the word length effect in the constrained condition, indicating phonological and, hence, the preceding lexical-semantic preplanning.

In two follow-up experiments, we set out to obtain further evidence for phonological preplanning and to investigate whether only the initial phonemes (Experiment 2A) or the whole phonological word (Experiment 2B) are typically preplanned. In both experiments, we found large phonological facilitation effects in the unconstrained sentences: Hearing either the initial or the final phonemes of the target word simultaneously with picture presentation sped up naming, presumably by decreasing the time required for the encoding of the phonological word form (e.g., Meyer & Schriefers, 1991). In constrained contexts, if speakers preplan as far as the phonological form, no such facilitation should be expected. While we still observed a facilitation effect at the aggregated RT level in constrained sentences in both experiments, the effect was smaller than in the unconstrained condition and driven by the differences in the tail (*tau*) but not the bulk (*mu*) of the distributions. Specifically, the phonological facilitation effects disappeared when only faster responses (i.e., the bulk of the distributions) were compared. Speakers likely did not preplan responses on the slower trials, with a considerable number of RTs in the constrained condition comparable to those in the unconstrained condition. Furthermore, these slower trials disproportionately impacted the mean-based statistics. Overall, the PWI-like experiments with auditory fragments demonstrated that speakers preplan phonological representations of the whole word in semantically constraining contexts but not on each trial.

Utterance preplanning and its extent in turn-taking and context-driven picture naming

Corpus studies, showing that responses can start as fast as 200 ms after the end of the interlocutor's turn in conversation (e.g., Heldner & Edlund, 2010), constituted the basis for the theories of utterance preplanning (Levinson & Torreira, 2015). Although there is a discussion about the extent to which the gaps between speech segments in turn-taking can inform us about utterance planning (Corps et al., 2022), there is accumulating evidence for preplanning.

Similarly to turn-taking tasks, context-driven picture naming is able to induce early response planning in semantically constraining contexts (e.g., Piai et al., 2014). The preplanning that we observed was of a lexical-semantic nature, as suggested by the tentative lexical frequency effect (Experiment 1), in line with previous findings (Gastaldon et al., 2020; Piai et al., 2014). Furthermore, the evidence for the completion of word form encoding prior to picture onset came from the reliable absence of the word length effect (Experiment 1) together with the lack of phonological facilitation in the fastest RTs (Experiments 2A and 2B) only in the constrained condition. These findings are consistent with Barthel and Levinson's (2020) results, who, similarly manipulating the phonological relatedness of distractors, concluded that speakers can plan phonological information in advance. The experiment with initial-phoneme fragments (2A) indicated that when speakers respond fast, they have preplanned at least the first fragments of the phonological representation. It should be noted that we cannot claim that in Experiment 2A only the initial phonemes were planned in advance since we would see no facilitation from the initial fragments in both cases of only-initial and the whole-word phonological preplanning. The results of the experiment with final-phoneme fragments (2B) furthermore showed that participants did not experience facilitation from the final phonemes in semantically constraining contexts. Since the end of the word can only be encoded after the beginning (Meyer & Schriefers, 1991), this indicates that the whole word was at least phonologically (and potentially phonetically, see section *Loci of lexical and phonological effects*) encoded.

Although the present findings convincingly suggest phonological preplanning, the literature indicates that the presence and extent of early planning depends on a number of variables that are situation-specific and, thus, will likely differ from task to task. First of all, preplanning depends on time constraints imposed by the communicative context. For instance, preplanning of longer and more content-rich utterances will require more time, thus becoming impossible under certain time pressure conditions (Meyer et al., 2018; for review see Meyer, 2023). In our context-driven naming paradigm, the 800-ms interval before picture presentation provided sufficient time to preplan the word-level representations related to the sentential context before initiating articulation (Meyer et al., 2018). The assumption of word preplanning as opposed to simply faster planning after picture presentation was supported by the naming RT distribution shift towards earlier responses observed in the semantically constraining compared to unconstraining sentences (Experiment 1).

Another factor that impacts advance planning is the interplay of comprehension and production in conversation. While response planning can occur in parallel with listening to the interlocuter, there is a cost (e.g., Barthel & Sauppe, 2019): Both understanding and producing language require attentional resources, leading to a competition that results in slower and more error-prone production (Bögels et al., 2018; Meyer, 2023). In dual tasks such as speech planning while finger-tapping, advance planning has been shown to occur late, with speakers only initiating more cognitively demanding planning processes at the end of the interlocuter's turn (Sjerps & Meyer, 2015). During concurrent comprehension, speakers likely preplan the lexical-semantic information (i.e., content) but not the word form, postponing phonological encoding to the moment when they prepare to initiate articulation (Corps & Pickering, 2024). In the present study, however, the 800-ms interval after the lead-in sentence removed the necessity to plan while comprehending, thus potentially providing beneficial conditions for more extensive preplanning. Together with the relative simplicity of the sentence completion task, this is what might have allowed speakers to plan as far as the whole-word phonological representation (Experiment 2B). Furthermore, this makes our results compatible with the accounts suggesting that the size of the phonological planning unit is modulated by the task demands (Kilbourn-Ceron & Goldrick, 2021; Schriefers, 1999). Namely, in less cognitively demanding cases, speakers may afford to retrieve the whole-word phonological representation before articulation (Damian & Dumay, 2007; Levelt et al., 1999; Meyer et al., 2003; Roelofs, 1997), whereas articulation of complex utterances or utterances planned under cognitively demanding circumstances (e.g., dual tasks, time pressure) may already start when less is planned.

Finally, the extent of preplanning up to and including phonological encoding observed in the present study might have been impacted by the potential repetition priming effects arising due to the familiarization of participants with pictures before the experiment, as well as the repetition of the target words in two conditions. Stimulus familiarity is known to facilitate language planning: Concepts made more accessible via repetition priming or due to them being of higher lexical frequency are retrieved more quickly (Bock & Warren, 1985; Konopka, 2012).

To sum up, the presence and scope of preplanning varies depending on multiple factors, including time constraints, complexity and length of the planned response, familiarity with the topic and cognitive processing load. Importantly, our results reconcile the findings that speakers start preplanning as early as possible (Barthel et al., 2017; Bögels et al., 2015) with the suggestions that preplanning minimally overlaps with comprehension (Sjerps & Meyer, 2015) and mostly entails lexical-semantic processes (Corps & Pickering, 2024): Less cognitively demanding communicative situations seem to allow earlier and more advanced planning, while dividing cognitive resources between preplanning and other tasks makes speakers choose an easier late-planning strategy.

Behavioural variability in preplanning within task

The variables impacting the presence and scope of preplanning across tasks can also affect the consistency of an individual's performance in similar situations, i.e., across trials of an experimental task. For instance, concepts recently activated through familiarization will ease access to conceptual and syntactic representations (e.g., Konopka, 2012), thus facilitating preplanning on the trial. Attentional fluctuations while listening to lead-in sentences, by contrast, will preclude it and, thus, lead to slower RTs (Balota et al., 2008; Wundt, 1918, as cited in Roelofs & Piai, 2017). In Experiment 2, many responses in the constrained condition were as slow as those in the unconstrained condition, suggesting no preplanning on a (large) subset of trials. These patterns are consistent with inter-individual and inter-trial variability as well as the differential impact of task demands on behaviour described in the literature (Kilbourn-Ceron & Goldrick, 2021; Schriefers, 1999). The results of Experiment 2 are especially illustrative of how the underlying variability in the data might be masked by analyses relying on aggregated mean- or median-based statistics. While the facilitation effect was prominent at the aggregated RT level in the constrained condition, the cumulative distributions and ex-Gaussian analyses clearly demonstrated its absence in the bulk of RTs represented by the *mu* parameter. Since the cognitive effects of interest are often represented by the central part but not the slower tail of the RT distribution, this highlights the importance of exploring the whole distribution beyond the mean, for instance, using distributional analysis based on quantile estimation (e.g., Rousselet & Wilcox, 2020). Such more nuanced approaches can help clarify behavioural and even neural (see next section) activity patterns.

Implications for interpreting neural signatures of word retrieval

Manipulations of sentential context, that is, sentences with varying degrees of cloze probability, are widely used in both comprehension and production to study the neural underpinnings of language (e.g., DeLong et al., 2005; Gastaldon et al., 2020; Nieuwland et al., 2020; Piai et al., 2014). Furthermore, the sentence completion task that utilizes lexical predictability of the final word has been used to investigate the potentially facilitatory role of context in word retrieval in aphasia (with mixed results, e.g., Berndt et al., 2002; Chupina et al., 2022; Pashek & Tompkins, 2002). In context-driven picture naming, the difference in the constraining relative to unconstraining lead-in sentences manifests itself as alpha-beta power decreases (8–25 Hz) in the pre-picture window where speakers have time to integrate the sentential information and preplan the final word (e.g., Piai et al., 2014; see also e.g., León-Cabrera et al., 2022, and the studies reviewed there for similar effects in comprehension). Piai and colleagues (2014, 2020) have interpreted these power decreases as word preplanning. Their 2020 study demonstrated that presenting words from the same semantic category in the pre-picture interval delays the alpha-beta signatures, effectively creating an electrophysiological semantic interference effect, that is, pre-picture lexical-semantic processes get disrupted by the semantically related information.

Whereas, to the best of our knowledge, the phonological nature of the alpha-beta power decreases in context-driven picture naming still remains an open question, the present behavioural findings suggest that this electrophysiological signature might reflect phonological-level processes as well. Descriptively, the spatial distribution of the alpha-beta preplanning signatures peaking in the posterior temporal and inferior parietal regions (Roos et al., 2023; Roos & Piai, 2020) is highly similar to the increased ERP amplitude after critical word presentation in the preplanning vs no-preplanning condition reported by Bögels and colleagues (2015) and hypothesized by the authors to be of a phonological nature. Furthermore, phonological encoding is associated with the left inferior frontal gyrus (e.g., Indefrey & Levelt, 2004). If participants preplan the phonological form of the name of the picture before picture onset in the constrained condition, and after picture onset in the unconstrained condition, the left inferior frontal gyrus should be more active before picture onset in the constrained condition than in the unconstrained condition, and vice versa after picture onset. This is what Roos et al. (2023) observed in an fMRI study. Finally, the variable presence of preplanning from trial to trial observed in the current study could help explain why electrophysiological signatures of phonological preplanning might not be consistently detectable in neuroimaging studies (e.g., Jongman et al., 2020): Preplanning might either happen only on a smaller subset of trials, thus weakening the signal-to-noise ratio at the aggregated level, or preplanning might stop earlier during lexical access without reaching the phonological encoding stage.

Loci of lexical and phonological effects

Interestingly, despite the lexical frequency effect often being used to tap lexical access (Almeida et al., 2007) – which, more broadly, includes both lexical selection and word form generation (Levelt et al., 1999),– its exact locus is under debate. There is evidence that lexical frequency impacts the speed of either or both lemma selection and word form encoding stages (Corps & Meyer, 2023; Kittredge et al., 2008), the latter in the form of morphological (lexeme) encoding, according to attempts at a more fine-grained explanation (Jescheniak & Levelt, 1994). This means that previous interpretations of preplanning in semantically constraining contexts as being lexical-semantic in nature, including the interpretation of the electrophysiological signatures of the effect (Gastaldon et al., 2020; Piai et al., 2014; see *Implications for interpreting neural signatures of word retrieval* section for more detail), could be extended to the word form generation processes based on the lexical frequency effect alone.

In a similar manner, throughout the paper, we referred to the processes tapped by the word length and phonological facilitation effects as phonological encoding. While these effects are known to arise at the combinatorial word generation stage (as opposed to the "holistic" retrieval of the word representations from the memory at the earlier stages), similarly to the lexical frequency effect, their precise locus has not been clearly established. The effects could stem from either phonological or phonetic (e.g., syllable motor programmes) encoding stages or both combined (Levelt et al., 1999; Meyer et al., 2003). In case of word length, the effect could even stem, to a lesser extent, from morphological encoding (but not in case of facilitation since the auditory fragments were not constructed to follow morphological rules).

In the present study, we highlight the importance of employing multiple operationalizations when investigating a cognitive process. This could be of interest in the context of other well-established psycholinguistic effects, which are lacking consensus with regard to their loci. Here, we were able to provide overlapping evidence for phonological encoding in context-driven picture naming through the use of two differing operationalizations in Experiment 1 (word length in phonemes) and Experiment 2 (phonological facilitation), thus allowing for a larger degree of confidence regarding the generalizability of the findings.

Future directions

As argued earlier, the presence and extent of advance planning are impacted by a range of factors, with certain features of context-driven picture naming being atypical for a natural conversation where listening to the interlocutor and preplanning might have to co-occur (Levinson & Torreira, 2015; see Meyer, 2023, for discussion). Thus, future studies should directly investigate specific stages of utterance preplanning in different communicative scenarios and under different conditions, manipulating, amongst others, variables such as time allowed for preplanning, complexity and length of preplanned utterances and levels of attention allocated to the task. Furthermore, when studying the extent of preplanning, it is important to tap cognitive processes of interest using multiple operationalizations. Given that the exact locus of the commonly used psycholinguistic effects might be difficult to pinpoint, using a variety of measures to replicate the findings is necessary for providing convincing evidence for preplanning. Finally, to account for the variability typical for linguistic behaviour in language production, we advocate for future studies to analyse the whole RT distribution and to choose statistical approaches that can adequately capture the shape of the underlying distributions.

Conclusion

Rapid turn-taking in conversation suggests that speakers plan part of their turn in advance, but evidence for this is scarce. Previous studies provided evidence that constraining linguistic contexts facilitate advance response planning up to and including the word form generation stage. In the present study employing context-driven object naming, we obtained evidence that speakers preplan both lexical-semantic and phonological representations at the word level, including the phonological form of the whole word. Importantly, speakers did not preplan responses consistently on all semantically constraining trials, demonstrating variability in behaviour, likely determined by factors such as communicative demands and individual strategies. This finding highlights that approaches that take into account the whole RT distribution beyond classical aggregated measures of central tendency, such as the mean and the median, allow for more nuanced claims about word production behaviour.

Notes

1. While the precise age for ten participants is missing, it falls within 18–50 y.o. recruitment criterion.
2. This was the case for all but three unrelated pairings in each experiment (Experiment 2A: out of 123 total targets; Experiment 2B: out of 105), given the impossibility of achieving zero phonological overlap with the target otherwise.

Acknowledgements

We are grateful to Kirstin Bleumink and Rosemarije Weterings for recording the auditory stimuli, Yunzhi Luo and Nory te Brake for help with data collection, and to Hubert Voogd for reprogramming the experiments in Python.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was supported by the VIDI grant from the Netherlands Organization for Scientific Research (Nederlandse Organisatie voor Wetenschappelijk Onderzoek [NWO]) to V. P. (VI.Vidi.201.081).

Data availability statement

The data are available at https://osf.io/5k2qj/.

ORCID

Irina Chupina http://orcid.org/0000-0002-7534-6642

References

Almeida, J., Knobel, M., Finkbeiner, M., & Caramazza, A. (2007). The locus of the frequency effect in picture naming: When recognizing is not enough. *Psychonomic Bulletin & Review*, *14*(6), 1177–1182. https://doi.org/10.3758/BF03193109

Audacity Team. (2014). Audacity®: Free Audio Editor and Recorder [Computer program]. http://audacity.sourceforge.net/

Balota, D. A., & Yap, M. J. (2011). Moving beyond the mean in studies of mental chronometry: The power of response time distributional analyses. *Current Directions in Psychological Science*, *20*(3), 160–166. https://doi.org/10.1177/0963721411408885

Balota, D. A., Yap, M. J., Cortese, M. J., & Watson, J. M. (2008). Beyond mean response latency: Response time distributional analyses of semantic priming. *Journal of Memory and Language*, *59*(4), 495–523. https://doi.org/10.1016/j.jml.2007.10.004

Barthel, M., & Levinson, S. C. (2020). Next speakers plan word forms in overlap with the incoming turn: Evidence from gaze-contingent switch task performance. *Language, Cognition and Neuroscience*, *35*(9), 1183–1202. https://doi.org/10.1080/23273798.2020.1716030

Barthel, M., Meyer, A. S., & Levinson, S. C. (2017). Next speakers plan their turn early and speak after turn-final "go-signals". *Frontiers in psychology*, *8*, 393. https://doi.org/10.3389/fpsyg.2017.00393

Barthel, M., & Sauppe, S. (2019). Speech planning at turn transitions in dialog is associated with increased processing load. *Cognitive Science*, *43*(7), e12768. https://doi.org/10.1111/cogs.12768

Berndt, R. S., Burton, M. W., Haendiges, A. N., & Mitchum, C. C. (2002). Production of nouns and verbs in aphasia: Effects of elicitation context. *Aphasiology*, *16*(1–2), 83–106. https://doi.org/10.1080/02687040143000212

Bi, Y., Xu, Y., & Caramazza, A. (2009). Orthographic and phonological effects in the picture–word interference paradigm: Evidence from a logographic language. *Applied Psycholinguistics*, *30*(4), 637–658. https://doi.org/10.1017/S0142716409990051

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67. https://doi.org/10.1016/0010-0277(85)90023-X

Boersma, P., & Weenink, D. (2013). Praat: Doing phonetics by computer [computer program]. Version 5.3. 51. http://www. praat.org

Bögels, S., Casillas, M., & Levinson, S. C. (2018). Planning versus comprehension in turn-taking: Fast responders show reduced anticipatory processing of the question. *Neuropsychologia*, *109*, 295–310. https://doi.org/10.1016/j.neuropsychologia.2017.12.028

Bögels, S., & Levinson, S. C. (2023). Ultrasound measurements of interactive turn-taking in question-answer sequences: Articulatory preparation is delayed but not tied to the response. *PLoS One*, *18*(7), e0276470. https://doi.org/10.1371/journal.pone.0276470

Bögels, S., Magyari, L., & Levinson, S. C. (2015). Neural signatures of response planning occur midway through an incoming question in conversation. *Scientific Reports*, *5*(1), 12881. https://doi.org/10.1038/srep12881

Brown, S., & Heathcote, A. (2003). QMLE: Fast, robust, and efficient estimation of distribution functions based on quantiles. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 485–492. https://doi.org/10.3758/BF03195527

Brysbaert, M., Mandera, P., & Keuleers, E. (2018). The word frequency effect in word processing: An updated review. *Current Directions in Psychological Science*, *27*(1), 45–50. https://doi.org/10.1177/0963721417727521

Bürkner, P.-C. (2017). **brms**: An *R* Package for Bayesian multilevel models using *Stan*. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Chupina, I., Piai, V., & Westner, B. U. Manuscript in preparation.

Chupina, I., Sierpowska, J., Zheng, X. Y., Dewenter, A., Piastra, M. C., & Piai, V. (2022). Time course of right-hemisphere recruitment during word production following left-hemisphere damage: A single case of young stroke. *European Journal of Neuroscience*, *56*(8), 5235–5259. https://doi.org/10.1111/ejn.15813

Corps, R. E., Knudsen, B., & Meyer, A. S. (2022). Overrated gaps: Inter-speaker gaps provide limited information about the timing of turns in conversation. *Cognition*, *223*, 105037. https://doi.org/10.1016/j.cognition.2022.105037

Corps, R. E., & Meyer, A. S. (2023). Word frequency has similar effects in picture naming and gender decision: A failure to replicate Jescheniak and Levelt (1994). *Acta Psychologica*, *241*, 104073. https://doi.org/10.1016/j.actpsy.2023.104073

Corps, R. E., & Pickering, M. J. (2024). The role of answer content and length when preparing answers to questions. *Scientific Reports*, *14*(1), 17110. https://doi.org/10.1038/s41598-024-68253-6

Cumming, G. (2013). *Understanding the new statistics: Effect sizes, confidence intervals, and meta-analysis* (1st ed.). Routledge. https://doi.org/10.4324/9780203807002

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

Damian, M., & Dumay, N. (2007). Time pressure and phonological advance planning in spoken production. *Journal of Memory and Language*, *57*(2), 195–209. https://doi.org/10.1016/j.jml.2006.11.001

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological review*, *93*(3), 283. https://doi.org/10.1037/0033-295X.93.3.283

DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*(8), 1117–1121. https://doi.org/10.1038/nn1504

Donders, F. C. (1968). Over de snelheid van psychische processen. *Onderzoekingen gedaan in het Physiologisch Laboratorium der Utrechtsche Hoogeschool (1868–1869)*, *2*, 92–120.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816. https://doi.org/10.1080/17470218.2017.1310261

Gastaldon, S., Arcara, G., Navarrete, E., & Peressotti, F. (2020). Commonalities in alpha and beta neural desynchronizations during prediction in language comprehension and production. *Cortex*, *133*, 328–345. https://doi.org/10.1016/j.cortex.2020.09.026

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis*. *Chapman and Hall/CRC*. https://doi.org/10.1201/b16018

Griffin, Z. M., & Bock, K. (1998). Constraint, word frequency, and the relationship between lexical processing levels in spoken word production. *Journal of Memory and Language*, *38*(3), 313–338. https://doi.org/10.1006/jmla.1997.2547

Heldner, M., & Edlund, J. (2010). Pauses, gaps and overlaps in conversations. *Journal of Phonetics*, *38*(4), 555–568. https://doi.org/10.1016/j.wocn.2010.08.002

Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*(1–2), 101–144. https://doi.org/10.1016/j.cognition.2002.06.001

Jescheniak, J. D., & Levelt, W. J. M. (1994). Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*(4), 824–843. https://doi.org/10.1037/0278-7393.20.4.824

Jongman, S. R., Piai, V., & Meyer, A. S. (2020). Planning for language production: The electrophysiological signature of attention to the cue to speak. *Language, Cognition and Neuroscience*, *35*(7), 915–932. https://doi.org/10.1080/23273798.2019.1690153

Kelley, K., Kelley, M. K., & Imports, M. A. S. S. (2018). The MBESS R Package. Computer software]. MBESS. Retrieved from: https://CRAN. R-project. org/package = MBESS

Keuleers, E., Brysbaert, M., & New, B. (2010). SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods*, *42*(3), 643–650. https://doi.org/10.3758/BRM.42.3.643

Kilbourn-Ceron, O., & Goldrick, M. (2021). Variable pronunciations reveal dynamic intra-speaker variation in speech planning. *Psychonomic Bulletin & Review*, *28*(4), 1365–1380. https://doi.org/10.3758/s13423-021-01886-0

Kittredge, A. K., Dell, G. S., Verkuilen, J., & Schwartz, M. F. (2008). Where is the effect of frequency in word production? Insights from aphasic picture-naming errors. *Cognitive Neuropsychology*, *25*(4), 463–492. https://doi.org/10.1080/02643290701674851

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*(1), 143–162. https://doi.org/10.1016/j.jml.2011.08.003

Kruschke, J. K. (2010). What to believe: Bayesian methods for data analysis. *Trends in Cognitive Sciences*, *14*(7), 293–300. https://doi.org/10.1016/j.tics.2010.05.001

Kruschke, J. K. (2015). *Doing Bayesian data analysis: A tutorial with R, JAGS, and Stan* (2nd ed.). Academic Press.

Kruschke, J. K., & Liddell, T. M. (2018). The Bayesian new statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review*, *25*(1), 178–206. https://doi.org/10.3758/s13423-016-1221-4

Lampe, L. F., Hameau, S., & Nickels, L. (2023). Are they really stronger? Comparing effects of semantic variables in speeded deadline and standard picture naming. *Quarterly Journal of Experimental Psychology*, *76*(4), 762–782. https://doi.org/10.1177/17470218221103356

Lenth, R. V. (2023). emmeans: Estimated marginal means, aka Least-Squares Means. R package version 1.8.5. https://CRAN.R-project.org/package = emmeans

León-Cabrera, P., Piai, V., Morís, J., & Rodríguez-Fornells, A. (2022). Alpha power decreases associated with prediction in written and spoken sentence comprehension. *Neuropsychologia*, *173*, 108286. https://doi.org/10.1016/j.neuropsychologia.2022.108286

Levelt, W. J., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*(1), 1–38. https://doi.org/10.1017/S0140525X99001776

Levinson, S. C. (2016). Turn-taking in human communication – origins and implications for language processing. *Trends in Cognitive Sciences*, *20*(1), 6–14. https://doi.org/10.1016/j.tics.2015.10.010

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*, 713. https://doi.org/10.3389/fpsyg.2015.00731

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171

Lupker, S. J. (1982). The role of phonetic and orthographic similarity in picture–word interference. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *36*(3), 349–367. https://doi.org/10.1037/h0080652

Marian, V., Bartolotti, J., Chabal, S., & Shook, A. (2012). CLEARPOND: Cross-linguistic easy-access resource for phonological and orthographic neighborhood densities. *PLoS One*, *7*(8), e43230. https://doi.org/10.1371/journal.pone.0043230

Meyer, A. S. (2023). Timing in conversation. *Journal of Cognition*, *6*(1), 20. https://doi.org/10.5334/joc.268

Meyer, A. S., Alday, P. M., Decuyper, C., & Knudsen, B. (2018). Working together: Contributions of corpus analyses and experimental psycholinguistics to understanding conversation. *Frontiers in Psychology*, *9*, 525. https://doi.org/10.3389/fpsyg.2018.00525

Meyer, A. S., Roelofs, A., & Levelt, W. J. M. (2003). Word length effects in object naming: The role of a response criterion. *Journal of Memory and Language*, *48*(1), 131–147. https://doi.org/10.1016/S0749-596X(02)00509-0

Meyer, A. S., & Schriefers, H. (1991). Phonological facilitation in picture-word interference experiments: Effects of stimulus onset asynchrony and types of interfering stimuli. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *17*(6), 1146. https://doi.org/10.1037/0278-7393.17.6.1146

Nardo, D., Holland, R., Leff, A. P., Price, C. J., & Crinion, J. T. (2017). Less is more: Neural mechanisms underlying anomia treatment in chronic aphasic patients. *Brain*, *140*(11), 3039–3054. https://doi.org/10.1093/brain/awx234

Neurobehavioral Systems, Inc. (2023). *Presentation® [Computer software]*. www.neurobs.com

Nieuwland, M. S., Arkhipova, Y., & Rodríguez-Gómez, P. (2020). Anticipating words during spoken discourse comprehension: A large-scale, pre-registered replication study using brain potentials. *Cortex*, *133*, 1–36. https://doi.org/10.1016/j.cortex.2020.09.007

Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences on lexical retrieval in aphasia. *Aphasiology*, *16*(3), 261–286. https://doi.org/10.1080/02687040143000573

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

Piai, V., Klaus, J., & Rossetto, E. (2020). The lexical nature of alpha-beta oscillations in context-driven word production. *Journal of Neurolinguistics*, *55*, 100905. https://doi.org/10.1016/j.jneuroling.2020.100905

Piai, V., Roelofs, A., & Maris, E. (2014). Oscillatory brain responses in spoken word production reflect lexical frequency and sentential constraint. *Neuropsychologia*, *53*, 146–156. https://doi.org/10.1016/j.neuropsychologia.2013.11.014

Rabovsky, M., Schad, D. J., & Abdel Rahman, R. (2016). Language production is facilitated by semantic richness but inhibited by semantic density: Evidence from picture naming. *Cognition*, *146*, 240–244. https://doi.org/10.1016/j.cognition.2015.09.016

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Roelofs, A. (1997). The WEAVER model of word-form encoding in speech production. *Cognition*, *64*(3), 249–284. https://doi.org/10.1016/S0010-0277(97)00027-9

Roelofs, A. (2018). One hundred fifty years after Donders: Insights from unpublished data, a replication, and modeling of his reaction times. *Acta Psychologica*, *191*, 228–233. https://doi.org/10.1016/j.actpsy.2018.10.002

Roelofs, A. (2021). Phonological cueing of word finding in aphasia: Insights from simulations of immediate and treatment effects. *Aphasiology*, *35*(2), 169–185. https://doi.org/10.1080/02687038.2019.1686748

Roelofs, A., & Piai, V. (2017). Distributional analysis of semantic interference in picture naming. *Quarterly Journal of Experimental Psychology*, *70*(4), 782–792. https://doi.org/10.1080/17470218.2016.1165264

Roos, N. M., Chauvet, J., & Piai, V. (2024). The Concise Language Paradigm (CLaP), a framework for studying the intersection of comprehension and production: Electrophysiological properties. *Brain Structure and Function*, 1–17. https://doi.org/10.1007/s00429-024-02801-8

Roos, N. M., & Piai, V. (2020). Across-session consistency of context-driven language processing: A magnetoencephalography study. *European Journal of Neuroscience*, *52*(5), 3457–3469. https://doi.org/10.1111/ejn.14785

Roos, N. M., Takashima, A., & Piai, V. (2023). Functional neuroanatomy of lexical access in contextually and visually guided spoken word production. *Cortex*, *159*, 254–267. https://doi.org/10.1016/j.cortex.2022.10.014

Rousselet, G. A., & Wilcox, R. R. (2020). Reaction times and other skewed distributions problems with the mean and the median. *Meta-psychology*, *4*. https://doi.org/10.15626/MP.2019.1630

Schriefers, H. (1999). Phonological facilitation in the production of two-word utterances. *European Journal of Cognitive Psychology*, *11*(1), 17–50. https://doi.org/10.1080/713752301

Shao, Z., & Rommers, J. (2020). How a question context aids word production: Evidence from the picture–word interference paradigm. *Quarterly Journal of Experimental Psychology*, *73*(2), 165–173. https://doi.org/10.1177/1747021819882911

Sjerps, M. J., & Meyer, A. S. (2015). Variation in dual-task performance reveals late initiation of speech planning in turn-taking. *Cognition*, *136*, 304–324. https://doi.org/10.1016/j.cognition.2014.10.008

Starreveld, P. A. (2000). On the interpretation of onsets of auditory context effects in word production. *Journal of Memory and Language*, *42*(4), 497–525. https://doi.org/10.1006/jmla.1999.2693

Taylor, W. L. (1953). "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly*, *30*(4), 415–433. https://doi.org/10.1177/107769905303000401

van Casteren, M., & Davis, M. H. (2006). Mix, a program for pseudorandomization. *Behavior Research Methods*, *38*(4), 584–589. https://doi.org/10.3758/BF03193889

Wagenmakers, E.-J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., Love, J., Selker, R., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Rouder, J. N., & Morey, R. D. (2018). Bayesian inference for psychology. Part I: Theoretical advantages and practical ramifications. *Psychonomic Bulletin & Review*, *25*(1), 35–57. https://doi.org/10.3758/s13423-017-1343-3

Westner, B. U., Luo, Y., & Piai, V. Manuscript in preparation.