# Advance planning in written and spoken sentence production

Jens Roeser, Mark Torrance, Thom Baguley

Department of Psychology, Nottingham Trent University, Nottingham, UK

Author Note

Correspondance concerning this article should be addressed to Jens Roeser, 50 Shakespeare Street, Nottingham NG1 4FQ, UK. E-mail: jens.roeser@ntu.ac.uk

Data and analysis: figshare.com/articles/Analysis/5097403

## Abstract

Response onset latencies for sentences that start with a conjoined noun phrase are typically longer than for sentences starting with a simple noun phrase. This suggests that advance planning has phrasal scope, which may or may not be lexically driven. All previous studies have involved spoken production, leaving open the possibility that effects are, in part, modality-specific. In three image-description experiments ($N$s = 32) subjects produced sentences with conjoined (e.g. *Peter and the hat*) and simple initial noun phrases (e.g. *Peter*) in both speech and writing. Production onset latencies and participants' eye movements were recorded. Ease of lexical retrieval of sentences' second noun was assessed by manipulating codability (Experiment 1) and by gaze-contingent name priming (Experiments 2 and 3). Findings confirmed a modality-independent phrasal scope for advance planning but did not support obligatory lexical retrieval beyond the sentence-initial noun. This research represents the first direct experimental comparison of sentence planning in speech and writing.

*Keywords:* Grammatical encoding; planning scope; language production; sentence processing; eye tracking

Advance planning in written and spoken sentence production

## Introduction

People can prepare entire sentences carefully in advance before they address an interlocutor or audience. However, sentences are typically planned in smaller units and planning is incomplete at speech onset. Some advance planning is obligated prior to output, while other planning can be delayed (V. S. Ferreira & Slevc, 2007; Levelt, 1989). The generation of this advance-planning unit requires planning on various levels. This paper addresses the minimum linguistic processing requirements to initiate sentence output.

A number of studies have examined advance planning in language production (Allum & Wheeldon, 2007, 2009; F. Ferreira, 1991; Griffin, 2001; Hardy, Segaert, & Wheeldon, 2018; Konopka, 2012; Konopka & Bock, 2009; E.-K. Lee, Brown-Schmidt, & Watson, 2013; Levelt & Maasen, 1981; Martin, Crowther, Knight, Tamborello II, & Yang, 2010; Martin, Yan, & Schnur, 2014; Meyer, 1996; Smith & Wheeldon, 1999, 2001, 2004; Swets, Jacovina, & Gerrig, 2014; Wagner, Jescheniak, & Schriefers, 2010; Wheeldon, Ohlson, Ashby, & Gator, 2013; Zhao & Yang, 2013, 2016). Conclusions concerning the minimal planning unit obligated by the language system have been mixed. Some authors conclude that sentence initiation requires only the first determiner-noun pair (e.g. Griffin, 2001; Zhao & Yang, 2016) or less (Bürki, Sadat, Dubarry, & Alario, 2016). Others suggest that the minimum planning unit comprises the smallest full phrase embracing the first nominal head (Schriefers & Teruel, 1999), the first thematically functional unit (Allum & Wheeldon, 2007, 2009; Zhao & Yang, 2013), the first noun phrase (e.g. Konopka, 2012; Martin et al., 2010; Smith & Wheeldon, 1999), or the entire clause (e.g. Lindsley, 1975; Meyer, 1996; Smith & Wheeldon, 2004).

These studies all involved participants generating spoken response and tend to use the terms language and speech interchangeably. Speech is not, however, the only way in which language can be output, and brings with it certain pragmatic and cognitive constraints that may not be shared by other output modalities (writing, signing), and are independent of underlying constraints imposed by the language production system. The research reported in this study addressed whether findings concerning the planning scope tested in the spoken domain also hold true for writing. Studying written sentence production is valuable in this context because (a) extending language production models to written production is, in itself, a worthwhile goal and (b) triangulating effects from speech and writing helps to disentangle general linguistic effects from modality-specific pragmatic effects. Knowing that sentence planning scope effects replicate in speech and writing provides evidence in support of the argument that these effects are fundamental to the language production system. We will address possible differences between the spoken and written modality first, before we turn to conclusions of planning scope debate that might be the product of the bias towards the spoken domain.

Alario, Costa, Ferreira, and Pickering (2006, p. 783–784) highlighted the modality bias in language production research and stressed the importance of studies on writing (and sign language) in developing a complete model of language production. Planning mechanisms in written sentence production have, however, been almost entirely neglected by cognitive-experimental researchers. It seems probable that writing and speech employ the same syntactic processing system, and this claim finds support in studies that show cross-modal syntactic priming(Branigan, Pickering, & Cleland, 1999; Cleland & Pickering, 2006; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008; Hartsuiker & Westenberg, 2000). Writing, however, differs from speech in a number of ways. Researchers has argued that both hand-writing (Van Galen, 1991) and keyboard-typing (Gentner, 1982; Rumelhart & Norman, 1982) are cognitive skills that are importantly different from and cannot be understood as a simple extension of speech. Writing a sentence typically takes more time than it it were spoken. Writing involves orthographic retrieval in addition to (or in substitution for) phonological retrieval.[1] Motor planning associated with typed output is learned later and probably never achieves the effortlessness of articulation (Gentner, Larochelle, & Grudin, 1988; Olive, 2014). This suggests either an increased need for mental buffering and/or benefits of shorter planning scope. Persistent visual feedback from the unfolding text is usually available in writing. This means that reviewing output can, in principle, be delayed. In speech, however, acoustic feedback is transitory and monitoring must occur at the time of utterance. Of particular relevance in the present context, spoken communication requires greater output fluency. Pauses in speech have a communicational effect (Clark & Fox Tree, 2002) and therefore hesitation in spoken output has potential implications for listeners' understanding and interpretation of the message. Speakers' advance planning is therefore potentially affected by the need to minimise intra-sentence pausing once speaking commences (Levelt & Meyer, 2000; Meyer, 1997), over and above the demands of (modality independent) language processes (see Allum & Wheeldon, 2007; Griffin, 2003). By contrast, hesitation in

---

[1] There is an extensive debate whether orthographic representations are activated via a phonological (e.g. Bonin & Fayol, 2000; Nottbusch, Grimm, Weingarten, & Will, 2005; Zhang & Damian, 2010) or a lexical route without access to phonology (e.g. Bonin, Fayol, & Gombert, 1998; Rapp, Benzing, & Caramazza, 1997; Sahel, Nottbusch, Blanken, & Weingarten, 2005) which can be characterised as a dual-route process (e.g. Barry, 1994; Damian, Dorjee, & Stadthagen-Gonzalez, 2011; Qu & Damian, 2017) in which phonology may or may not serve as a mediator.

(1) a. The dog and the foot move above the kite.

b. The dog moves above the foot and the kite.

written production, in most contexts, has no bearing on the text's eventual communicational effect. Writers are therefore free to hesitate and plan without concern for how this will affect their perceived message.

We argue, therefore, that there is at least the possibility that findings from previous research examining advance planning scope in the production of simple sentences result in part from speech-specific processes, and therefore that triangulation with written production is valuable. Existing, spoken production research has shown that the extent of pre-sentence planning depends on the structure that is being planned. Several researchers have suggested that advance syntactic planning scopes over the first verb-argument phrase (e.g. Martin et al., 2010, 2014; Wagner et al., 2010; Wheeldon et al., 2013). For example Smith and Wheeldon (1999) manipulated the syntactic complexity of sentence-initial subject noun phrase. Participants were presented with arrays of three images which then moved in opposite directions to elicited sentences with either a complex, conjoined subject noun phrase as in example (1a) or a simple subject noun phrase as in example (1b) while the overall complexity of the stimulus array and the target sentence were held constant.

They found longer sentence onset latency for sentences with complex NPs (for similar effects see Martin et al., 2010, 2014; Wheeldon et al., 2013). There is strong evidence, however, that advance lexical processing is restricted to sentence-initial nouns and does not extend to subsequent nouns in the same NP (Allum & Wheeldon, 2009; Griffin, 2001; Konopka, 2012; Zhao & Yang, 2013, 2016). For instance, Griffin (2001) elicited sentences with subject phrases similar to those in example (1a). She manipulated the frequency of all image names and the codability of the second and third image name: frequency was used to manipulate difficulty of phonological encoding and codability (the number of names associated with an image) to manipulate difficulty of lexical selection. She found evidence for lexical preparation of the sentence-initial noun but no effects on later nouns. Similarly, Zhao and Yang (2016) presented evidence from event-related potentials showing semantic blocking effects for sentence-initial nouns only.

In comparison to speech, written production in general, and written sentence production in particular, has received relatively little attention from cognitive-experimental researchers. There is evidence that written picture-naming latencies are affected both by factors that affect spoken picture-naming latency (image familiarity and agreement, word frequency and age-of-acquisition) and by orthographic features that are specific to the written form of the name (Bonin, Méot, Laroche, Bugaiska, & Perret, 2017; Bonin, Peereman, & Fayol, 2001; Kandel & Perret, 2015; Perret, Bonin, & Laganaro, 2014; Scaltritti, Arfé, Torrance, & Peressotti, 2016; Torrance et al., 2017). At above the word level, studies of spontaneous, multi-sentence text production – in the tradition of analyses of pause patterns in spoken monologues (e.g. Butterworth, 1975; Henderson, Goldman-Eisler, & Skarbek, 1966) – suggest a greater tendency to pause at sentence and clause boundaries than before or within mid-clause words (Ailhaud & Chenu, 2018; Foulin, 1998; Immonen & Mäkisalo, 2017; Kaufer, Hayes, & Flower, 1986; Van Hell, Verhoeven, & Van Beijsterveldt, 2008). There is evidence that probability and duration of eye-movement back into already-produced text follows a similar pattern (Chukharev-Hudilainen, Saricaoglu, Torrance, & Feng, 2018; Torrance, Johansson, Johansson, & Wengelin, 2016). Controlled experimental studies that have explored written production of unrelated and syntactically-related word pairs (adjective-noun phrases) suggest that, in both cases, these are planned as a unit in advance of output (Damian & Stadthagen-Gonzalez, 2009), although there is some evidence that planning of the second noun is less complete at production onset for the second noun in unrelated noun-noun pairs (Bonin, Malardier, Méot, & Fayol, 2006).

To our knowledge research exploring planning scope in written sentence production is limited to three papers reporting preliminary findings (Nottbusch, 2010; Nottbusch, Weingarten, & Sahel, 2007; Torrance & Nottbusch, 2012). Nottbusch et al. (2007) and Nottbusch (2010) found evidence for increased sentence-initial planning time associated with producing noun phrases with a prepositional phrase modifier, than to producing conjoined noun phrases. Interestingly, Allum and Wheeldon (2007, 2009) found the opposite pattern for speech. This may have been due to a number of factors, including experimental design and language tested. This effect is at least consistent with the possibility that spoken production may result in increased planning scope. Torrance and Nottbusch (2012) describe an additional preliminary study comparing writing and speech in an experimental paradigm similar to that used by Griffin (2001). Findings paralleled those of Griffin in spoken production, with eye movement evidence suggesting planning scope rarely extended beyond the first noun of the sentence. This effect was also present, and was stronger, when sentences were written.

An additional specific question, which again might plausibly have different answers depending on whether output is spoken or written, concerns the extent to which planning beyond a single noun, in particular in the case of coordinated noun phrases, entails lexical encoding of a second noun. Extensive research in the spoken domain, and preliminary findings on the planning scope in writing appear to suggest a planning-scope hierarchy, with syntax planning extending over the sentence-initial verb-argument phrase, and lexical planning proceeding incrementally, on a word-by-word basis.

Several studies provide fairly direct evidence that, in speech, syntactic structure does not rely on lexical specification (Allum & Wheeldon, 2007, 2009; Konopka & Bock, 2009; E.-K. Lee et al., 2013; Wheeldon, 2011, 2012; Wheeldon et al., 2013; Wheeldon, Smith, & Apperly, 2011). In Wheeldon et al. (2013) the authors used the Smith and Wheeldon (1999) design described above, but allowed participants to preview images representing either the second or third noun. These nouns were either within or outside of the sentence-initial phrase. If syntactic planning is lexically-mediated, the phrasal scope effect should be modulated by preview for images that are named as part of the sentence-initial phrase. However, this was not what they found. No preview benefit was observed for the third noun, regardless of its syntactic position. Preview benefit was found for the second noun as part of the sentence-initial phrase only. The authors concluded that phrasal scope limits but does not require lexical activation. Allum and Wheeldon (2007, 2009) found consistently longer latencies for conjoined noun phrases compared to noun phrases with prepositional phrases in both head-initial and head-final languages. They concluded that the linearisation of lexical items in noun phrases with prepositional phrase modifiers is syntactically determined, while the order of nouns in complex noun phrases is arbitrary and requires lexical buffering. These findings suggest that syntactic planning guides lexical activation (for further evidence see e.g. Konopka & Bock, 2009; E.-K. Lee et al., 2013).

These results were taken as evidence for *syntax-based* theories of language production (e.g. Bock, 1990; Bock & Ferreira, 2014; Chang, Bock, & Goldberg, 2003; Chang, Dell, & Bock, 2006; Chang, Dell, Bock, & Griffin, 2000; Costa & Caramazza, 2002; Dell, 1986; Dell & O'Seaghdha, 1992; V. S. Ferreira & Dell, 2000; Garrett, 1975). This conclusion rests on the assumption that syntactic structure derives directly from conceptual representations but that lexical access is post-syntactic. Syntactic and lexical representations therefore have a degree of independence, with retrieved lexical items filling an, independently retrieved or constructed syntactic frame. These theories oppose *lexically-based* theories (e.g. Bock, 1982; Bock & Levelt, 1994; F. Ferreira, 2000; Levelt, 1989, 2001) that assume syntactic structure to be derived in response to morpho-syntactic information associated with specific lexical items. In these theories conceptual properties (e.g. animacy, saliency) rather than syntactic properties of the target language determine order of lexical activation. Syntactic representations can only be derived after retrieval of lexical items, and thus syntactic planning scope cannot extend beyond lexical planning scope.

The scope of syntactic planning may or may not coincide exactly with lexical planning scope. While it has generally been assumed that the phrasal scope effect reflects advance grammatical planning that is independent of lexical processing, there are alternative explanations that are consistent with lexical planning theories. Phrasal scope effects might be lexically driven rather than resulting from syntactic complexity (Wheeldon et al., 2013; Zhao & Yang, 2013, 2016). Allum and Wheeldon (2009) and Zhao, Alario, and Yang (2015) found that increased planning difficulty for conjoined noun phrases disappears if participants were provided with an image preview. This suggests that the phrasal scope effect may have a lexical rather than syntactic basis. Thus syntax may rely upon or emerge from lexical retrieval, and so share scope.

In line with this view sentence planning may be strictly lexical and incremental (Griffin, 2001, 2003, 2004; Zhao & Yang, 2013, 2016). Additional planning effort for conjoined noun phrases might result from non-linguistic, contextual, and communicational factors imposed on speech which expand planning scope beyond the minimum planning unit obligated by the language production system (F. Ferreira & Swets, 2002; Griffin, 2003; Wagner et al., 2010). Our discussion above suggests that the first planning unit might require less processing in writing compared to speech. Hence, the phrasal scope effect and generally planning beyond the first determiner-noun pair may be specific to spoken utterances and may not generalise to advance planning in writing.

The present research aimed to confirm phrasal scope for advance planning in written production of simple sentences and to examine whether advance planning beyond the single noun is associated with lexical encoding. We report three experiments in which participants generated short spoken and written sentences in response to image arrays. Summarising our argument: If the minimally-obligated sentence-initial planning unit is phrasal but lexically independent, it can extend beyond minimally obligated lexical planning scope. Therefore non sentence-initial nouns within this scope do not need to be lexically specified in advance of output onset. Conversely, the minimally-obligated planning unit in sentence production may be based upon lexical retrieval. All previous studies have been in the spoken modality. It is possible that this results in an extension of planning scope beyond that obligated by the language system. Phrasal scope might therefore result from speech-specific rather than modality-general constraints on the production system and may therefore not provide evidence against lexical accounts. Replication of effects in speech and writing will provide support for modality-independent constraints on planning scope.

## Experiment 1

Experiment 1 aimed to confirm phrase-level scope of advance planning of simple sentences. In an experimental paradigm similar to that adopted by Smith and Wheeldon (1999) participants performed image-description tasks in both writing and speech, producing sentences with subject noun phrases that were either *Simple* with a single noun phrase (NP) (i.e. *N1 moved up and the N2 and N3 moved down*) or *Complex* NPs (i.e. *N1 and the N2 moved up and N3 moved down*) in sentence-initial position. The codability of the image corresponding to the second noun (noun N2) was manipulated. This second noun was either within the sentence-initial subject noun phrase (the Complex NP condition) or outside this phrase (the Simple NP condition). Onset latencies and participants' eye movements were recorded.

Phrasal planning scope[2], consistent with previous findings, would be indicated by longer latency between stimulus onset and output (speech, typing) onset in the complex condition. Lexical accounts of syntax generation argue that lexical processing is a prerequisite for creating syntactic structure. Therefore, if planning has a phrasal scope, N2 codability will affect planning latency only if N2 is part of the subject noun phrase (i.e. in the Complex NP condition). Syntax-based accounts hold that lexical preparation for non sentence-initial nouns is not obligatory.

In addition to onset latencies we recorded participants' eye movements to each referent of the to-be produced structure. Eye movements, in contrast to onset latencies, provide information about whether or not participants looked at a referent and about

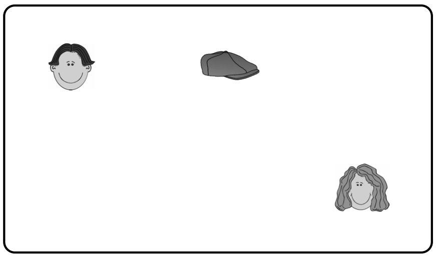
[2] In this experiment, and in Experiment 2, phrase and clause boundaries coincide while N2 was in the same clause from N1 for Complex NPs but in a different clauses for Simple NPs. It is, therefore, strictly not possible on the basis of findings from these experiments to disambiguate phrasal and clausal planning scope. Note, however, that this is disambiguated in Experiment 3.

the time and order of gaze shifted from one referent to another and thus, their consideration as part of the advance plan. As the time course of advance planning does not necessarily unfold over the linear order of the surface string (see e.g. Do, 2018; Momma, Slevc, & Phillips, 2015), eye movements provide important information on whether or not a phrase referent was preplanned.

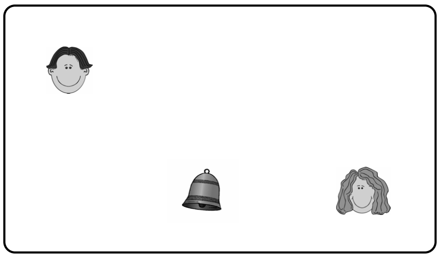
## Method

**Participants.** 32 psychology students (26 female, 6 male, mean age = 19.1 years, $SD = 1.4$, range: 18–25) participated as part of a research-reward scheme to collect research credits. All participants self-reported as native speakers of British English, as free of linguistic impairments, and having normal or corrected-to-normal vision. This research was approved by the College Research Ethics Committee (CREC) of Nottingham Trent University.

**Design.** Descriptions were elicited in response to arrays of three images. The images were presented horizontally aligned and then immediately separated with a rapid vertical movement (Figure 1). Images reached the target position after 100 ms and then stopped moving. Participants were asked to produce sentences of the form shown in Figure 1, with the order of the nouns in the sentence preserving the left-to-right order of images on the screen. In arrays as shown in Figure 1a, the leftmost image and the image in the centre of the screen moved up while the rightmost image moved down. In other arrays as shown in Figure 1b the leftmost image moved up and the other two images moved down. The target sentences differed with respect to the complexity of the first noun phrase while the overall complexity (i.e. number of noun phrases, VPs, and propositions) was held constant (Smith & Wheeldon, 1999). The subject phrase of the target sentence for Figure 1a is a conjoined noun phrase (*Peter and the hat*) and is, therefore, more complex than the subject phrase of the target sentence for Figure 1b which comprised just a single, proper name (*Peter*). All sentences were of identical length, included both a Complex and a Simple NP, and comprised three lexical items. Very rapid initial movement and exemplar sentences encouraged the use of past tense verbs, thus avoiding the need for the verb to agree with the number of the subject.



(a) Target sentence: *Peter and the hat moved up and Tania moved down*; Condition: Complex NP, low codable N2



(b) Target sentence: *Peter moved up and the bell and Tania moved down*; Condition: Simple NP, high codable N2

*Figure 1*. Example stimulus screens. The image in the centre is the referent for *N2*.

In a full factorial 2 × 2 × 2 design NP complexity (Simple vs. Complex) was crossed with N2 codability (high vs. low), and output modality (written vs. spoken). NP complexity represents whether the initial subject phrase of the target sentence was Complex or Simple. N2 codability was based on the number of names available for the image and was manipulated for the image corresponding to the linearly second noun in the elicited sentence (i.e. N2). For example, the image of a cap (Figure 1a), which is low-codable, has more associated names (e.g. *hat*, *cap*, *bonnet*) than high codable images such as the image of a bell (Figure 1b). In the written output modality participants typed their responses via a computer keyboard.

Both onset latency and participants' eye movements were recorded as indicators of advance planning. Onset latency was timed from appearance of the stimulus array on the computer screen to the start of spoken or written output. As an indicator of N2 planning we used gaze shift away from the image corresponding to the first noun and towards the referent of the second noun. Although all three images were areas of interest (henceforth, AOI) for the eye movement data, the critical variables were calculated for N1 and N2. In particular we calculated (1) the time relative to production onset when gaze shifted away from N1 and towards N2 and (2) the proportion of trials for which this shift happened before production onset (see results section of Experiment 1 for details). This gaze shift can be taken as evidence for a processing shift away from the first noun and towards the second noun (see e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer & Lethaus, 2004). Eye samples within the first 100 ms of each trial (the duration of image movement) were ignored.

**Materials.** To permit manipulation of N2 codability, estimates were obtained for images of everyday objects from the colourized version of the Snodgrass picture set images (Rossion & Pourtois, 2004; Snodgrass & Vanderwart, 1980). As part of a larger study (Torrance et al., 2017) 103 students from the same psychology department as those sampled in the main experiment (75 female, mean age = 22.1, $SD = 6.5$) provided written names for all 260 images. Codability was then calculated from the variability of different names used for an image weighted by the number of participants using each name ($H$; Lachman, 1973). Images were selected by first excluding images that elicited a high proportion of non-responses and images with very low ($< .3$) or very high ($> .95$) proportions of the most commonly used name. The remaining images were then divided into sets with $H$ scores ranging from 0 to 0.08 and from 1 to 2.48. 48 high codable images ($M = 0.02$, $SD = 0.04$) and 48 low codable images ($M = 1.34$, $SD = 0.35$) were then sampled from these sets. The images used for N2 can be found in Appendix A.

The resulting 96 images were combined with images of *Peter* and *Tania*, the boy and the girl in Figure 1. First names do not permit the participant strategically to start a sentence while delaying planning processes either by typing *the* or by extending its articulation (e.g. /theeee/). The plosive onsets of /peter/ and /tania/ permitted more precise onset timing in the spoken condition.

Item sets were counterbalanced for position of the images of *Peter* and *Tania* (left, right), for NP complexity and for modality such that each of the 96 images appeared just once per participant. The direction of the movement of the leftmost image (up, down) was counterbalanced across items within subjects. Participants performed blocks of trials in a single output modality with order (spoken-first or written-first) counterbalanced across subjects. 44 filler trials were added that elicited syntactically different descriptions from those elicited by the experimental items to prevent strategic sentence production and structural priming. Fillers included horizontal movement (*Tania and the cow swapped position*, *The plug moved to the left*), movement of less than three images (*Peter moved up*), all images moving into the same direction (*All pictures moved up*), and empty screens in which case participants generated the sentence, e.g., *No picture appeared*. The filler list was separated into two sets and counterbalanced by modality and order of session. Trial order was randomised. Each subject saw 96 stimulus trials and 44 filler trials (i.e. 48 stimulus and 22 filler trials per modality).

**Procedure.** Participants were tested individually. Experimental sessions started with nine-point eye tracker calibration and validation. Participants received instructions on the computer screen asking them to describe the action of the images from left to right. During the instruction phase, examples of image arrays and the associated target sentences were intermingled with examples of fillers. Participants were also taught with the names of the *Peter* and *Tania* images. The size of each image was 200 × 150 pixels (including transparent margins). Trials were then completed in separate writing and speaking blocks. Each block started with 10 practice trials during which the experimenter monitored descriptions and reminded the participant of the target sentence structure when necessary. After the training phase, the participant had the opportunity to ask questions and the eye tracker was recalibrated.

Each trial began with a blank screen (300 ms) followed by a screen-centred fixation point (a 21 × 21 pixel circle). Fixating this point for 200 ms triggered display of the image array, and also checked the spatial accuracy of the eye recordings. If the trigger did not respond, the experimenter performed a recalibration. The images appeared horizontally aligned just above the vertical centre of the screen, started moving immediately on display and arrived at their final positions after 100 ms. In the written session a text box (896 × 50 pixels) was shown on the bottom of the screen were the participant could monitor the production of his/her sentence. All images remained on the screen until the participant finished the end of the trial by pressing return. A blank screen followed. Participants were able to pause either before or after any trial. The duration of the entire experiment was approximately one hour.

**Apparatus.** Eye movements were recorded using a desk mounted SR Research EyeLink 1000 remote eye tracker to ensure free jaw and head movements. Eye data were sampled at 500 Hz sampling with recordings of just the right eye. The experiment was created in SR Research Experiment Builder, with custom code permitting keystroke display and capture in the written output condition. Keystrokes were recorded on a Steelseries Cherry (Black) MX gaming keyboard. Stimuli were displayed on a 19” ViewSonic Graphic Series (G90fB) CRT monitor with a screen resolution of

1,280 × 1,024 pixels and 85 Hz refresh rate using an Intel Core 2 PC. The spoken sentences were recorded with a Logitec headset using an ASIO audio driver supported by the Creative SB X-Fi sound card.

**Results**

We excluded trials in which participants produced structures that differed from the target sentence structure, used vague image names, e.g., *the thing*, or were output with considerable disfluency and/or extensive correction (17.1%). Trials with exceptionally long or short onset latencies were removed. For speech, trials with onset latencies shorter than 50 ms (0.9%) or longer than 4,000 ms (0.2%) were removed as were trials with sentence output durations shorter than 1,500 ms (0.13%) or longer than 10,000 ms (0.5%). In the written condition, trials with onset latencies longer than 5,000 ms (0.9%) and trials with total production durations longer than 40,000 ms (0.4%) were removed. For the analysis of eye data a further 11.6% were removed owing to a proportion of eye samples larger than .75 outside of AOIs.

Data were analysed by means of hierarchical Bayesian linear mixed effects models (Gelman et al., 2014; Kruschke, 2014; McElreath, 2016) using the probabilistic programming language Stan and the R interface Rstan (Carpenter et al., 2016; Hoffman & Gelman, 2014; Stan Development Team, 2015).[3] All models were fitted with maximal random effects structures (Barr, Levy, Scheepers, & Tily, 2013; Bates et al., 2015). The model was fitted with predictors for main effects of NP complexity, N2 codability, and modality, and for their interactions. Predictors were sum coded (±1). Inferential

[3] An adapted version of the code for linear mixed models presented in Bates, Kliegl, Vasishth, and Baayen (2015) was used for analyses. The Stan code for the binomial Bayesian linear mixed model is based on Sorensen, Hohenstein, and Vasishth (2016) and was kindly provided by Bruno Nicenboim.

statistics are summarised as the most probable *maximum a posteriori* parameter value $\hat{\mu}$ and the posterior 95% credible intervals (henceforth, CrI). 95% CrIs that do not contain zero are evidence for an effect of the predictor variable (see Kruschke, Aguinis, & Joo, 2012; Nicenboim & Vasishth, 2016; Sorensen et al., 2016). The strength of support for an effect was expressed in Bayes Factors calculated using the Savage-Dickey method (Dickey, Lientz, et al., 1970) (henceforth, BF signifying the evidence for the alternative hypothesis over the null hypothesis). A BF of 2, for example, means that the data are twice as likely under the alternative hypothesis than under the null hypothesis. While the logic of BFs preclude the existence of specific cut-off values that determine "statistical significance", we considered BFs around 10 or larger strong support for and small BFs (around 0.3 or smaller) evidence against the alternative hypothesis (see e.g. Baguley, 2012; M. D. Lee & Wagenmakers, 2014; Wagenmakers, Lodewyckx, Kuriyal, & Grasman, 2010).

All models were fitted with weak, locally uniform priors and by-subject and by-item adjustments using an LKJ prior on the correlation matrix of the variance-covariance matrix (see Sorensen et al., 2016).[4] Model convergence was confirmed by visual inspection of traceplots of the Markov chain Monte Carlo chains and the distribution of the posterior samples, and using the Rubin-Gelman statistic ($\hat{R} = 1$) (Gelman & Rubin, 1992). Code and data are available on figshare.com/articles/Analysis/5097403.

**Onset latency.** The onset latency data are summarized in Table 1. For a visualisation of the distribution of the data see Appendix B.

[4] For onset latency, models were run with four chains with 2,000 iterations per chain, 1,000 iterations warm-up and no thinning.

Table 1

*Descriptive summary of onset latency in ms (Experiment 1)*

| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 codability | *M* | *SE* | *N* | *M* | *SE* | *N* |
| Complex | high | 1245 | 22 | 342 | 1271 | 28 | 302 |
| | low | 1355 | 29 | 332 | 1403 | 38 | 287 |
| Simple | high | 1183 | 19 | 339 | 1245 | 30 | 302 |
| | low | 1228 | 25 | 335 | 1250 | 29 | 280 |

Latencies were positively skewed and were therefore square-root transformed prior to analysis as indicated by the Box-Cox power-transformation (Box & Cox, 1964). The results of the Bayesian linear mixed model are presented in Table 2. The model gave strong evidence for longer onset latencies for Complex NPs compared to Simple NPs (BF > 100). There was some evidence of longer onset latencies for low codable N2 images compared to high codable images (BF = 5). The interaction of NP complexity and N2 codability weakly supported (BF = 1.8). The interaction of N2 codability by NP complexity was inspected in nested comparisons within NP complexity contrasting low and high N2 codability. Strong evidence was found for longer onset latencies for low compared to high codable N2 images when sentences started with Complex NPs ($\hat{\mu}$ = 2.82, 95% CrI[1.26, 4.36], BF > 100) but not when sentences started with Simple first NPs (and therefore did not contain N2) ($\hat{\mu}$ = 0.5, 95% CrI[-1.1, 2.11], BF < 1). There was no evidence for any other model predictor (all BFs < 0.6).

The absence of by-modality interactions suggests that NP complexity and N2 codability have similar effects in writing and speech. To confirm the presence of NP

Table 2

*Main effects of first NP complexity, codability of N2, modality and their interactions inferred by the Bayesian linear mixed model on onset latency (Experiment 1)*

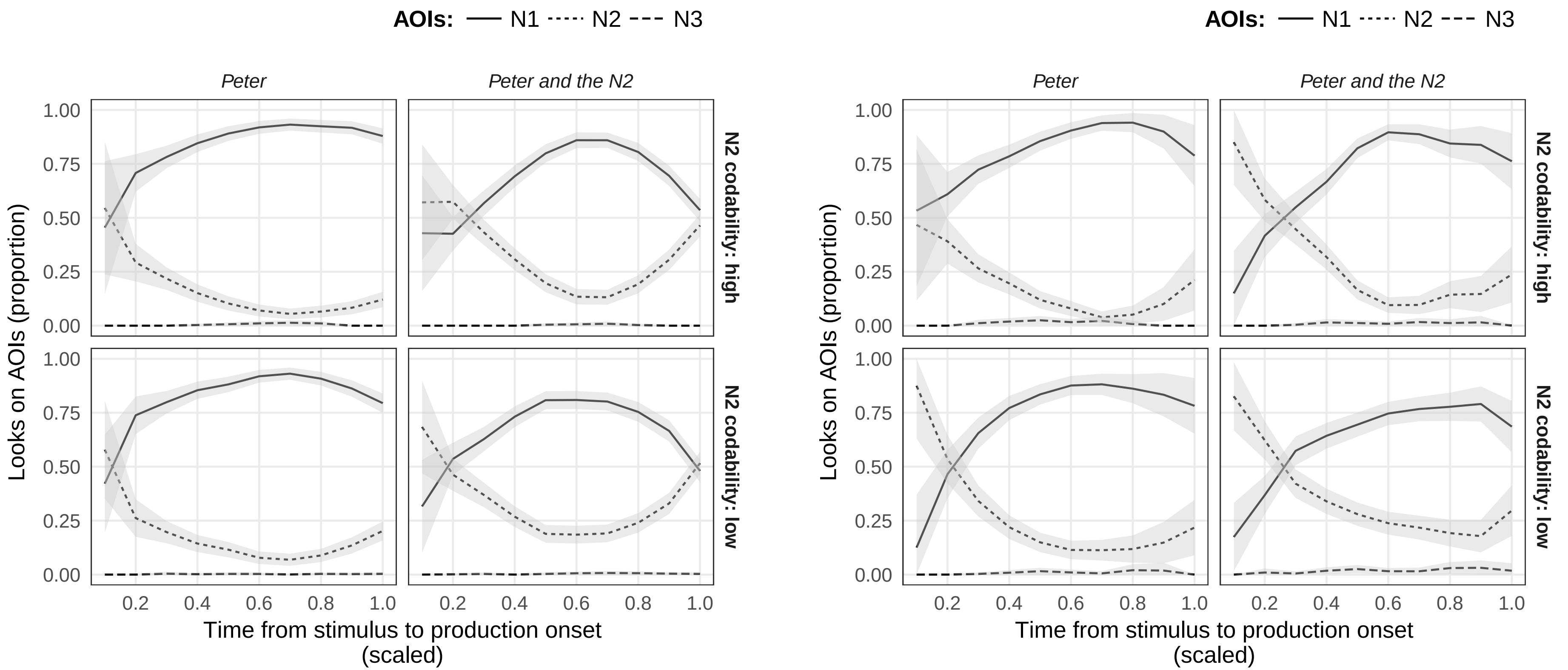
| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 0.59 | 0.31 | 0.86 |
| N2 codability | 0.41 | 0.11 | 0.72 |
| Modality | 0.25 | -0.60 | 1.09 |
| NP complexity × N2 codability | 0.29 | 0.04 | 0.55 |
| NP complexity × Modality | 0.00 | -0.25 | 0.26 |
| N2 codability × Modality | -0.04 | -0.29 | 0.21 |
| NP complexity × N2 codability × Modality | 0.07 | -0.18 | 0.32 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

complexity and N2 codability effects in writing, simple by-modality effects were tested. Similar NP complexity effects were found across modality with strong evidence supporting longer latencies in Complex NPs in writing ($\hat{\mu}$ = 2.36, 95% CrI[0.81, 3.86], BF = 67) and speech ($\hat{\mu}$ = 2.32, 95% CrI[0.83, 3.76], BF = 80). Further there was strong evidence for such a N2 codability effect in Complex NPs in both writing ($\hat{\mu}$ = 1.47, 95% CrI[0.41, 2.54], BF = 19) and speech ($\hat{\mu}$ = 1.35, 95% CrI[0.29, 2.34], BF = 13) but negligible evidence for N2 codability effects in Simple NPs, again, in both writing ($\hat{\mu}$ = 0.02, 95% CrI[-1.08, 1.12], BF < 1) and speech ($\hat{\mu}$ = .47, 95% CrI[-0.55, 1.47], BF < 1).

**Eye movements.** The proportions of eye samples in each AOI is illustrated across time to production onset in Figure 2. AOIs correspond to the image representing the first noun (N1), the second noun (N2) and the third noun (N3).

These graphs, which follow an approach to analysis of eye-movement data adopted in several previous studies (see e.g. Griffin, 2004; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer & Lethaus, 2004), show an initial increase in proportion of looks to N1 followed by a decrease, and the reverse pattern for N2. The point of inflection, where looks to the image representing N2 start to increase, indicates that processing is shifting from processing of the first noun to processing the second noun. Two dependent variables were calculated from these data: (1) the time relative to production onset for when the gaze shifted from the image of N1 to N2 and (2) the proportion of trials for which this gaze shift happened before production onset. Gaze shift from the N1 referent to N2 was defined as the first fixation (with a minimum duration of 100 ms) on the image of N2 after the gaze left N1.




(a) Speech

(b) Writing

*Figure 2*. Proportion of eye samples to AOIs from stimulus to production onset illustrated by condition. AOIs are the noun referents (i.e. N1, N2, N3). The time axis was scaled between 0 and 1 within trial and binned – 0 = stimulus onset and 1 = production onset. Bands indicate 95% confidence intervals (Experiment 1).

This gaze shift was detected in 87% of the data either before or after production onset. The time of gaze shift data were log transformed to account for positive skew. The data were analysed in a Bayesian linear mixed model. The results are summarized in Table 3. The model revealed substantial support for a main effect of NP complexity (BF $>$ 100) supporting earlier gaze shift in complex NPs. Further, there was strong support for a main effect of modality (BF $>$ 100) showing earlier gaze shift in speech compared to writing. The support for the modality by NP complexity interaction was moderate (BF = 5). This interaction was inspected in pairwise comparisons revealing strong support for NP complexity effects in speech ($\hat{\mu}$ = -0.16, 95% CrI[-0.22, -0.1], BF $>$ 100) and writing ($\hat{\mu}$ = -0.47, 95% CrI[-0.62, -0.32], BF $>$ 100) with a larger effect magnitude in the latter. Both effects indicate earlier gaze shift from N1 to N2 in Complex NPs. The evidence for the other model predictors was negligible (all BFs $<$ 0.2).

The proportion of trials in which gaze shift occurred before production onset is summarized by condition in Table 4. A Bayesian linear mixed model with Bernoulli distribution was fitted on whether or not gaze shift happened before production onset. The model outcome is summarized in Table 5. The model revealed strong support for a main effect of NP complexity (BF $>$ 100) showing larger proportions of gaze shift before production onset in Complex NPs. Strong evidence was found for a main effect of modality (BF $>$ 100) indicating larger proportions of pre-onset gaze shift to AOI N2 for speech. There was strong support for an interaction of these two main effects (BF $>$ 100). This interaction was inspected in pairwise comparisons revealing NP complexity effects in both speech ($\hat{\mu}$ = 4.66, 95% CrI[3.69, 5.69], BF $>$ 100) and writing ($\hat{\mu}$ = 2.11, 95% CrI[0.82, 3.46], BF $>$ 100) showing a larger probability of pre-onset gaze

Table 3

*Main effects of NP complexity, N2 codability, modality and their interactions inferred by Bayesian linear mixed model on the time of gaze shift data – time relative to production onset (Experiment 1)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | -0.08 | -0.10 | -0.06 |
| N2 codability | -0.02 | -0.05 | -0.00 |
| Modality | 0.18 | 0.15 | 0.21 |
| NP complexity × N2 codability | -0.02 | -0.04 | -0.00 |
| NP complexity × Modality | -0.04 | -0.06 | -0.02 |
| N2 codability × Modality | -0.01 | -0.03 | 0.01 |
| NP complexity × N2 codability × Modality | -0.02 | -0.04 | -0.00 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

shift in Complex NPs for speech compared to writing. We found only weak evidence for the main effect of N2 codability (BF = 1.5). The evidence for all other model predictors was negligible (all BF $<$ 0.2).

## Discussion

Findings from Experiment 1 indicate that sentences starting with Complex NPs were associated with longer onset latencies. Eye tracking data demonstrated that N2 received more attention prior to writing/speech onset when it was contained in the initial noun phrase (the subject of the first clause). Taken together these findings suggest initial planning extends to include the entire sentence-initial subject NP. Lower codability of the image associated with N2 resulted in longer onset latencies relative to

Table 4

*Descriptive summary for the proportions of trials in which the gaze shift from AOI N1 to N2 occurred before production onset (Experiment 1)*

| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 codability | *M* | *SE* | *N* | *M* | *SE* | *N* |
| complex | high | .53 | .03 | 322 | .17 | .03 | 127 |
| | low | .65 | .03 | 315 | .28 | .03 | 193 |
| simple | high | .15 | .02 | 316 | .11 | .02 | 173 |
| | low | .22 | .02 | 309 | .11 | .02 | 176 |

Table 5

*Main effects of NP complexity, N2 codability, modality and their interactions inferred by Bayesian linear mixed model on the proportion of trials with gaze shift from AOI N1 to N2 before production onset (Experiment 1)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 0.85 | 0.63 | 1.07 |
| N2 codability | 0.23 | 0.03 | 0.43 |
| Modality | -0.69 | -0.96 | -0.43 |
| NP complexity × N2 codability | 0.11 | -0.07 | 0.29 |
| NP complexity × Modality | -0.32 | -0.51 | -0.13 |
| N2 codability × Modality | -0.05 | -0.24 | 0.11 |
| NP complexity × N2 codability × Modality | 0.11 | -0.07 | 0.30 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

more easily coded images, but only when it was contained in the initial noun phrase. This suggests that advance planning of the initial noun phrase involved processing constituent nouns at a lexical level, and not just retrieval of associated concepts.

This finding does not strictly contradict theories that claim independence of lexical and syntactic planning (e.g. Chang et al., 2006; Garrett, 1975). The eye data suggest that N2 was typically attended only if it was part of a complex subject NP. Additionally N2 codability effects were observed in the onset latency for Complex NPs suggesting additional processing of the name of N2. Hence, the lexical entry of image N2 is prepared before production onset but only when it is contained in the sentence-initial subject noun phrase. These effects were present in both speech and writing.

There was proportionally more gaze dedicated to the image representing N2 in speech compared to writing before production began. Also looks to the image of N2 occurred earlier in speech than in writing. This may indicate that the second noun in Complex NPs is more likely to be planned in speech than in writing, possibly to satisfy speech-specific production demands on the output. However speech-specific requirements do not adequately explain the phrasal scope effect. No evidence was found that effects of NP complexity and N2 codability differed across modality. Therefore Experiment 1 concludes that phrasal planning occurs both in speech and in writing. Language producers, whether speaking or writing, plan lexical content and syntactic structure to the extent of the entire first noun phrase. The presence of NP complexity and N2 codability effects in writing as well as speech suggests that speech-specific demands on the output do not account for this more extended planning.

Codability effects found in Experiment 1 are, therefore, consistent with the theory that advance planning of syntax is lexically mediated. This findings is, however, open to alternative, methodological explanation. The present methods extended research to the written modality, but otherwise closely followed the design of previous studies in this area (e.g. Griffin, 2001; Martin et al., 2010; Smith & Wheeldon, 1999; Wagner et al., 2010; Wheeldon et al., 2013). Features of the methods used in this and previous studies – specifically regarding the gaze position at trial onset and the use of image-name agreement as a proxy for ease of lexical retrieval – potentially bias findings associated with N2 processing. Experiment 2 addresses these methodological issues.

## Experiment 2

Findings from Experiment 1 are consistent with both syntactic and lexical planning embracing the whole of the subject noun phrase. Experiment 2 eliminated methodological issues that may have influenced these findings.

First, it is possible that name agreement variation – the basis for the codability manipulation – was an inadequate proxy for ease of lexical retrieval. The lexical manipulation in Experiment 1 used images with low and high name agreement (codability). Some possible confounds were controlled. However visual (lexically independent) characteristics might have facilitated the processing of high codable images. Images with high name agreement may, for example, be more visually salient (e.g. brighter, more colours) than images with low agreement and therefore, receive more attention during early visual apprehension. To avoid this problem, Experiment 2 manipulated lexical availability via lexical primes that were activated when the participant fixated the image corresponding to the second noun.

Second, there is evidence that the starting point of linguistic processing can be controlled by, for example, subliminal visual cues that increase the salience of particular features of the display (see e.g. Gleitman, January, Nappa, & Trueswell, 2007;

Kuchinsky, Bock, & Irwin, 2011) or by the prominence of a particular referent (Prat-Sala & Branigan, 2000). In Experiment 1 N2 was the only novel image on the stimulus screen. This might have increased the prior attention to the image representing N2 and, hence, its incorporation into planning scope. To reduce this problem, Experiment 2 used a novel image in the rightmost position on the stimulus screens.

Third, the position of the fixation target that appeared at the start of each trial overlapped with the starting position of the image corresponding to the second noun. Therefore, advance sentence planning in previous studies might have been biased for any image located at the centre of the stimulus screen. Although the first 100 ms of every trial were removed from the analysis this fixation target may have cued early lexical processing of N2. Similar criticism can be levelled at previous studies in which the stimulus screen was preceded by a fixation cross in the centre of the screen (Allum & Wheeldon, 2007, 2009; Martin et al., 2010; Wagner et al., 2010), the top middle (Griffin, 2001), at the location where the first image is going to appear (Wheeldon et al., 2013), or by a frame (Smith & Wheeldon, 1999, 2001, 2004). To avoid this problem, Experiment 2 randomised the location of the trial-initial fixation targets. Gaze triggers allow to control the participants' gaze position before stimulus onset.

## Method

### Participants.

32 psychology students (28 female, 4 male, mean age = 18.9 years, *SD* = 0.8, range: 18–21) participated participated as part of a voluntary research-participation program. All participants were self-reported as native speakers of British English, as free of linguistic impairments, and having normal or corrected-to-normal vision.

(2) a. Peter and the hat moved up and the sock moved down.

b. Peter moved up and the hat and the sock move down.

**Design.** Experiment 2 followed the same general design as Experiment 1. In a full factorial 2 × 2 × 2 design we manipulated NP complexity, ease of N2 retrieval and output modality. Participants were instructed to use descriptions as shown in example (2). NP complexity was manipulated in the same way as in Experiment 1:

The ease of lexical processing of N2 was manipulated by using a gaze-contingent written name prime (Bock, 1986; Dell & O'Seaghdha, 1992; Levelt et al., 1991). Fixations on the image corresponding to the second noun triggered display of a printed prime word superimposed on the image followed by a mask (#######). The prime word was the most commonly given name for the image, derived from the naming data described in Experiment 1. In the control condition these were replaced with a length-matched non-word (e.g. *qji* vs. *hat*).

**Materials.** For the image corresponding to the second noun (N2), only images with medium to low codability were employed ($M$ = 1.1, $SD$ = 0.51, range: 0.4–2.5). Images were not included if they frequently elicited non-responses or for which the proportion of subjects giving the most commonly given name was smaller than .3 or larger than .95, or with a most commonly used name longer than 10 letters. A total of 96 items were sampled from the remaining images. The CELEX data base was used to generate non-words with unconstrained combinations of letters and sampled strings that matched the length of the image names (Medler & Binder, 2005). Stimulus items can be found in Appendix C.

Prime/picture pairs were piloted in a typed image naming task performed by ten native speakers of British English. Images were presented with image-name primes, with non-words, or without any additional information, overlaying the image for either 50 ms or 80 ms. The results showed that, compared to the no prime condition ($M$ = 1555, $SD$ = 1103), onset latencies were shorter for image name primes showing a mean posterior difference of -178 ms for 50 ms priming duration ($M$ = 1378, $SD$ = 907) and -326 ms for 80 ms priming duration ($M$ = 1229, $SD$ = 709). Non-words led to longer onset latencies showing a mean posterior difference of 68 ms for 50 ms primes ($M$ = 1624, $SD$ = 914) and 109 ms for 80 ms primes ($M$ = 1687, $SD$ = 846) compared to the unprimed condition. The probability of using the prime word increased, compared to the no prime condition ($M$ = .68, $SD$ = 0.47), for image-name primes by .16 mean posterior difference for 50 ms ($M$ = .84, $SD$ = .37) and by .28 mean posterior difference for 80 ms prime duration ($M$ = .89, $SD$ = .31). These differences were negligible for non-word primes showing a decrease of -.06 mean posterior difference for 50 ms prime duration ($M$ = .72, $SD$ = .45) and a mean posterior difference of -.1 for 80 ms ($M$ = .69, $SD$ = .46). In sum the priming manipulation facilitated lexical retrieval. Details on this pilot study and the analysis can be found in Appendix D.

For the main experiment 96 images were shown in each prime by NP complexity by modality condition, counterbalanced in a Latin square design. Item sets were counterbalanced for whether *Peter* or *Tania* appeared in the left most position. The rightmost image was sampled from coloured Snodgrass images, excluding complex images and the 96 images used for N2. Session order was counterbalanced between subjects and direction of movement of the left most image was counterbalanced between items. 44 fillers were created targeting structurally different sentences as described in Experiment 1 New images were sampled for filler trials and horizontal image movement was omitted. Fillers were allocated to item lists as described in Experiment 1. Trial order was randomised. Each subject saw 96 experimental and 44 filler trials (i.e. 48 experimental trials and 22 filler trials per modality).

**Procedure.** The procedure followed that of Experiment 1 with the following differences. The location of the fixation target – the target that the participant had to fixate in order to initiate the trial – was randomized within the screen area, excluding the margins and an area of 160 by 170 pixels around the centre of the screen. Fixations on N2 triggered primes. Both prime and mask were displayed superimposed on the image in green 24 pt Arial font (RGB = [0, 255, 0]) to avoid interference with the image's colour. Primes were triggered immediately when gaze entered the image area. The prime was then displayed for 80 ms followed by a 20 ms mask. Primes were re-triggered if gaze left and then returned to the image, but only if the delay since the offset of the last fixation on the image was greater than 500 ms. This avoided successive primes for eye blinks which would make the prime readable.

**Apparatus.** The keyboard was replaced by a Microsoft Sidewinder X4 gaming keyboard (because participants reported that the size of the backspace of the Steelseries keyboard caused errors while editing). This was modified by removing various extraneous function keys. Otherwise apparatus was identical to that used in Experiment 1.

## Results

Prior to analysis trials where the produced sentence did not match the target structure, included vague image names, or contained a considerable amount of disfluency or editing were removed (13.4%). For speech, trials with onset latencies shorter than 50 ms (0.6%) or longer than 4,000 ms (0.2%) were removed as well as sentence durations shorter than 1,500 ms (0.07%) or longer than 10,000 ms (0.9%). In the written trials, responses with onset latencies longer than 5,000 ms (1.5%) were removed as well as trials with durations longer than 40,000 ms (0.6%). Statistical analysis methods were the same as detailed for Experiment 1. For the analysis of eye data a further 10.1% trials were removed owing to a proportion of eye samples larger than .75 outside of AOIs. Statistical analysis followed the same methods as those described for Experiment 1.

**Onset latency.** The onset latency data are shown in Table 6. A visualization of this can be found in Appendix E. To correct for positive skew the onset latency was logarithmically transformed for the analysis as suggested by the Box-Cox test (Box & Cox, 1964). The results of the Bayesian linear mixed model are presented in Table 7. The model revealed strong evidence (BF $>$ 100) showing longer onset latencies for Complex NPs compared to Simple NPs. This NP complexity effect was tested as simple main effects within modality revealing strong support in both writing ($\hat{\mu} = 0.19$, 95% CrI[0.12, 0.26], BF $>$ 100) and speech ($\hat{\mu} = 0.22$, 95% CrI[0.15, 0.29], BF $>$ 100). Moreover, weak evidence was found for longer onset latencies in writing compared to speech (BF = 2). The data did not support an effect of name priming of N2, and did not support any interaction effects (all BFs $<$ 0.1).

**Eye movements.** The time course of proportions of eye samples to each AOI before production onset is illustrated in Figure 3. These graphs illustrate the shift of attention away from N1 and towards N2 indicating processing shift from the first noun to the second noun.

Table 6

*Descriptive summary of onset latency in ms (Experiment 2)*

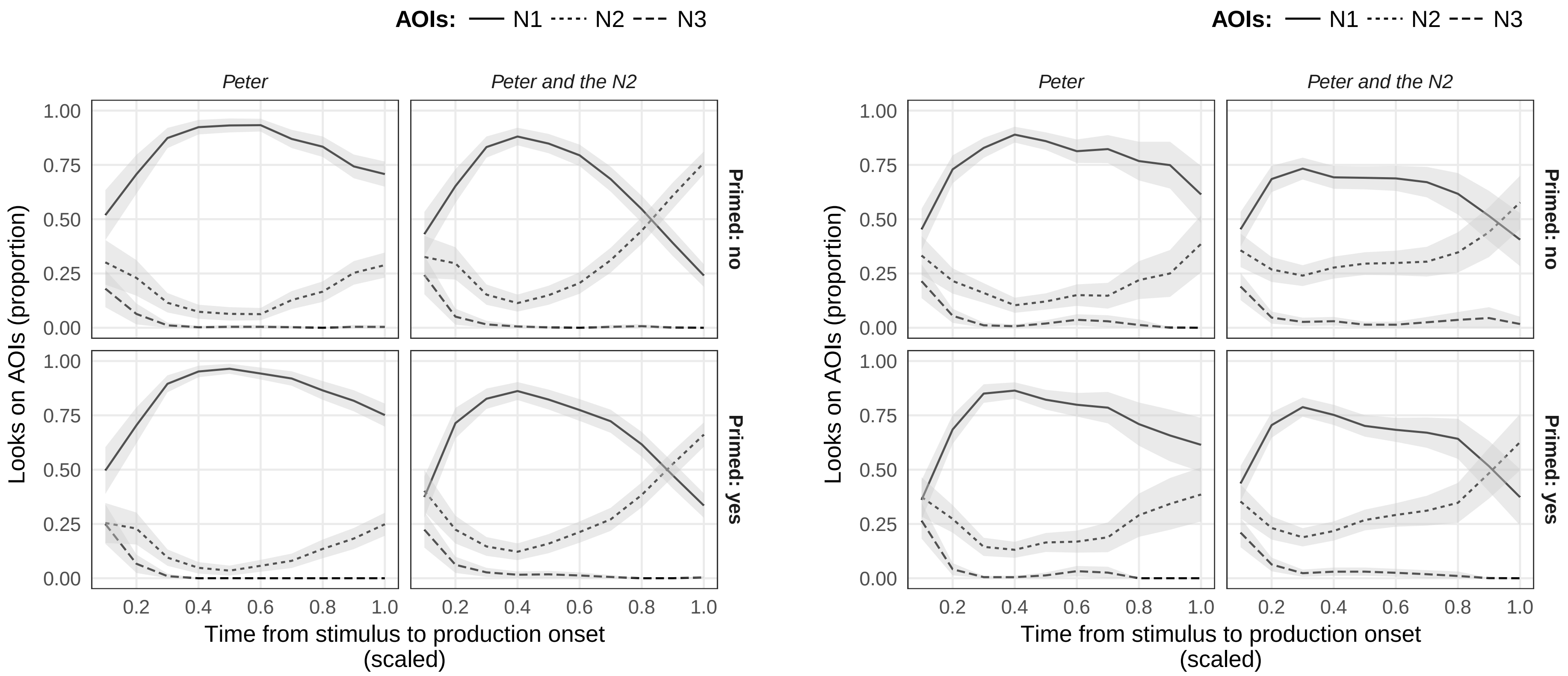
| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 prime | *M* | *SE* | *N* | *M* | *SE* | *N* |
| Complex | image name | 1303 | 23 | 339 | 1462 | 33 | 336 |
| | non-word | 1286 | 24 | 324 | 1469 | 32 | 321 |
| Simple | image name | 1155 | 19 | 330 | 1312 | 27 | 328 |
| | non-word | 1142 | 24 | 324 | 1326 | 28 | 326 |

Table 7

*Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on onset latency (Experiment 2)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 0.05 | 0.04 | 0.06 |
| N2 prime | 0.00 | -0.01 | 0.02 |
| Modality | 0.07 | 0.02 | 0.11 |
| NP complexity × N2 prime | -0.00 | -0.01 | 0.01 |
| NP complexity × Modality | -0.00 | -0.02 | 0.01 |
| N2 prime × Modality | -0.01 | -0.02 | 0.00 |
| NP complexity × N2 prime × Modality | 0.00 | -0.01 | 0.02 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI




(a) Speech

(b) Writing

*Figure 3*. Proportion of eye samples to AOIs by condition from stimulus to production onset. AOIs are the noun referents (i.e. N1, N2, N3). The time axis was scaled between 0 and 1 within trial and binned – 0 = stimulus onset and 1 = production onset. Bands indicate 95% confidence intervals (Experiment 2).

As in Experiment 1 pre-onset planning of N2 was assessed by calculating the time relative to production onset for when the gaze shift from the image representing N1 to N2 occurred and whether or not this gaze shift happened before production onset. Gaze shift from N1 to N2 was defined as the first fixation (minimum duration of 100 ms) on N2 after the gaze moved away from AOI N1.

In 97% of the data gaze shift was detected either before or after production onset. The data were log transformed to account for positive skew and analysed in a Bayesian linear mixed model. The results are summarized in Table 8. The model revealed substantial support for a main effect of NP complexity (BF $>$ 100) supporting earlier gaze shift in complex NPs. This effect was confirmed for both modalities showing strong evidence for writing ($\hat{\mu}$ = -0.4, 95% CrI[-0.54, -0.26], BF $>$ 100) and speech ($\hat{\mu}$ = -0.17, 95% CrI[-0.26, -0.08], BF = 22). Further, there was strong support for a main effect of modality (BF $>$ 100) showing earlier gaze shift in speech compared to writing. The support for all other predictors was negligible (all BFs $<$ 0.1).

The proportion of trials in which gaze shift occurred before production onset is summarized by condition in Table 9. A logistic Bayesian mixed model was fitted on whether or not gaze shift happened before production onset. The model outcome is summarized in Table 10. The model revealed strong support for a main effect of NP complexity (BF $>$ 100) showing larger proportions of gaze shift before production onset in Complex NPs. Strong evidence was found for a main effect of modality (BF $>$ 100) indicating larger proportions in speech compared to writing. There was strong support for the interaction of NP complexity and modality (BF $>$ 100). This interaction was inspected in pairwise comparisons revealing NP complexity effects in both speech ($\hat{\mu}$ = 6.31, 95% CrI[4.92, 7.87], BF $>$ 100) and writing ($\hat{\mu}$ = 2.52, 95% CrI[1.44, 3.65],

Table 8

*Predictors main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on the time of gaze shift data (Experiment 2)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | -0.07 | -0.09 | -0.05 |
| N2 prime | 0.01 | -0.01 | 0.02 |
| Modality | 0.08 | 0.06 | 0.10 |
| NP complexity × N2 prime | -0.00 | -0.02 | 0.01 |
| NP complexity × Modality | -0.03 | -0.05 | -0.01 |
| N2 prime × Modality | 0.00 | -0.01 | 0.02 |
| NP complexity × N2 prime × Modality | -0.00 | -0.01 | 0.01 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

BF > 100) with a larger magnitude in speech. The evidence for all other model predictors was negligible (all BFs < 0.2).

**Discussion**

The aim of Experiment 2 was to establish whether results from Experiment 1 could be replicated after removing features of methods used in Experiment 1 (and previous similar studies) that might encourage greater attention to the referent of N2 prior to production onset. Experiment 2 replicated the NP complexity effect of Experiment 1 in both writing and speech after controlling for factors that might have encouraged planning beyond the first noun. This was found for both initial latencies and in eye measures.

The lexical planning effect found on N2 in Experiment 1 – easily codable N2s gave

Table 9

*Descriptive summary for the proportion of trial in which the gaze shift from the referent of the first noun to the referent of the second noun – N1 to N2 – occurred before production began (Experiment 2)*

| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 prime | *M* | *SE* | *N* | *M* | *SE* | *N* |
| complex | image name | .83 | .02 | 242 | .48 | .03 | 285 |
| | non-word | .86 | .02 | 227 | .50 | .03 | 257 |
| simple | image name | .33 | .03 | 239 | .29 | .03 | 246 |
| | non-word | .40 | .03 | 230 | .26 | .03 | 247 |

Table 10

*Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on the proportion of trials with gaze shift from first noun referent N1 to second noun referent N2 before production onset (Experiment 2)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 1.10 | 0.87 | 1.35 |
| N2 prime | -0.09 | -0.25 | 0.07 |
| Modality | -0.86 | -1.13 | -0.60 |
| NP complexity × N2 prime | -0.01 | -0.17 | 0.15 |
| NP complexity × Modality | -0.47 | -0.70 | -0.26 |
| N2 prime × Modality | 0.09 | -0.08 | 0.26 |
| NP complexity × N2 prime × Modality | -0.03 | -0.20 | 0.13 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

shorter onset latencies in the Complex NP condition and earlier gaze shifts – was not replicated in Experiment 2. There was also no effect of N2 priming observed on any eye measure. One possible explanation for this may simply be a failure to prime lexical retrieval. We think this explanation is unlikely. Pilot data from image naming gave evidence of the effectiveness of the priming manipulation in speeding lexical access although evidence from single object naming may not generalise to naming in the context of sentence production. However, findings from Experiment 3 provide more direct support for the efficacy of our priming manipulation. We will return to this issue in the general discussion.

The absence of effects of N2 priming therefore suggest that lexical retrieval beyond the first noun of the initial noun phrase is not obligated by the language production system. This may suggest that any lexical processing of N2 prior to production onset did not, in fact, play a role in preparing the subsequent utterance. This finding is in line with syntax-based models of sentence production (e.g. Chang et al., 2006).

Experiment 2 therefore replicated the finding of phrasal (or possibly clausal) scope for sentence-initial planning from Experiment 1, but did not provide support for obligatory lexical retrieval beyond the first noun. These findings reproduced in both writing and speech. This suggests that the syntax of the first phrase is always planned. Lexical retrieval, on the other hand, is only required for the first noun but might go beyond the first noun depending on extra-linguistic factors, e.g. the production context or the experimental set-up (Wheeldon et al., 2013).

## Experiment 3

Findings from Experiments 1 and 2 provide strong evidence that, independent of output modality, syntactic planning prior to production onset necessarily extends beyond the sentence-initial noun when the sentence starts with a coordinated noun phrase. After controlling for the methodological issue found in Experiment 1 (and various previous studies), Experiment 2 concluded that there is, however, no obligation for advance *lexical* planning beyond the sentence initial noun.

In the discussions above, extended initial planning for sentences with Complex subject NPs was interpreted as evidence for phrasal planning scope. However in Experiments 1 and 2 the elicited sentence structure comprised two intransitive clauses in which the complexity of the sentence-initial clause differed with regard to the first noun phrase. Evidence from Experiments 1 and 2 does no rule out the possibility that obligatory advance sentence planning scopes has clausal rather that phrasal scope This was suggested by some early studies (Bock & Cutting, 1992; Bock & Miller, 1991). Although several subsequent studies of spoken production found have found evidence against clausal scope (see Smith & Wheeldon, 1999; Wheeldon et al., 2013) this has yet to be tested in writing. Experiment 3 ruled out this possibility by eliciting transitive single-clause sentences (e.g. *N1 and the N2 moved above the N3*). In all other respects design was the same as in Experiment 2.

## Method

**Participants.** 32 psychology students (30 female, 2 male, mean age = 19.3 years, *SD* = 2, range: 18–29) participated as part of a research-reward scheme to collect research credits. All participants self-reported as native speakers of British English, as free of linguistic impairments, and as having normal or corrected-to-normal vision.

**Design & Material.** This experiment used the same design and materials as Experiment 2, with the exception that we elicited single-clause sentences. NP complexity was manipulated and crossed with the ease of N2 retrieval and output

(3) a. Peter and the hat moved above the sock.

b. Peter moved above the hat and the sock.

modality. Participants were instructed to describe the stimulus array with sentences of the form illustrated in example (3). While the descriptions in Experiments 1 and 2 contained two intransitive propositions (i.e. clauses) with all noun phrases in subject position, the descriptions in example (3) consist of one transitive proposition.

**Procedure & Apparatus.** The procedure and the apparatus were the same as in Experiment 2.

**Results**

Prior to analysis trials where the elicited sentence did not match the structure of the target sentence, and where image names were imprecise, or were produced with considerable disfluency or editing were removed (10.2%). For speech, trials with onset latencies shorter than 50 ms (3.1%) or longer than 4,000 ms (0.5%) were removed as were trials with sentence durations longer than 10,000 ms (0.3%). In the written condition, responses with onset latencies longer than 5,000 ms (0.6%) and trials with durations longer than 40,000 ms (0.1%) were removed. For analysis of eye data a further 12.5% were removed because proportion of total samples outside of defined AOIs was greater than .75.

Statistical analysis followed the same methods as those described for Experiment 1.

**Onset latency.** Observed onset latencies are summarized in Table 11. For a visualisation of the entire distribution by-condition see Appendix F.

For statistical analysis the onset latency was square-root transformed to correct

Table 11

*Descriptive summary of onset latency in ms (Experiment 3)*

| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 prime | *M* | *SE* | *N* | *M* | *SE* | *N* |
| Complex | image name | 1165 | 23 | 352 | 1335 | 24 | 308 |
| | non-word | 1216 | 27 | 345 | 1309 | 27 | 330 |
| Simple | image name | 1089 | 20 | 346 | 1240 | 26 | 316 |
| | non-word | 1073 | 18 | 354 | 1254 | 23 | 325 |

for positive skew as determined by the Box-Cox power-transformation (Box & Cox, 1964). The results of the Bayesian linear mixed model are shown in Table 12. The model gave compelling evidence for longer onset latencies in Complex NPs compared to Simple NPs (BF $>$ 100), and longer onset latencies in the written compared to the spoken output condition (BF $>$ 100). The NP complexity effect was tested within modality calculated as simple main effects from the posterior samples of the model. Strong evidence for NP complexity effects was found in both writing ($\hat{\mu}$ = 2.35, 95% CrI[1.06, 3.69], BF $>$ 100) and speech ($\hat{\mu}$ = 2.94, 95% CrI[1.58, 4.27], BF $>$ 100). There was negligible support for a main effect of N2 prime (BF $<$ 1). The posterior samples support a three-way interaction of NP complexity, N2 prime-type and output modality which was, however, not substantial (BF = 0.77). A follow-up inspection of N2 prime-type comparisons within NP complexity and output modality revealed non-substantially shorter latencies in image name primes for Complex NPs when responses were spoken ($\hat{\mu}$ = -0.73, 95% CrI[-1.63, 0.13], BF = 1.5) but negligible evidence of priming was found in the remaining model contrasts (BF $<$ 1); either in

Simple NPs in speech ($\hat{\mu}$ = 0.32, 95% CrI[-0.52, 1.17]), or in Complex ($\hat{\mu}$ = -0.17, 95% CrI[-1.07, 0.72]) or in Simple NPs in writing ($\hat{\mu}$ = 0.47, 95% CrI[-0.4, 1.36]). The evidence for all other model predictors was negligible (all BFs < 0.2).

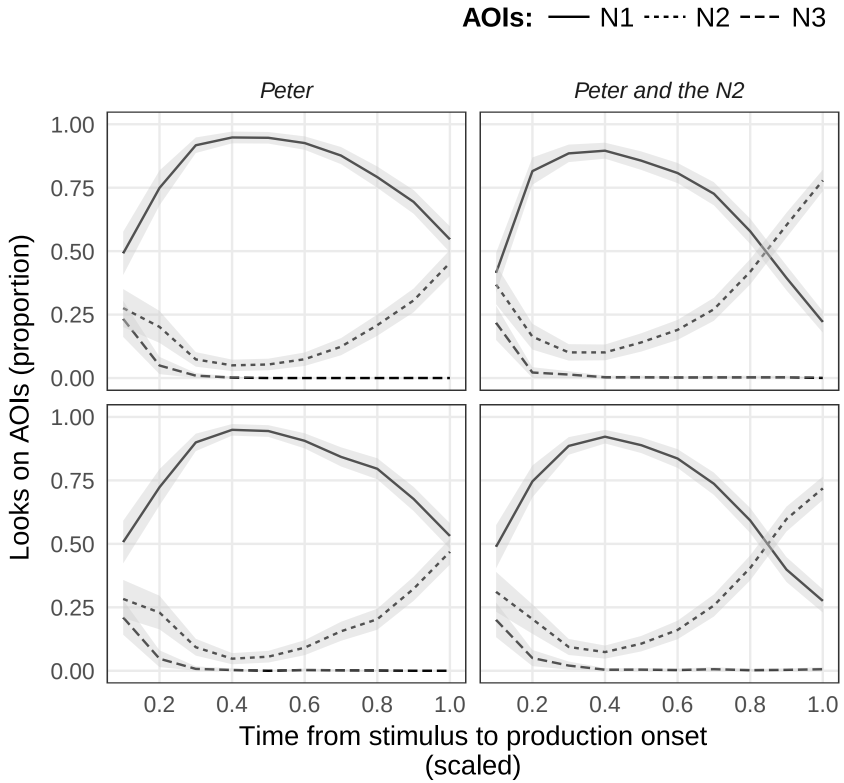
Table 12

*Main effects of first NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on onset latency (Experiment 3)*

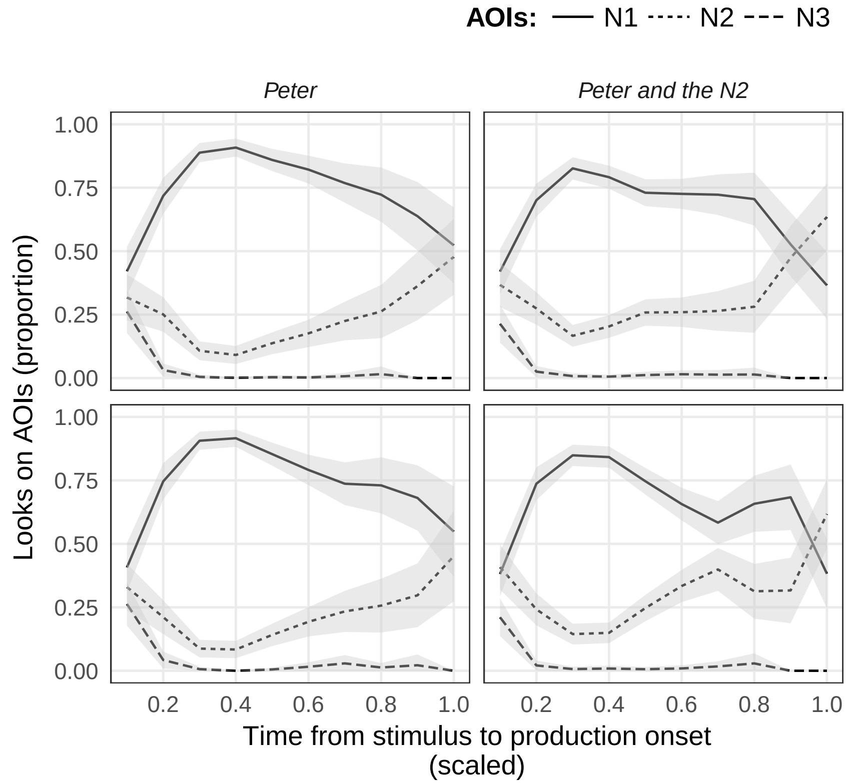
| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 0.66 | 0.43 | 0.91 |
| N2 prime | -0.01 | -0.22 | 0.20 |
| Modality | 1.25 | 0.74 | 1.73 |
| NP complexity × N2 prime | -0.05 | -0.26 | 0.15 |
| NP complexity × Modality | -0.07 | -0.30 | 0.15 |
| N2 prime × Modality | 0.09 | -0.16 | 0.34 |
| NP complexity × N2 prime × Modality | 0.21 | -0.01 | 0.43 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

**Eye movements.** The time course of proportions of eye samples to each AOI for the time period before production onset is illustrated in Figure 4. These graphs illustrate the change of attention dedicated to N1 and N2 while there were only few eye samples on N3.




(a) Speech




(b) Writing

*Figure 4.* Proportion of eye samples to AOIs by condition from stimulus to production onset. AOIs are the referents (i.e. N1, N2, N3) as mentioned in the target sentence. The time axis was scaled between 0 and 1 within trial and binned – 0 = stimulus onset and 1 = production onset. Bands indicate 95% confidence intervals (Experiment 3).

As in the previous experiments, the time relative to production onset for when the gaze shift from the image representing N1 to N2 occurred was calculated as well as whether or not this gaze shift happened before production onset. Gaze shift from AOI N1 to N2 was defined as the first fixation (minimum duration of 100 ms) on N2 after N1 was fixated. In 97% of the data this gaze shift from image N1 to N2 was detected either before or after production onset. These timing data were log transformed to account for positive skew. The data were analysed in a Bayesian linear mixed model. The results are summarized in Table 13.

Table 13

*Model summary showing main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on the time of gaze sift data – time relative to production onset (Experiment 3)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | -0.07 | -0.08 | -0.05 |
| N2 prime | -0.00 | -0.01 | 0.01 |
| Modality | 0.09 | 0.07 | 0.12 |
| NP complexity × N2 prime | 0.00 | -0.01 | 0.01 |
| NP complexity × Modality | -0.04 | -0.05 | -0.02 |
| N2 prime × Modality | -0.01 | -0.02 | 0.00 |
| NP complexity × N2 prime × Modality | -0.00 | -0.01 | 0.01 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

The model revealed substantial support for a main effect of NP complexity (BF > 100) supporting earlier gaze shift in Complex NPs. Further, there was strong support for a main effect of modality (BF > 100) showing earlier gaze shift in speech compared to writing. The model showed strong support for an interaction of these two main effects (BF = 23). The interaction was inspected in pairwise comparisons revealing strong support for earlier gaze shift for Complex NPs in speech ($\hat{\mu}$ = -0.11, 95% CrI[-0.16, -0.06], BF = 17) and writing ($\hat{\mu}$ = -0.41, 95% CrI[-0.54, -0.29], BF > 100) with a larger magnitude in the latter. The support for all other predictors was negligible (all BFs < 0.02).

The proportion of trials in which gaze shift occurred before rather than after production onset is summarized in Table 14.

Table 14

*Descriptive summary for the proportions of trials in which the gaze shift from AOI N1 for the first noun to the second noun N2 occurred before production onset (Experiment 3)*

| | | Speech | | | Writing | | |
|---|---|---|---|---|---|---|---|
| NP complexity | N2 prime | *M* | *SE* | *N* | *M* | *SE* | *N* |
| complex | image name | .78 | .02 | 343 | .48 | .03 | 232 |
| | non-word | .82 | .02 | 335 | .47 | .03 | 239 |
| simple | image name | .56 | .03 | 339 | .26 | .03 | 210 |
| | non-word | .54 | .03 | 342 | .25 | .03 | 219 |

A logistic Bayesian mixed model was fitted on these data. The model outcome is summarized in Table 15. The analysis revealed strong support for the main effect of NP complexity (BF > 100) showing larger proportions of gaze shift before production onset in Complex NPs. This effect was assessed in simple main effects within modality. These comparisons supported NP complexity effects for both writing ($\hat{\mu}$ = 2.66, 95% CrI[1.61, 3.76], BF > 100) and speech ($\hat{\mu}$ = 3.09, 95% CrI[2.08, 4.14], BF > 100). Further, strong evidence (BF > 100) was found for a main effect of modality indicating larger proportions in speech compared to writing. The evidence for all other model predictors was negligible (all BFs < 0.2).

Table 15

*Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on the proportion of trials with gaze shift from the image representing N1 to N2 occurring before production onset (Experiment 3)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| NP complexity | 0.72 | 0.52 | 0.93 |
| N2 prime | 0.04 | -0.09 | 0.18 |
| Modality | -1.03 | -1.43 | -0.68 |
| NP complexity × N2 prime | -0.04 | -0.16 | 0.08 |
| NP complexity × Modality | -0.05 | -0.22 | 0.12 |
| N2 prime × Modality | 0.08 | -0.05 | 0.21 |
| NP complexity × N2 prime × Modality | 0.05 | -0.08 | 0.18 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

## Discussion

The aim of Experiment 3 was to eliminate the possibility that tendency to plan syntax beyond the sentence-initial in Experiments 1 and 2 was associated with a language production system requirement to advance-plan the whole initial clause rather than just the subject noun phrase. If advance planning scopes over the clause rather than the phrase, one would predict no difference between Complex and Simple subject

NPs for single clause sentences. This prediction was not borne out. Instead the data provide a replication of the NP complexity effects in single-clause sentences, supporting the phrase as the unit of advance planning (e.g. Smith & Wheeldon, 1999). This replicated the phrasal scope effect observed for two-clause utterances in Experiments 1 and 2. The present findings confirm that this is true for spoken output, consistent with the conclusions of several previous studies (e.g. Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Wagner et al., 2010; Wheeldon et al., 2013) and, for the first time, demonstrate that it is also true for writing. Phrasal scope therefore appears to hold for both speech and writing and is therefore plausibly a basic feature of the language production system.

In contrast with findings from Experiment 2 we found, in the spoken condition only, shorter onset latencies for lexically primed referents in Complex NPs. This suggests that lexical advance planning beyond N1 is dependent on a combination of both output modality and the syntax of the elicited sentence. As argued before, the extent of advance planning in language production is likely, in part, to be dependent on speech-specific output requirements. Previous research has also suggested that lexical advance planning is dependent (in part) on syntactic factors (e.g. Konopka, 2012; Wagner et al., 2010; Wheeldon et al., 2013). The transitive, single clause structure elicited in Experiment 3 is more likely to require some planning across the whole sentence; this contrasts with the two-clause structures elicited in Experiments 1 and 2 in which pausing before the second clause is permitted: Hesitation is more common, and therefore more permissible, at clause boundaries than within clauses in spontaneous speech (e.g. Boomer, 1965; Goldman-Eisler, 1972; Hawkins, 1971). It is worth noting, however, that if this account is correct it requires that some advance planning must scope beyond the initial phrase (Meyer, 1996; Smith & Wheeldon, 1999). This anticipation in necessary in order to make an advance judgment about the production requirements of the to-be-produced sentence (Griffin, 2003). If lexical advance-planning beyond N1 is contingent on structure beyond the initial phrase then the production system must have some knowledge of this prior to output onset.

## General Discussion

The research presented here had two objectives: First, these experiments sought to confirm that planning of the initial subject noun phrase is obligatory in sentence production for reasons that are independent of output modality. Specifically this study aimed to exclude the possibility that previously published findings were specific to demands imposed by the spoken output modality. Second, this research tested whether phrase-level planning is lexically mediated, or whether planning of lexical items can potentially be delayed until after production onset by reproducing this effect for speech (Wheeldon et al., 2013) and, for the first time, in writing. Again, previous research suggests that advance planning beyond the first noun might be the result of speech-specific production demands. Thus, crucially, establishing that results are common to both output modalities is provides clear support for the claim that the planning scope that the present data imply derives from fundamental properties of a common (modality independent) language production system.

In all three experiments grammatical encoding was found to embrace the entire first coordinated subject noun phrase in both speech and writing. However, non sentence-initial nouns, even when part of the sentence-initial phrase, typically remained lexically unspecified. Advance planning in writing and in speech followed similar planning patterns, with the exception that there was evidence that non-sentence initial nouns were more likely to be retrieved in advance of output onset when the output was spoken. This points toward a fundamental requirement for advance planning of just the syntax of the initial subject noun phrase. Lexical specification, on the other hand, is not required, but may occur to meet speech-specific production requirements such as output-fluency (Griffin, 2003; Levelt & Meyer, 2000). Before accepting this conclusion possible alternative explanations for the presented findings are going to be discussed.

The reported experiments differ from previous studies (Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Wagner et al., 2010; Wheeldon et al., 2013) in that the sentence initial noun did not require a determiner and was repeated throughout the experiment making, it very easy to retrieve. It is possible that this will have encouraged more advance planning than is minimally required by the language production system. For example Konopka (2012) found advance lexical retrieval in coordinated noun phrases starting with a high frequency noun followed by a low frequency noun rather than the other way around. However, there are two reasons why this is unlikely to account for these data. First, if planning beyond the first noun was encouraged by ease of retrieval of the first noun, then one would not expect NP complexity effects. Rather, one would expect processing of the second noun's referent in both Complex and Simple NPs. Eye movement data confirm that looks to the second noun's referent were indeed rare for Simple NPs. Instead NP complexity effects were found even in writing and even when the lexical name of the first noun was easy to retrieve. This is strong evidence for an obligatory phrasal scope. Also, in line with Konopka (2012), one would expect evidence for advance lexical retrieval of N2 (priming effects), at least, for Complex NPs. The sparse evidence for lexical retrieval suggests that the ease of activating the first noun did not increase the planning span. Second, although there may be benefits for advance sentence planning beyond the obligatory unit in speech this is not true for writing. In speech, planning may go beyond the first noun, for example, to ensure fluency after production onset. However, while aiming to maximise output fluency, there is also a general tendency to minimise the need for buffering of linguistic material (Levelt & Meyer, 2000; Meyer, 1997). This is likely to be particularly important in writing as the difference in production speed – resulting from the more complex processing associated with orthographic retrieval and motor planning of typed output that requires writers to buffer information over a longer period of time (Gentner et al., 1988; Olive, 2014). Therefore although in speech ease of retrieval of the sentence-initial noun might have encouraged more advance planning, e.g. because of pressure to maintain fluency, in writing the opposite effect would be expected (i.e. the reduction of buffering demands). The same effects were observed in both conditions.

The conclusion that advance planning does not require the lexical specification of non sentence-initial nouns is based on (a) the failure to find priming effects in onset latencies in either modality in Experiment 2 or in writing in Experiment 3, and (b) on the assumption that codability effects in Experiment 1 was most parsimoniously attributed to increased prominence of the second noun's referent induced by the experimental setup. Alternatively however the absence of name priming effects in Experiment 2 and 3 may mean that the priming manipulation was not an effective strategy for increasing ease of lexical retrieval. The second noun may have been lexically prepared but the priming manipulation did not result in sufficient difference between speed of retrieval of primed and unprimed nouns for this to be detectable in production onset latencies. This is unlikely for three reasons. First, name priming effects were in fact observed in Experiment 3. Second, pilot data (see Appendix D) indicated name priming using the same materials in an image naming experiment. Third, in analyses not reported in this paper priming effects on production duration and of the attention to the image representing N2 *after* production onset were observed: The duration of the post-onset production process was generally shorter when N2 was lexically primed, and the proportion of eye samples to the image N2 was reduced if primed. These three reasons suggest that the lack of evidence for an effect of prime on production onset latency did not result simply from an ineffective priming manipulation.

Another possibility is that advance-planning may have been syntactically primed. Language users tend to recycle syntactic structures they heard or used recently (see Pickering & Ferreira, 2008). This might have affected our results in two ways. It may be that that participants did not engage in syntactic processing but rather learned to retrieve an intact syntactic frame in response to particular array movement patterns. However, if participants had repeatedly recalled structural templates from memory rather than actively engaging the linguistic processor, and assuming retrieval of Complex and Simple syntactic frames is equally time consuming, then there would be no NP complexity effects observed. Note also that to reduce the possibility of syntactic priming effects target arrays with either upwards or downwards movements were included as well as filler arrays that targeted structurally different sentences. The experimental design therefore made it impossible for participants to predict upcoming syntactic structures or movement patterns. As syntactic priming is subject to interference (Branigan et al., 1999), the variety of different movement patterns prevented sentence planning by mere retrieval of syntactic frames.

Increased production-onset latency for Complex NPs is therefore not readily explained as an artefact of the experimental design, but rather points towards obligatory, modality-independent planning of the initial noun phrase. It is possible that planning beyond the initial noun is perceptually or conceptually motivated. Griffin and Bock (2000) suggested that a visual "apprehension" of the stimulus screen serves the conceptualisation of the message. This apprehension is guided by the perceptual attraction of the larger moving unit increasing onset latencies and eye movements towards the target image. Martin et al. (2010, Experiment 4) addressed this concern directly by comparing a condition in which participants generated sentences similar to those elicited in the present study with a condition in which participants produced simple lists. They observed effects for sentences only (see also Zhao et al., 2015, Experiment 2 and 3). Note also that in the present context these effects would be similar in both Simple and Complex conditions, because the apprehension explanation does not differentiate between larger moving units on the left and right side of the screen.

In this paper we have used the term "writing" in a rather general way. A reviewer commented that findings from keyboard typing may not generalise to handwriting, or to other output modalities such as typing with thumbs on a smartphone keyboard. The main aim of our research was to determine whether planning scope findings are speech-specific. Our findings, from typed production, provide evidence against this modality-specificity hypothesis. It may indeed be that we would not find similar effects in handwriting. Note, though, that the features we identified in speech as possible alternative (non-syntactic) explanations for extended onset latencies for more complex sentences are absent in both typing and handwriting. Our prediction, therefore, is that effects are also present when output is handwritten.

A related issue, raised by the same reviewer, relates to effects of variation in typing proficiency. Planning scope findings from the relatively proficient typists that made up our sample might not generalise to people who type more slowly. For written naming, there is evidence that quite substantial differences in reported typing skill do not affect patterns of findings (see, for example, findings and discussion in Torrance et al., 2017). That notwithstanding, we think that is quite possible that we would have observed different effects had our sample comprised participants who did not type relatively fluently. There is, of course, also considerable variation is speaking speed (and the possibility that findings from advanced planning do not generalise to speakers with speech impairment but see Martin, Miller, & Vu, 2004). The possibility of different findings for people with different levels of competence or rates of output does not contradict our claim that we are able to draw sensible conclusions about obligatory planning scope on the basis of data from typists and speakers who are functionally-competent.

In this and previous papers a strong distinction has been made between syntax-based and lexical theories of sentence planning (e.g. Konopka & Bock, 2009; E.-K. Lee et al., 2013; Wheeldon et al., 2013). This distinction is arguably oversimplified. Planning scope is typically considered over the linear sequence of the surface string. In principle, however, there is no reason for planning scope to map linearly onto the output string (see e.g. Do, 2018; E.-K. Lee et al., 2013). Our present data, and results from previous results, give convincing evidence that there is some planning beyond the initial noun in coordinated noun phrases. However, this doesn't rule out the possibility that in both conditions retrieval of the verb was necessary prior to planning the initial phrase. This would be consistent with a (limited) lexical account of syntax planning (Bock & Levelt, 1994; F. Ferreira, 2000), although existing evidence for anticipation of the verb in sentence planning is, in fact, rather mixed (Do, 2018; Konopka & Meyer, 2014; Momma et al., 2015, but see, Brown-Schmidt & Konopka, 2008; Griffin, 2001; Schriefers, Teruel, & Meinshausen, 1998).

It also remains possible, however, that syntactic planning scope is driven by variation in the semantic representation of what needs to be expressed, i.e. scope effects are essentially a semantic rather than a syntactic (or lexical) effect (see e.g. Bunger, Papafragou, & Trueswell, 2013; Chang et al., 2006; Do, 2018; Konopka & Brown-Schmidt, 2014; Konopka & Meyer, 2014; Momma et al., 2015; Roeser, 2017). In this study planning dedicated to the second noun in complex NPs may remained pre-lexical. The present findings suggest that the presence of N2 in the initial noun phrase affected advance planning even in the absence of effects indicating lexical retrieval of N2s name. This pre-lexical identification of a placeholder may then serve to support the building of a syntactic "scaffold" (Bock & Ferreira, 2014) – a basic identification of the thematic agent (i.e. N1 and N2). The identification of the sentence's agent might underlie a semantic representation. To output a conjoined noun phrase the simultaneity of the entities' action needs to be encoded (i.e. two entities, the N1 and the N2, perform a mutual, in contrast to, for instance, an exclusive action of a single entity). As semantic conceptualisation is fundamental to build a syntactic representation, one cannot rule out that the NP complexity effect, here and in previous research, represents pre-syntactic semantic processing difficulty. Future research will be needed to determine the role of semantic or conceptual structure in sentence planning.

## Conclusion

The most parsimonious explanation of the present results is that sentence initial processing obligates advance planning of the syntactic structure of the sentence-initial phrase but permits lexical retrieval to be delayed until after production onset. Grammatical encoding beyond the first noun in conjoined noun phrases is therefore modality independent and best attributed to basic requirements of the language production system. The reported experiments are the only direct comparison of advance planning in spoken and written sentence production and the first systematic investigation of planning in written sentence production. This comparison provides strong evidence that, for reasons fundamental to the language system, planning in short sentence production has syntax-driven phrasal scope.

## Acknowledgements

This research was funded by the Nottingham Trent University VC Scholarship scheme awarded to the first author and is part of his PhD thesis.

# References

Ailhaud, E., & Chenu, F. (2018). Variations of chronometric measures of written production depending on clause packaging. *CogniTextes*, *17*. (in press)

Alario, F.-X., Costa, A., Ferreira, V. S., & Pickering, M. J. (2006). Architectures, representations and processes of language production. *Language and Cognitive Processes*, *21*(7-8), 777–789.

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

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

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

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

Barry, C. (1994). Spelling routes (or roots or rutes). In G. D. A. Brown & N. C. Ellis (Eds.), *Handbook of spelling: Theory, process and intervention* (pp. 27–49). Chichester, UK: Wiley.

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

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, *89*(1), 1–47.

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*(4), 575–586.

Bock, J. K. (1990). Structure in language: Creating form in talk. *American Psychologist*, *45*(11), 1221–1236.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*(1), 99–127.

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

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bock, J. K., & Miller, C. A. (1991). Broken agreement. *Cognitive psychology*, *23*(1), 45–93.

Bonin, P., & Fayol, M. (2000). Writing words from pictures: What representations are activated, and when? *Memory & Cognition*, *28*(4), 677–689.

Bonin, P., Fayol, M., & Gombert, J.-E. (1998). An experimental study of lexical access in the writing and naming of isolated words. *International Journal of Psychology*, *33*(4), 269–286.

Bonin, P., Malardier, N., Méot, A., & Fayol, M. (2006). The scope of advance planning in written picture naming. *Language and Cognitive Processes*, *21*(1–3), 205–237. doi: 10.1080/01690960400002141

Bonin, P., Méot, A., Laroche, B., Bugaiska, A., & Perret, C. (2017). The impact of image characteristics on written naming in adults. *Reading and Writing*, 1–19.

Bonin, P., Peereman, R., & Fayol, M. (2001). Do phonological codes constrain the selection of orthographic codes in written picture naming? *Journal of Memory and Language*, *45*(4), 688–720. doi: 10.1006/jmla.2000.2786

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech*, *8*(3), 148–158.

Box, G. E. P., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, *26*(2), 211–252.

Branigan, H. P., Pickering, M. J., & Cleland, A. A. (1999). Syntactic priming in written production: Evidence for rapid decay. *Psychonomic Bulletin & Review*, *6*(4), 635–640.

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

Bunger, A., Papafragou, A., & Trueswell, J. C. (2013). Event structure influences language production: Evidence from structural priming in motion event description. *Journal of Memory and Language*, *69*(3), 299–323.

Bürki, A., Sadat, J., Dubarry, A.-S., & Alario, F.-X. (2016). Sequential processing during noun phrase production. *Cognition*, *146*, 90–99.

Butterworth, B. (1975). Hesitation and semantic planning in speech. *Journal of Psycholinguistic research*, *4*(1), 75–87.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., ... Riddell, A. (2016). Stan: A probabilistic programming language. *Journal of Statistical Software*, *20*.

Chang, F., Bock, J. K., & Goldberg, A. E. (2003). Can thematic roles leave traces of their places? *Cognition*, *90*(1), 29–49.

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

Chang, F., Dell, G. S., Bock, J. K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. *Journal of Psycholinguistic Research*, *29*(2), 217–230.

Chukharev-Hudilainen, E., Saricaoglu, A., Torrance, M., & Feng, H.-H. (2018). Combined deployable keystroke logging and eyetracking for investigating L2 writing fluency. *Studies in Second Language Acquisition.* (in press)

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

Cleland, A. A., & Pickering, M. J. (2006). Do writing and employ the same syntactic representations? *Journal of Memory and Language*, *54*(2), 185–198.

Costa, A., & Caramazza, A. (2002). The production of noun phrases in English and Spanish: Implications for the scope of phonological encoding in speech production. *Journal of Memory and Language*, *46*(1), 178–198.

Damian, M. F., Dorjee, D., & Stadthagen-Gonzalez, H. (2011). Long-term repetition priming in spoken and written word production: Evidence for a contribution of phonology to handwriting. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(4), 813–826.

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

Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, *93*(3), 283–321.

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

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

Do, M. L. (2018). *How to think before you speak: Getting from abstract thoughts to sentences* (Unpublished doctoral dissertation). University of Southern California.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language*, *30*(2), 210–233.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. R. Wheeldon (Ed.), *Aspects of language production* (pp. 291–330). Hove, UK: Psychology Press.

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

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

Ferreira, V. S., & Slevc, L. R. (2007). Grammatical encoding. In M. G. Gaskell (Ed.), *The Oxford Handbook of Psycholinguistics* (pp. 453–470). Oxford, UK: Oxford University Press.

Foulin, J.-N. (1998). To what extent does pause location predict pause duration in adults' and children's writing? *Cahiers de Psychologie Cognitive/Current*

*Psychology of Cognition*, *17*(3), 601–620.

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

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

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

Gentner, D. R. (1982). Testing the central control model of typing: Comments on the reply by viviani and terzuolo. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(6), 814–816.

Gentner, D. R., Larochelle, S., & Grudin, J. (1988). Lexical, sublexical, and peripheral effects in skilled typewriting. *Cognitive Psychology*, *20*(4), 524–548.

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

Goldman-Eisler, F. (1972). Pauses, clauses, sentences. *Language and Speech*, *15*(2), 103–113. doi: 10.1177/002383097201500201

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

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

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

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

Hardy, S. M., Segaert, K., & Wheeldon, L. (2018). Healthy ageing and sentence production: Impaired lexical access in the context of intact syntactic planning. *bioRxiv*.

Hartsuiker, R. J., Bernolet, S., Schoonbaert, S., Speybroeck, S., & Vanderelst, D. (2008). Syntactic priming persists while the lexical boost decays: Evidence from written and spoken dialogue. *Journal of Memory and Language*, *58*(2), 214–238.

Hartsuiker, R. J., & Westenberg, C. (2000). Word order priming in written and spoken sentence production. *Cognition*, *75*(2), B27–B39.

Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, *14*(3), 277–288.

Henderson, A., Goldman-Eisler, F., & Skarbek, A. (1966). Sequential temporal patterns in spontaneous speech. *Language and Speech*, *9*(4), 207–216.

Hoffman, M. D., & Gelman, A. (2014). The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, *15*(1), 1593–1623.

Immonen, S., & Mäkisalo, J. (2017). Pauses reflecting the processing of syntactic units in monolingual text production and translation. *HERMES – Journal of Language and Communication in Business*, *23*(44), 45–61. doi: 10.7146/hjlcb.v23i44.97266

Kandel, S., & Perret, C. (2015). How does the interaction between spelling and motor processes build up during writing acquisition? *Cognition*, *136*, 325–336.

Kaufer, D. S., Hayes, J. R., & Flower, L. (1986). Composing written sentences. *Research in the Teaching of English*, *20*(2), 121–140.

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

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

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

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

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

Kruschke, J. K., Aguinis, H., & Joo, H. (2012). The time has come: Bayesian methods for data analysis in the organizational sciences. *Organizational Research Methods*, *15*(4), 722–752.

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

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology*, *99*(2), 199–208. doi: 10.1037/h0034633

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

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

Levelt, W. J. M. (1989). *Speaking: From intention to articulation* (Vol. 1). MIT press.

Levelt, W. J. M. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*(23), 13464–13471.

Levelt, W. J. M., & Maasen, B. (1981). Lexical search and order of mention in sentence production. In W. Kleine & W. J. M. Levelt (Eds.), *Crossing the boundaries in linguistics* (pp. 221–252). Dordrecht: Reidel.

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

Levelt, W. J. M., Schriefers, H., Vorberg, D., Meyer, A. S., Pechmann, T., & Havinga, J. (1991). The time course of lexical access in speech production: A study of picture naming. *Psychological review*, *98*(1), 122–142.

Lindsley, J. R. (1975). Producing simple utterances: How far ahead do we plan? *Cognitive Psychology*, *7*(1), 1–19.

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

Martin, R. C., Miller, M., & Vu, H. (2004). Lexical-semantic retention and speech production: Further evidence from normal and brain-damaged participants for a phrasal scope of planning. *Cognitive Neuropsychology*, *21*(6), 625–644.

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

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

Medler, D. A., & Binder, J. R. (2005). *MCWord: An on-line orthographic database of the English language.* `http://www.neuro.mcw.edu/mcword/`.

Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture–word interference experiments. *Journal of Memory and Language*, *35*(4), 477–496.

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

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

Momma, S., Slevc, L. R., & Phillips, C. (2015). The timing of verb selection in japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 813–824.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas – Part II. *arXiv preprint arXiv:1602.00245*.

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

Nottbusch, G., Grimm, A., Weingarten, R., & Will, U. (2005). Syllabic sructures in typing: Evidence from deaf writers. *Reading & Writing*, *18*(6), 497–526.

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

Olive, T. (2014). Toward a parallel and cascading model of the writing system: A review of research on writing processes coordination. *Journal of Writing Research*, *6*(2), 173–194.

Perret, C., Bonin, P., & Laganaro, M. (2014). Exploring the multiple-level hypothesis of AoA effects in spoken and written object naming using a topographic ERP analysis. *Brain and Language*, *135*, 20–31. doi: 10.1016/j.bandl.2014.04.006

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological bulletin*, *134*(3), 427–459.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in english and spanish. *Journal of Memory and Language*, *42*(2), 168–182.

Qu, Q., & Damian, M. F. (2017). Orthographic effects in spoken word recognition: Evidence from Chinese. *Psychonomic Bulletin & Review*, *24*(3), 901–906.

Rapp, B., Benzing, L., & Caramazza, A. (1997). The autonomy of lexical orthography. *Cognitive Neuropsychology*, *14*, 71–104.

Roeser, J. (2017). *Planning scope in spoken and written sentence production* (Unpublished doctoral dissertation). Nottingham Trent University.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: The role of surface detail in basic-level object recognition. *Perception*, *33*(2), 217–236.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science*, *6*(1), 1–36.

Sahel, S., Nottbusch, G., Blanken, G., & Weingarten, R. (2005). The role of phonology in syllabic structure in the time course of typing: Evidence from aphasia. *Linguistische Berichte*, *201*, 65–87.

Scaltritti, M., Arfé, B., Torrance, M., & Peressotti, F. (2016). Typing pictures: Linguistic processing cascades into finger movements. *Cognition*, *156*, 16–29. doi: 10.1016/j.cognition.2016.07.006

Schriefers, H., & Teruel, E. (1999). The production of noun phrases: A cross-linguistic comparison of French and German. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st annual conference of the cognitive science society* (pp. 637–642). Mahwah: Erlbaum.

Schriefers, H., Teruel, E., & Meinshausen, R. M. (1998). Producing simple sentences: Results from picture-word interference experiments. *Journal of Memory and Language*, *39*(4), 609–632.

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

Smith, M., & Wheeldon, L. R. (2001). Syntactic priming in spoken sentence production – an online study. *Cognition*, *78*, 123–164.

Smith, M., & Wheeldon, L. R. (2004). Horizontal information flow in spoken sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*(3), 675–686.

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

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

Stan Development Team. (2015). *Stan modeling language user's guide and reference manual.* http://mc-stan.org/. (R package version 2.6.1.)

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

Torrance, M., Johansson, R., Johansson, V., & Wengelin, Å. (2016). Reading during the composition of multi-sentence texts: an eye-movement study. *Psychological Research*, *80*(5), 729–743. doi: 10.1007/s00426-015-0683-8

Torrance, M., & Nottbusch, G. (2012). Written production of single words and simple sentences. In V. Berninger (Ed.), *Past, Present, and Future Contributions of Cognitive Writing Research to Cognitive Psychology* (pp. 403–422). New York: Taylor Francis.

Torrance, M., Nottbusch, G., Alves, R. A., Arfé, B., Chanquoy, L., Chukharev-Hudilainen, E., … others (2017). Timed written picture naming in 14 European languages. *Behavior Research Methods*, 1–15. doi: 10.3758/s13428-017-0902-x

Van Galen, G. P. (1991). Handwriting: Issues for a psychomotor theory. *Human movement science*, *10*(2), 165–191.

Van Hell, J. G., Verhoeven, L., & Van Beijsterveldt, L. M. (2008). Pause time patterns in writing narrative and expository texts by children and adults. *Discourse Processes*, *45*(4–5), 406–427. doi: 10.1080/01638530802070080

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

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

Wheeldon, L. R. (2011). Generating spoken sentences: The relationship between words and syntax. *Language and Linguistics Compass*, *5/6*, 310–321.

Wheeldon, L. R. (2012). Producing spoken sentences: The scope of incremental planning. In S. Fuchs, M. Weirich, D. Pape, & P. Perrier (Eds.), *Speech planning and dynamics* (Vol. 1, pp. 97–118). Frankfurt, Germany: Peter Lang.

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

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

Zhang, Q., & Damian, M. F. (2010). Impact of phonology on the generation of handwritten responses: Evidence from picture-word interference tasks. *Memory & Cognition*, *38*(4), 519–528.

Zhao, L.-M., Alario, F.-X., & Yang, Y.-F. (2015). Grammatical planning scope in sentence production: Further evidence for the functional phrase hypothesis. *Applied Psycholinguistics*, *36*, 1059–1075.

Zhao, L.-M., & Yang, Y.-F. (2013). Grammatical planning scope in sentence production: Evidence from Chinese sentences. *Acta Psychologica Sinica*, *45*, 599–613.

Zhao, L.-M., & Yang, Y.-F. (2016). Lexical planning in sentence production is highly incremental: Evidence from ERPs. *PloS one*, *11*(1), 1–20.

## Appendix A

List of stimulus images for second noun (Experiment 1)

| Item | File ID | Image | *H* | Codability | Movement N1 |
|---|---|---|---|---|---|
| 1 | 007 | arm | 0.1 | high | down |
| 1 | 002 | plane | 1.3 | low | down |
| 2 | 015 | balloon | 0.0 | high | down |
| 2 | 023 | fly | 2.5 | low | down |
| 3 | 016 | banana | 0.0 | high | down |
| 3 | 024 | beetle | 1.7 | low | down |
| 4 | 021 | bear | 0.1 | high | down |
| 4 | 029 | shirt | 2.2 | low | down |
| 5 | 022 | bed | 0.1 | high | down |
| 5 | 037 | broom | 1.3 | low | down |
| 6 | 025 | bell | 0.0 | high | down |
| 6 | 046 | hat | 1.4 | low | down |
| 7 | 030 | book | 0.0 | high | down |
| 7 | 055 | chicken | 1.0 | low | down |
| 8 | 040 | butterfly | 0.0 | high | down |
| 8 | 064 | coat | 1.4 | low | down |
| 9 | 042 | cake | 0.0 | high | down |
| 9 | 066 | corn | 1.3 | low | down |
| 10 | 043 | camel | 0.0 | high | down |
| 10 | 067 | sofa | 1.1 | low | down |
| 11 | 044 | candle | 0.0 | high | down |

*(continued)*

| Item | File ID | Image | $H$ | Codability | Movement N1 |
|---|---|---|---|---|---|
| 11 | 070 | cup | 1.1 | low | down |
| 12 | 048 | carrot | 0.0 | high | down |
| 12 | 071 | deer | 1.1 | low | down |
| 13 | 049 | cat | 0.0 | high | up |
| 13 | 072 | desk | 1.3 | low | up |
| 14 | 052 | chain | 0.1 | high | up |
| 14 | 077 | door knob | 1.5 | low | up |
| 15 | 053 | chair | 0.1 | high | up |
| 15 | 079 | drawers | 1.1 | low | up |
| 16 | 054 | cherry | 0.1 | high | up |
| 16 | 082 | eagle | 1.6 | low | up |
| 17 | 060 | clock | 0.0 | high | up |
| 17 | 092 | flute | 1.0 | low | up |
| 18 | 063 | clown | 0.0 | high | up |
| 18 | 101 | pan | 1.2 | low | up |
| 19 | 069 | crown | 0.1 | high | up |
| 19 | 116 | hanger | 1.3 | low | up |
| 20 | 073 | dog | 0.1 | high | up |
| 20 | 136 | leopard | 1.6 | low | up |
| 21 | 076 | door | 0.0 | high | up |
| 21 | 137 | lettuce | 1.1 | low | up |

*(continued)*

| Item | File ID | Image | $H$ | Codability | Movement N1 |
|---|---|---|---|---|---|
| 22 | 078 | dress | 0.1 | high | up |
| 22 | 138 | lightbulb | 1.1 | low | up |
| 23 | 083 | ear | 0.0 | high | up |
| 23 | 139 | switch | 1.0 | low | up |
| 24 | 084 | elephant | 0.0 | high | up |
| 24 | 143 | padlock | 1.1 | low | up |
| 25 | 089 | fish | 0.0 | high | up |
| 25 | 144 | glove | 1.7 | low | up |
| 26 | 090 | flag | 0.1 | high | up |
| 26 | 147 | motorbike | 1.0 | low | up |
| 27 | 097 | fork | 0.0 | high | up |
| 27 | 151 | nail | 1.1 | low | up |
| 28 | 098 | fox | 0.1 | high | up |
| 28 | 152 | nail file | 2.4 | low | up |
| 29 | 103 | giraffe | 0.0 | high | up |
| 29 | 153 | necklace | 1.1 | low | up |
| 30 | 105 | glasses | 0.1 | high | up |
| 30 | 161 | paint brush | 1.1 | low | up |
| 31 | 106 | glove | 0.0 | high | up |
| 31 | 163 | peach | 1.6 | low | up |
| 32 | 114 | hammer | 0.0 | high | up |

*(continued)*

| Item | File ID | Image | *H* | Codability | Movement N1 |
|---|---|---|---|---|---|
| 32 | 178 | bag | 1.1 | low | up |
| 33 | 115 | hand | 0.0 | high | up |
| 33 | 179 | pan | 1.5 | low | up |
| 34 | 118 | hat | 0.0 | high | up |
| 34 | 183 | racoon | 1.7 | low | up |
| 35 | 121 | horse | 0.0 | high | up |
| 35 | 189 | roller skate | 2.3 | low | up |
| 36 | 123 | iron | 0.0 | high | up |
| 36 | 191 | chicken | 2.3 | low | up |
| 37 | 128 | key | 0.0 | high | down |
| 37 | 193 | boat | 1.2 | low | down |
| 38 | 129 | kite | 0.0 | high | down |
| 38 | 194 | salt | 1.4 | low | down |
| 39 | 131 | ladder | 0.1 | high | down |
| 39 | 214 | thread | 2.4 | low | down |
| 40 | 135 | lemon | 0.1 | high | down |
| 40 | 221 | suitcase | 1.2 | low | down |
| 41 | 140 | lion | 0.1 | high | down |
| 41 | 228 | television | 1.2 | low | down |
| 42 | 150 | mushroom | 0.0 | high | down |
| 42 | 229 | racket | 2.0 | low | down |

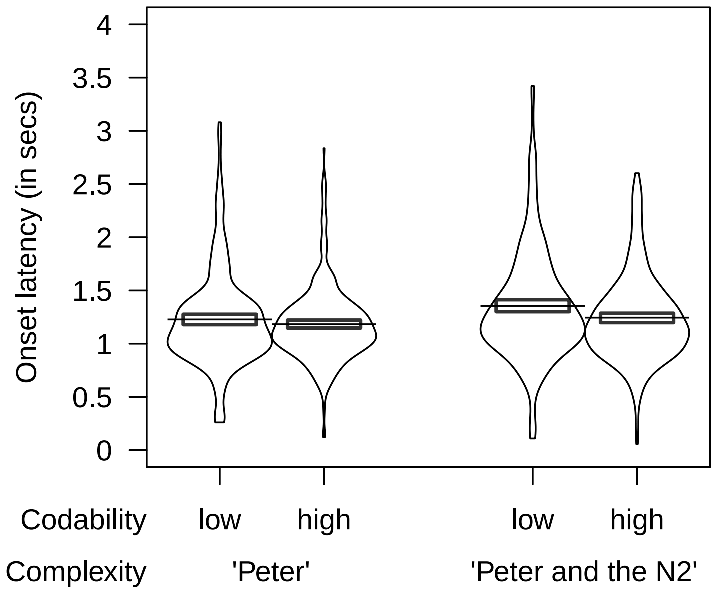
*(continued)*

| Item | File ID | Image | *H* | Codability | Movement N1 |
|---|---|---|---|---|---|
| 43 | 155 | nose | 0.0 | high | down |
| 43 | 235 | thumb | 1.1 | low | down |
| 44 | 158 | orange | 0.0 | high | down |
| 44 | 239 | traffic lights | 1.5 | low | down |
| 45 | 160 | owl | 0.0 | high | down |
| 45 | 242 | lorry | 1.1 | low | down |
| 46 | 166 | pear | 0.0 | high | down |
| 46 | 247 | waistcoat | 1.0 | low | down |
| 47 | 167 | pen | 0.1 | high | down |
| 47 | 252 | watermelon | 1.1 | low | down |
| 48 | 168 | pencil | 0.0 | high | down |
| 48 | 258 | glass | 1.2 | low | down |

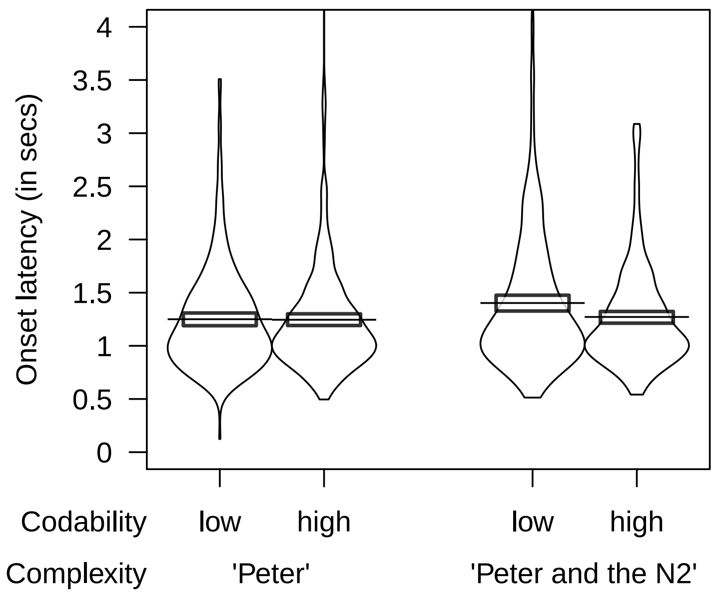
*Note:* File ID indicates the Rossion and Pourtois (2004) image

# Appendix B

## Onset latency: Experiment 1




(a) Speech




(b) Writing

Appendix C

List of stimulus images for second noun (Experiment 2, 3)

| Item | File ID | Image/prime | non-word | Movement N1 |
|---|---|---|---|---|
| 1 | 002 | plane | btzjv | up |
| 2 | 003 | crocodile | ljegvomxp | up |
| 3 | 005 | ant | hhv | up |
| 4 | 008 | arrow | dgnms | up |
| 5 | 010 | ashtray | dsfbphn | up |
| 6 | 013 | pram | wxfk | up |
| 7 | 017 | barn | tllw | up |
| 8 | 018 | barrel | rmdyzv | up |
| 9 | 019 | bat | zhj | up |
| 10 | 023 | fly | xgf | up |
| 11 | 024 | beetle | auxinh | up |
| 12 | 027 | bike | lnwv | up |
| 13 | 029 | shirt | kyyig | up |
| 14 | 031 | boot | ejzn | up |
| 15 | 033 | bow | qfi | up |
| 16 | 037 | broom | wwmgi | up |
| 17 | 038 | brush | lkscl | up |
| 18 | 046 | hat | qji | up |
| 19 | 055 | chicken | phbgnoz | up |
| 20 | 059 | cigarette | lkpcsoddi | up |
| 21 | 064 | coat | hhfm | up |

*(continued)*

| Item | File ID | Image/prime | non-word | Movement N1 |
|---|---|---|---|---|
| 22 | 066 | corn | ieqv | up |
| 23 | 067 | sofa | uqcm | up |
| 24 | 070 | cup | jyd | up |
| 25 | 071 | deer | jlra | up |
| 26 | 072 | desk | xjeu | up |
| 27 | 074 | doll | cgxl | up |
| 28 | 077 | door knob | lrgi gzjc | up |
| 29 | 079 | drawers | rhxljgc | up |
| 30 | 080 | drum | mohy | up |
| 31 | 082 | eagle | kdivy | up |
| 32 | 085 | envelope | dnrrmeef | up |
| 33 | 087 | fence | aewgv | up |
| 34 | 088 | finger | vkkvbl | up |
| 35 | 092 | flute | nnqxe | up |
| 36 | 093 | fly | kzb | up |
| 37 | 099 | trumpet | dmxvzhd | up |
| 38 | 101 | pan | wbv | up |
| 39 | 102 | bin | fnh | up |
| 40 | 107 | goat | pkdp | up |
| 41 | 108 | gorilla | opvjvac | up |
| 42 | 116 | hanger | fqctgn | up |

*(continued)*

| Item | File ID | Image/prime | non-word | Movement N1 |
|---|---|---|---|---|
| 43 | 117 | harp | dqnm | up |
| 44 | 122 | house | ytusn | up |
| 45 | 125 | coat | pivv | up |
| 46 | 127 | kettle | mzbksz | up |
| 47 | 136 | leopard | duskwlf | up |
| 48 | 137 | lettuce | zhbqzmz | up |
| 49 | 138 | lightbulb | ekfzhqfof | down |
| 50 | 139 | switch | mjqaha | down |
| 51 | 142 | lobster | ygvpqat | down |
| 52 | 143 | padlock | bpegwik | down |
| 53 | 144 | glove | rlgmw | down |
| 54 | 145 | monkey | hfiqie | down |
| 55 | 147 | motorbike | odjhpcihb | down |
| 56 | 148 | mountain | srfnrqbg | down |
| 57 | 149 | mouse | ufyjk | down |
| 58 | 151 | nail | kvld | down |
| 59 | 152 | nail file | vhtl dupj | down |
| 60 | 153 | necklace | ddulweod | down |
| 61 | 154 | needle | rhvxbp | down |
| 62 | 156 | nut | dfj | down |
| 63 | 159 | ostrich | bdisddl | down |

*(continued)*

| Item | File ID | Image/prime | non-word | Movement N1 |
|---|---|---|---|---|
| 64 | 163 | peach | duuaq | down |
| 65 | 165 | peanut | hzhwiy | down |
| 66 | 170 | pepper | bttbts | down |
| 67 | 174 | pipe | jtqk | down |
| 68 | 175 | jug | uzc | down |
| 69 | 177 | plug | wstb | down |
| 70 | 178 | bag | yqr | down |
| 71 | 179 | pan | svo | down |
| 72 | 182 | rabbit | aocxrr | down |
| 73 | 183 | racoon | jssrcc | down |
| 74 | 186 | rhino | teweg | down |
| 75 | 193 | boat | mtjq | down |
| 76 | 194 | salt | esga | down |
| 77 | 198 | screw | kiqpm | down |
| 78 | 201 | seal | qsxo | down |
| 79 | 202 | sheep | yunow | down |
| 80 | 206 | skunk | djyvs | down |
| 81 | 207 | sledge | finokx | down |
| 82 | 214 | thread | pojvmr | down |
| 83 | 219 | oven | hzeb | down |
| 84 | 221 | suitcase | izdumtrx | down |

*(continued)*

| Item | File ID | Image/prime | non-word | Movement N1 |
|---|---|---|---|---|
| 85 | 223 | swan | gdnm | down |
| 86 | 227 | telephone | cvpqiqsnn | down |
| 87 | 228 | television | vzibzsnbkc | down |
| 88 | 229 | racket | axdlpk | down |
| 89 | 235 | thumb | qgeex | down |
| 90 | 242 | lorry | gxbys | down |
| 91 | 243 | trumpet | gwutkfr | down |
| 92 | 244 | turtle | isicre | down |
| 93 | 247 | waistcoat | moyghnwqe | down |
| 94 | 248 | violin | ulofar | down |
| 95 | 252 | watermelon | tmpbeneklp | down |
| 96 | 258 | glass | bzlfz | down |

*Note:* File ID indicates the Rossion and Pourtois (2004) image

# Appendix D

## Pilot: priming experiment

In a pilot study we tested whether name priming task facilitates image naming and hence, the access of the image's name in the mental storage. Ten native speakers of British English (6 female, mean age = 27, $SD = 6.6$, range: 20–43) were asked to write (i.e. keyboard typing) the names of 95 low codable (mean $H = 1.1$, $SD = 0.51$, range: 0.4–2.5) coloured Snodgrass images (Rossion & Pourtois, 2004; Snodgrass & Vanderwart, 1980). Each image was either presented with or without prime. The prime was either the most commonly given name of the image – extracted from naming data recorded for the same population (Torrance et al., 2017) – or a length matched non-word – generated by the CELEX data base (Medler & Binder, 2005). Every trial started with a centred fixation cross on the position where the image will appear (800 ms). Images were presented in the centre of the screen simultaneously with the superimposed prime. The prime was presented either 50 ms or 80 ms followed by a mask (20 ms). Each of the 95 images was presented in all condition but only presented once per participant. Image items were distributed across five Latin square lists and presented in random order. 95 out of the 96 images used as stimulus material for Experiments 2 and 3 (see Appendix C) were tested due to counterbalancing constraints.

Prior to analysis we removed trials with onset latency longer than 10,000 ms (0.53%). Table D1 shows the descriptive data of the onset latency and the proportion of responses using the most commonly given name by condition.

For analysis we used the reciprocal of the onset latency (multiplied by 1000) to account for skew. Treatment contrasts were used with the no prime condition as baseline – each condition was compared to the no prime baseline. The results of the

Table D1

*Descriptive data summary of the onset latency (in ms) and the proportion of responses using the most commonly given name by prime type and prime duration (pilot)*

| Prime | | Latency | | Pr(MCN) | | |
|---|---|---|---|---|---|---|
| Type | Duration | *M* | *SE* | *M* | *SE* | *N* |
| no prime | *NA* | 1555 | 80 | .68 | .03 | 189 |
| image name | 50 | 1378 | 66 | .84 | .03 | 189 |
| image name | 80 | 1229 | 51 | .89 | .02 | 190 |
| non-word | 50 | 1624 | 67 | .72 | .03 | 188 |
| non-word | 80 | 1688 | 62 | .69 | .03 | 189 |

*Note:* Pr(MCN) = proportion of responses using the most commonly given name for a particular image

Bayesian linear mixed model are summarised in Table D2. The model revealed unsubstantial support (BF $<$ 1) for image name primes presented 50 ms, in spite of the numerically larger proportion of positive samples, but weak evidence supporting a priming effect for 80 ms presentation duration (BF = 2.4) showing shorter latencies. For non-word primes the model showed negligible evidence (BF $<$ 1) for the negative priming effect for 50 ms as indicated by the distribution of posterior samples but strong evidence supporting this effect for 80 ms priming duration (BF = 53). Also we calculated priming effects from the posterior samples of the model comparing image names and non-word primes. For 50 ms there was weak evidence (BF = 2.8) for a priming effect showing shorter latencies for image names compared to non-words ($\hat{\mu}$ = 0.16, 95% CrI[0.06, 0.26]). Strong evidence (BF $>$ 100) for a priming effect was found for 80 ms priming duration ($\hat{\mu}$ = 0.27, 95% CrI[0.16, 0.38]).

Table D2

*Bayesian linear mixed model on onset latency. Contrasts were treatment coded with no prime as baseline condition, i.e. estimates show the difference of each condition compared to the no prime responses (pilot)*

| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| image name (50 ms) | 0.09 | -0.01 | 0.19 |
| image name (80 ms) | 0.16 | 0.05 | 0.27 |
| non-word (50 ms) | -0.07 | -0.12 | -0.02 |
| non-word (80 ms) | -0.11 | -0.15 | -0.07 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

Further the proportion of responses using the most commonly given name was analysed in a Bayesian generalized mixed effects model using a Bernoulli distribution for binomial data. The results are shown in Table D3. The proportion of using the most commonly given name increased for image name primes for both 50 ms (BF = 26) and for 80 ms (BF $>$ 100) priming duration. The proportion of responses using the most commonly given name remained unchanged for non-word primes at both 50 ms (BF $<$ 1) and 80 ms (BF $<$ 1) priming duration. Comparisons between image name and non-word primes support this effect moderately (BF = 5) for 50 ms priming during ($\hat{\mu}$ = 1.09, 95% CrI[0.02, 2.33]) and substantially (BF = 82) for 80 ms priming duration ($\hat{\mu}$ = 2.28, 95% CrI[0.87, 4.09]).

In sum, image name primes showed shorter onset latencies and led to a larger probability of using the most commonly used image name as response. Non word

Table D3

*Bayesian generalized mixed model on the proportion of responses corresponding to the most commonly given names. Contrasts were treatment coded with no prime as baseline condition, i.e. all conditions were compared to the no prime condition (pilot)*

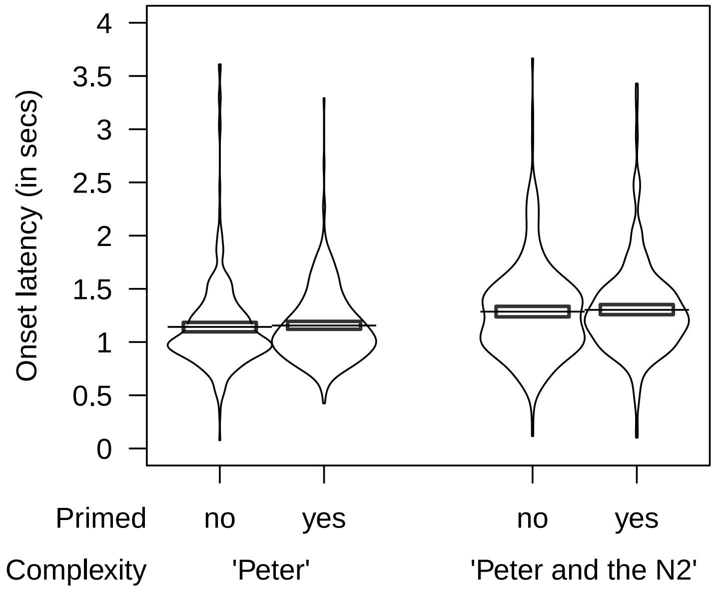
| | $\hat{\mu}$ | 2.5% | 97.5% |
|---|---|---|---|
| image name (50ms) | 1.34 | 0.41 | 2.49 |
| image name (80ms) | 2.35 | 1.02 | 4.16 |
| non-word (50ms) | 0.25 | -0.37 | 0.89 |
| non-word (80ms) | 0.06 | -0.55 | 0.72 |

*Note:* $\hat{\mu}$ = effect magnitude; 2.5% and 97.5% = 95% CrI

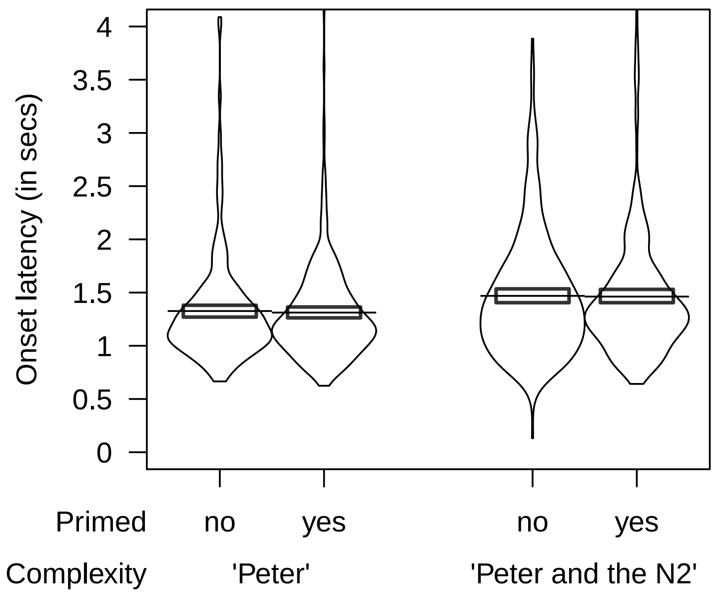
primes increased to onset latency while there was no change in the probability of using the most common image name compared to the no prime baseline. These results demonstrate that the prime facilitated naming and hence, lexical retrieval.

## Appendix E

### Onset latency: Experiment 2



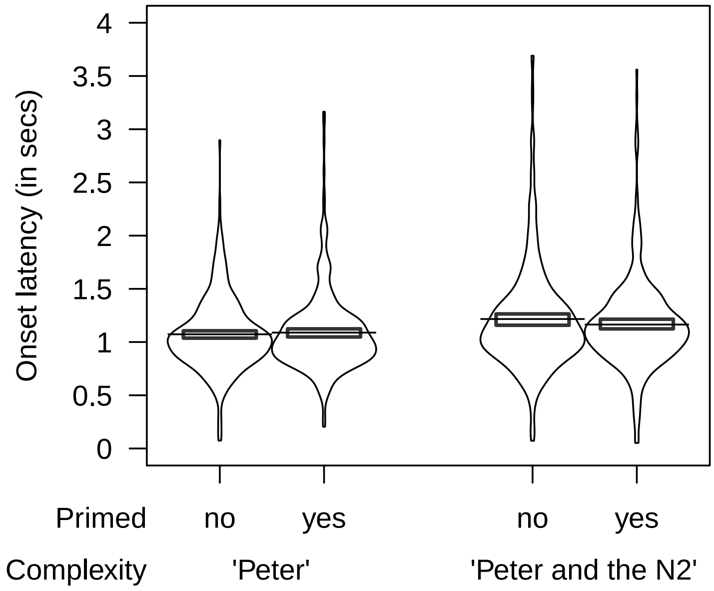

(a) Speech



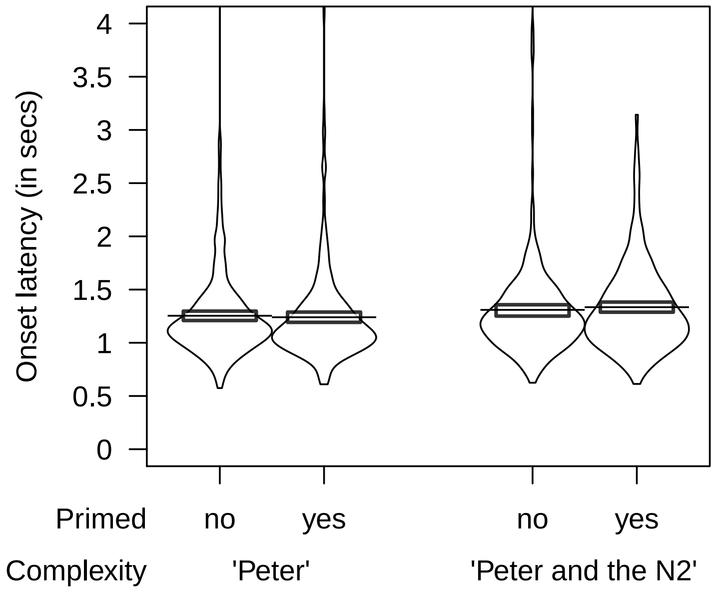

(b) Writing

# Appendix F

## Onset latency: Experiment 3




(a) Speech




(b) Writing