Planning scope in spoken and written sentence production

JENS ROESER

A thesis submitted in partial fulfilment of the requirements
of Nottingham Trent University for the degree of

Doctor of Philosophy

September 2017

Acknowledgements

This PhD would have been impossible without Mark Torrance who (has) spent uncountable numbers of hours on discussing every aspect of this research with me. Thanks Mark, for your critical thinking and your support before and throughout (and after) my PhD, for encouraging and believing in me, for providing me with the opportunity to freely development my own ideas, and for being an awesome advisor and person. Also, I want to thank Thom Baguley for his advice on the design and statistics of this thesis and for getting administrative obstacles out of my way. Further, thanks to my examiners Markus Damian and Clare Wood for the thorough review of, many helpful comments on, and a great discussion about this thesis. I really appreciate the time you invested in reviewing all aspects about my thesis. I'm also grateful for Natalie Braber's help and Gary Jones's support during my viva preparations. I also want to express my gratitude to Mark Andrews for all encouragement and inspiration he is (unknowingly) providing, many discussions about more than statistical modelling and lots of daily wisdom.

Furthermore, I want to thank Nottingham Trent University for supporting me in form of the Vice-Chancellors's Scholar-

ship. I'm very proud for having received this bursary as it opened up many possibilities to embrace my research and to present my work at many great conferences. My life would have been different without. More financial support was provided by the Experimental Psychology Society (Grindley grant), the European Literacy Network (COST Action IS1401), and by Santander (travelling bursary). All of these groups provided the necessary financial support to present my work at international conferences.

Thanks to all conference attendees that engaged with my work, especially at the *Writing Words Workshops* (Poitier, 2015; Bristol, 2017), *Architectures and Mechanisms of Language Production* (at Valletta, 2015; Bilbao, 2016; Lancaster, 2017), and *The International Workshop of Language Production* (San Diego, 2016), and at NTU's PhD Psychology conference (2015–2017). Although I'm grateful for everybody's comments and questions, I want to name a few people who are a constant source of inspiration and out-of-the-box thinking – my colleagues and friends Monica Do, Sara Farshchi and Rowena "Weng" Garcia.

I want to thank my family – for being my family – especially my mother Michaela, my brothers Hanko and Jan, my grandmother Charlotte Wesemann and my cousin Robert Giebel. There is no way in which I could possibly express my gratitude to you, Alexandra Arntsen. Thanks for all your patience.

Finally, thanks to all participants that took part in my experiments, to the three anonymous JML reviewers that engaged with my work and many other anonymous reviewers of conference abstracts.

To my mother – Michaela Roeser.

Abstract

This thesis investigates two questions about the cognitive mechanisms underlying the advance preparation of sentences. First, how much planning does the language system require to begin outputting a sentence and second, how is this scope determined. Previous research has concluded that advance planning embraces less than the sentence, is determined by either content or structure of some minimal linguistic unit, and is subject to variation (V. S. Ferreira & Slevc, 2007). Unlike previous research, the presented hypotheses were evaluated in both speech and writing. This eliminates explanations in terms of mechanisms that are modality specific, and therefore not fundamental to the language production system (see Alario, Costa, Ferreira, & Pickering, 2006). In two series of three experiments I elicited short sentences in speech and writing (keyboard typing). Under controlled conditions I manipulated (a) structural and lexical properties of elicited sentences (first series, Chapter 2) and (b) conceptual properties of the sentence's message (second series, Chapter 3). Hypotheses were evaluated by measurement of the time required to initiate output of the target sentence and of eye movements to referents of this sentence (arrays of simple line drawings)

shown on the computer screen. These suggested two main conclusions: (1) Consistent with some previous research advance planning scopes over coordinated noun phrases (*A and the B*) while lexical content requires planning for the first noun but not beyond (Chapter 2), demonstrating for the first time that this effect replicates in writing. (2) Whether or not noun phrases are preplanned beyond the first noun is determined at a conceptual level, and not at a syntactic level (Chapter 3). These findings are in line with current models of language production (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014) and constitute a first step towards confirming the modality independence of these models.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Aims & objectives

This thesis aims at developing an understanding of the cognitive processing mechanism underlying advance planning of sentences in the context of writing compared to speech. This is important for the following reasons: (1) Written communication is ubiquitous. Alongside traditional written output, SMS, email, social media and messaging have meant that a very broad cross section of the population use writing on a daily basis. However our understanding of the basic underlying processes of writing is poor. (2) Psycholinguistic theories of advance planning in language production are based on speech conflating fundamental and speech-specific language processing constraints. Findings from speech require a triangulation with those from written production. (3) Understanding basic processes in writing has the potential to inform research and practice in education, digital media research and other applied disciplines (e.g. human-computer interaction).

To develop an understanding of written sentence production, this thesis will address key issues regarding the architecture of higher order processes underlying advance planning in sentence production. Modality differences are taken into account by directly comparing data from both speech and writing. Specifically, I am going to establish whether findings from advance planning of simple sentences reproduce across output modality. This will be achieved by testing explicit hypotheses regarding the mechanisms underlying two obligatory stages of advance planning. First I examined how messages are encoded by the language production system before an utterance can begin (Chapter 2). Second I tested whether the extent of advance planning that is required by the production system is determined at message conceptualisation (i.e. prior to processing of syntax) (Chapter 3).

Writing, as I will show, imposes constraints that are different from those associated with speech. As sentence planning is well known to anticipate contextual non-linguistic factors during advance planning (F. Ferreira & Swets, 2002; Swets, Jacovina, & Gerrig, 2014; Wagner, Jescheniak, & Schriefers, 2010), the choice of output modality may determine the way that mental representations of sentences are prepared. If sentence planning were indeed modality specific, existing theories about underlying mechanisms, that were largely derived from spoken data, would not be general to the domain of human language but merely speech-specific. Comparing data from writing and speech is therefore essential to tease apart planning requirements that are imposed by the language processor and those that address modality-specific constraints.

1.2 Language planning

Humans with normal language ability are able to produce simple (and even complex) sentences fluently and without effort. We take communication for granted, which is remarkable, as the production of a sentence involves the orchestration of several complex cognitive processes in order to produce language. Exploring these processes is both interesting and difficult because people do generally lack conscious awareness and therefore introspective access to the processing involved.

Psycholinguistics in general aims to provide a theory that accounts for various aspects of the human language processing that enable us to communicate to each other by producing and understanding language. This ability allows us to convert ideas or "thoughts" into an utterance that is understandable by a receiver or interlocutor. To be able to produce an utterance we need to complete a number of mental processes. The idea we wish to communicate needs to be conceptualised into a lexical representation. For instance imagine I show you an image and I ask you to tell me what you see on it. This image depicts a cat, so you might correctly say "cat". To be able to complete this task, the image with the cat needs to be visually encoded; i.e. you need to recognise it. Once you have recognised the image, conceptual-semantic information is activated: domesticated, furry animal that is notorious for chasing mice. At this stage a conceptual mental representation has been created. This conceptual representation might still be lexically unspecified. If the image did not contain a cat but, say, a set of uilleann pipes you might recognise that this is a musical instrument but you might not remember the word "bagpipes" or you might experience the tip-of-the-tongue phenomenon (Brown & McNeill, 1966). This phenomenon refers to the activation of a not unfamiliar conceptual representation and possibly features of the word (i.e. starts in a /b/, has two syllables), but not a lexical representation of it. Once the name of the image is retrieved, it can be submitted to encoding in a phonological or orthographic representation that activates the associated motor plan that controls all physiological components involved to pronounce or to write the name of the image.

Uttering the single word "cat" might not be informative enough, if the cat shown in the image is engaged in an action that is sufficiently relevant to be mentioned. Let's say the cat in the image is biting its own tail. The target sentence might, thus, be "the cat is biting its tail". "The cat" became the *topic* of an event that is encoded as the *comment* "is biting its tail". This extension adds additional complexity to the planning process.

This raises a question that is central to the research that I report in this thesis. To what extent does this planning need to be completed in advance of output (i.e. before the speaker starts speaking or writer starts writing). The whole sentence could be mentally planned before we start addressing an interlocutor. On the other hand, we could plan the first noun, and think about the remainder after we said "the cat" but how do we know that the sentence needs to start with "cat"? It seems intuitive that we often start talking or writing before we actually know what we want to express or how we want to say it. Usually we manage somehow to add all information and end up with a more or less coherent sentence. This raises the interesting question as to how we prepare the production of language in our mind before we start talking, or writing. The umbrella term that arches over these processes is *advance planning* in language production.

Each utterance has to start with a message – a thought or a conceptual representation of what is going to be expressed – which is triggered by a communicative need or intention. This conceptual representation involves entities and objects, e.g. CAT and TAIL (capitals indicate prelexical representations), and a semantic representation of their relation (**who** does **what** to **whom**), i.e. BITING. These representations are typically thought of as being unordered (Bock, 1982; Konopka & Brown-Schmidt, 2014; Levelt, Le Page, & Longuet-Higgins, 1981). The conceptual representation needs to undergo a process in the mental language production system that is commonly referred to as *grammatical encoding* (Bock & Ferreira, 2014; Bock & Levelt, 1994; V. S. Ferreira & Slevc, 2007; Levelt, 1989). This process involves the generation of appropriate lexical representations with the morphosyntactic properties (e.g. gender) and the correct form of the target language, e.g. "cat", "katt", or "pusa". Further the syntactic structure that expresses the correct relationship between the message elements needs to be generated, e.g. "cat" – not "tail" – needs to be in the first position of the sentence. Finally, the phonetically and/or orthographically specified representations are submitted to the generation of motor codes that allow the mouth and the vocal cords to produce speech or the hands to output writing.

Theoretically speaking, one could preplan short or sometimes even long sentences in mind before outputting the first word. However, in practice – e.g. lively discussions with friends – we have barely any time to think about how to say something without over-stretching the patience of our interlocutor or having somebody else intrude into the discussion. Planning large amounts of language in advance is time consuming and it is cognitively exhausting to buffer large chunks of linguistic information (Christiansen & Chater, 2016; Levelt & Meyer, 2000). Producing a syntactically coherent utterance typically involves neither a lot of time

(Levelt & Maasen, 1981; Lindsley, 1975; Wagner et al., 2010; Wheeldon, Ohlson, Ashby, & Gator, 2013) nor conscious knowledge about the grammatical constraints of our mother tongue. Language planning is commonly understood as unfolding in a piece-meal – incremental – fashion (see V. S. Ferreira & Slevc, 2007) with linguistic planning units smaller than a clause (e.g. Smith & Wheeldon, 1999) or even smaller than a phrase (e.g. Griffin, 2001). While some information needs to be planned in advance to begin the utterance, further language planning is postponed until after production onset (V. S. Ferreira & Slevc, 2007; Kempen & Hoenkamp, 1987; Levelt, 1989). Which cognitive processes need to be completed at minimum before the production of a sentence begins is subject to an ongoing discussion. The extent to which these processes need to be completed before production onset will be referred to as *planning scope*.

The processes involved in language production are generally described as hierarchically organised into discrete modules – the conceptualisation, grammatical encoding, output (Bock & Ferreira, 2014; V. S. Ferreira & Slevc, 2007; Levelt, 1989). An illustration of the information flow between these modules can be found in Figure 1.1. This model is organised in hierarchically separated modules. Each module is associated with a process that has to be completed before the information flow into the next, lower process can begin. Each processing stage creates the input for the subsequent stage. Higher levels of organisation are commonly assumed to process larger language units than the lower levels (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; Garrett, 1975; Olive, 2014; Van Galen, 1991). This means that the conceptualisation stage generates propositions or messages that are then translated into lexical/syntactic units or a word or phrase, which are then processed as morphemes, phonemes or graphemes to be produced sequentially. The information flow between these modules is described as cascading from central to peripheral processes and occurs in parallel for different planning units (Alario et al., 2004; Humphreys, Riddoch, & Quinlan, 1988; Olive, 2014). Central processes – conceptualisation and grammatical encoding – generate information that flows continuously into the peripheral process – execution of speech or writing. In cascading architectures processing modules are not encapsulated but allow feedback from lower to higher levels. While each linguistic unit (or segment) has to be processed in a serial order from concept to output, processing of subsequent units can unfold in parallel and thus before processing for the first increment was completed.

Parallel cascading architectures have been proposed for language production in different modalities such as handwriting (Olive, 2014; Van Galen, 1991), typing (Gentner, 1983; Rumelhart & Norman, 1982) and speech (Dell, 1986; Levelt, 1989). Theories differ, however, in the extent to which information flow is cascaded: whether this occurs within lexical representations (Levelt, Roelofs, & Meyer, 1999), from conceptual representations to the syntactic frame (Bock & Ferreira, 2014), or from central to peripheral processes (Dell, 1986; Humphreys et al., 1988). While planning and execution can occur in parallel, sentence planning needs to unfold incrementally, at least to some extent (V. S. Ferreira &

Slevc, 2007; Kempen & Hoenkamp, 1987; Levelt, 1989). For example the product of the conceptual processing stage has be submitted to the grammatical encoder before processing of the next language increment can begin.

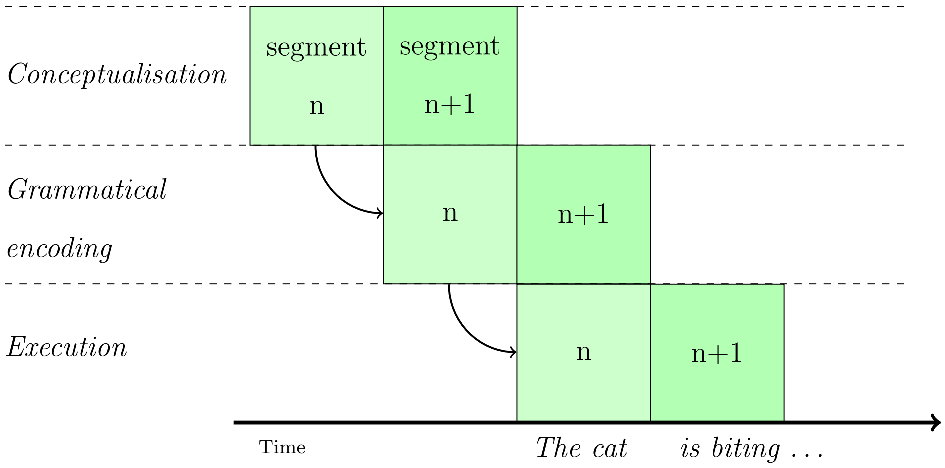


Figure 1.1: Hierarchical parallel model of language production with cascading information flow for the sentence *The cat is biting in its tail* adapted from Olive (2014, p. 178). The planning unit is referred to as "segment" and "n" indicates the sequential order; first unit is n, second unit is n+1 and so forth.

In this thesis, I am going to focus on the first increment. In particular, I will examine how this increment is generated during the highest two planning stages illustrated in Figure 1.1. There are varying conclusions about the size of this first increment ranging from the first noun of the sentence (Brown-Schmidt & Konopka, 2008; Griffin, 2001) or less (Bürki, Sadat, Dubarry, & Alario, 2016) to planning units beyond the utterance-initial elements (Allum & Wheeldon, 2007; E.-K. Lee, Brown-Schmidt, & Watson, 2013; Smith & Wheeldon, 1999, 2001) and across the entire clause (Bock & Cutting, 1992; Bock & Miller, 1991; Meyer, 1996, 1997).

All these studies agree that the part of the message that requires linguistic processing is dependent more on sentence position (i.e. the fact that it occurs at the start of the sentence) than on its functional role in the sentence. Also the scope of advance planning appears to vary. Depending on various environmental factors, speakers are likely to expand or contract advance planning. Planning of complex linguistic units is less likely if the speaker is under time pressure to initiate an utterance (F. Ferreira & Swets, 2002, 2005) and under increased processing demands (Martin, Yan, & Schnur, 2014; Wagner et al., 2010), and varies across individuals (Griffin & Spieler, 2006; Swets et al., 2014; Van de Velde & Meyer, 2014; Wagner et al., 2010).

If the scope of advance planning is forced to contract under some conditions and is permitted to expand in other situations, the planning system seems to anticipate the processing situation (context) before it decides how much advance processing is feasible or necessary (Griffin, 2003; Levelt & Meyer, 2000). As the extent of advance planning is subject to variation, it is difficult to tease apart whether a planning unit larger than or as large as the first word is obligated by the language production system or by non-linguistic factors. Therefore the influence of non-linguistic factors that are imposed onto the cognitive process that prepares language needs to be taken into consideration. This will be addressed using output modality comparisons, for reasons I will introduce in the following section.

1.3 Writing and language planning

Data from spoken sentence planning have shown that contextual factors influence the extent of linguistic preplanning (F. Ferreira & Swets, 2002; Wagner et al., 2010). Among those the choice of output modality might affect how sentences are prepared in advance and production in general. While there is some understanding of how sentence planning processes are coordinated in speech (e.g. Bock & Levelt, 1994; Bock & Ferreira, 2014; V. S. Ferreira & Slevc, 2007; Levelt, 1989), advance planning in writing has been largely neglected except for some preliminary studies (Damian & Stadthagen-Gonzalez, 2009; Nottbusch, 2010; Nottbusch, Weingarten, & Sahel, 2007; Torrance & Nottbusch, 2012). Research that aims to understand the mechanisms underlying language production has focused almost entirely on data from speech. Therefore it is impossible to determine whether existing data represent language-general or speech-specific processes. Alario et al. (2006) summarise this modality bias in an overview on the state-of-arts of language production research from the *International Workshop on Language Production* from 2004, arguably the most important summit in the field of language production research:

> "[P]sycholinguistic investigations should not overlook a valuable opportunity to gain fuller insights about language processing by exploiting these [i.e. writing and sign-language] distinct linguistic modes. [. . .] A full integration of speaking and writing models with the mechanisms of typical lan-

> guage production should be a high priority for the field [. . .]"
> (p. 783–784)

In more than a decade after this meeting little has changed. In the most recent edition of *the Oxford Handbook of Language Production* (Goldrick, Ferreira, & Miozzo, 2014), writing (and sign-language) is dedicated a chapter with topics distinct from other areas of language (i.e. speech) production research. Also, speech production has often been used as a synonym for language production (see e.g. Eysenck & Keane, 2015; Martin, Crowther, Knight, Tamborello II, & Yang, 2010). For instance, the chapter on language production in Eysenck and Keane's (2015) student's handbook on cognitive psychology the authors used the title "speech planning" (p. 455–456), presumably because all of the work they cited focused on speech. However, studying writing in the context of language planning should be elementary – to reiterate Alario et al. (2006) – to understand the mechanisms underlying language production more generally.

Writing research has largely focused on writing-specific rather than language-general domains. For example there is an educationally-focused literature on the production of extended texts, which draws on ideas about strategic processing first presented by Hayes (Flower & Hayes, 1980; Kaufer, Flower, & Hayes, 1986). Also there is research exploring working memory use in writing (e.g. Kellogg, 1988, 2001; McCutchen, 1996; Olive, Alves, & Castro, 2009; Olive & Kellogg, 2002). Reading during the production of (narrative) texts has been studied with mostly impaired speakers (e.g. Asker-Árnason, Wengelin, Sahlén, & Ibertsson, 2010; Behrns, Ahlsén, & Wengelin, 2010; Wengelin, 2002; Wengelin, Leijten, & Van Waes, 2010). There are some studies on motor planning in typing (e.g. Gentner, 1982; Gentner, Larochelle, & Grudin, 1988a; Rumelhart & Norman, 1982; Terzuolo & Viviani, 1980) and there is a relatively extensive literature exploring basic processes in the written production of single words (Bonin & Fayol, 2000; Bonin, Fayol, & Chalard, 2001; Bonin, Malardier, Meot, & Fayol, 2006; Bonin, Peereman, & Fayol, 2001; Damian & Stadthagen-Gonzalez, 2009; Delattre, Bonin, & Barry, 2006; Torrance et al., 2017; Will, Nottbusch, & Weingarten, 2006; Zesiger, Orliaguet, Boë, & Mounoud, 1995; Zhang & Damian, 2010) which broadly replicated findings from spoken naming. However, basic cognitive processes associated with written production above the word level have received very little attention in research.

Writing differs from speech in a number of ways. Writing places additional demands on the language processor. The grammatically encoded unit needs to be mapped onto symbols that allow the communication of a message. In speech we use symbols that are acoustic by nature (sounds). Writing, instead, uses graphical symbols or orthography (spelling) to convey information. These symbols might be representations of the sounds that are used in speech in alphabetic languages (e.g. Russian, Tagalog) or representations of entire syllables or words in logosyllabic languages (e.g. Mandarin Chinese, Japanese). Both hand-writing (Van Galen, 1991) and keyboard-typing (Gentner, 1982; Rumelhart & Norman, 1982) are cognitive skills that are substantially different from speech and cannot be understood as a simple extension of spoken language. Writing does not necessarily involve the activation of phonological representations: The generation of orthographic codes (Bourdin & Fayol, 1994; Van Galen,

1991) is typically acquired subsequent to the phonological inventory of the target language. The activation of orthographic representations can be characterised by a dual-route process (Barry, 1994; Damian, Dorjee, & Stadthagen-Gonzalez, 2011; Qu & Damian, 2017) in which phonology may or may not serve as a mediator. Some research suggests that orthographic representations activate phonology (e.g. Bonin & Fayol, 2000; Nottbusch, Grimm, Weingarten, & Will, 2005; Zhang & Damian, 2010) but also other researchers have found that orthographic representations can be activated via a lexical route without access to phonology (e.g. Bonin, Fayol, & Gombert, 1998; Rapp, Benzing, & Caramazza, 1997; Sahel, Nottbusch, Blanken, & Weingarten, 2005). Written words are planned prior to output (Damian & Stadthagen-Gonzalez, 2009; Shen, Damian, & Stadthagen-Gonzalez, 2013) while some lexical processing continues after word onset (Nottbusch et al., 2005, 2007; Nottbusch, Weingarten, & Will, 1998; Sahel et al., 2005; Sahel, Nottbusch, Grimm, & Weingarten, 2008; Weingarten, Nottbusch, & Will, 2004; Will et al., 2006), which argues against complete advance planning on a lexical level.

Further, planning sentences is considerably more complex than the production of simple words. The additional requirements that are involved in planning sentences might load differently onto the processor in writing than in speech. Although the mental processes underlying the generation of syntactic representations in speech and writing are assumed to overlap (e.g. Branigan, Pickering, & Cleland, 1999; Cleland & Pickering, 2006; Hartsuiker & Westenberg, 2000; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008), planning and executing the written production of a sentence involves different contextual conditions from speech. Consider, for instance, the following three examples and their implications for the language processor. First, writing typically takes more time than speaking. This implies that – all else being equal – information that was preplanned at higher levels of representation has to be mentally buffered over a longer period of time. In a cascading parallel planning architecture this will increase the overall production difficulty. Alternatively, the processor might either plan smaller units in advance and rely on parallel processing or planning might unfold serially to avoid buffering of large chunks of information. Less experienced writers tend to operate serially while more advanced writers use parallel processing strategies (Olive, 2014; Van Galen, 1991).

Second, persistent visual feedback from the unfolding sentence (even dialogues and entire texts) is usually available in writing. In speech, however, acoustic feedback is only immediately available. Linguistic memory traces are known to decay quickly if not rehearsed, and are difficult to access afterwards (Christiansen & Chater, 2016; Lewis & Vasishth, 2005). Thus, monitoring the output needs to happen immediately in speech (Levelt, 1989) but may be delayed or, in fact, not happening at all in writing (Olive, 2014).

Third, speech imposes fluency demands on the output. Speech requires a certain degree of fluency. This is because pauses, and particularly frequent and/or long pauses have communicative effects: They may change the interpretation of the intended sentence (Clark & Fox Tree, 2002). Interruption of the speech stream may occur when parallel planning is not possible, i.e. the speaker pauses to prepare the next speech unit. To maintain fluency the processor may prepare more information in advance of output onset. By contrast writing allows, in principle, pausing and editing during the production process, even within words. This affects only the process but not the product.

Taken together these speech/writing differences suggest that writing is a valuable context for the investigation of the first planning unit in sentence production. In writing there is less overlap with subsequent planning units, there is no need for immediate monitoring, and, possibly most importantly, advance planning needs only to address linguistic requirements whereas speech imposes additional fluency demands.

In summary, underlying both speech and writing there are cognitive processes that are associated with a modality-independent language production system. However, the focus on a single modality (speech) and the absence of data from writing in production research is problematic as it does not allow us to establish whether theoretical models are indeed language-general or speech-specific. Speech is generally known to be subject to non-linguistic and contextual factors, which affect the way language is planned (F. Ferreira & Swets, 2002; Wagner et al., 2010), and communicative factors, which may require more advance preparation than required by the linguistic processor.

Studying language production in writing makes it possible to remove speech-specific demands such as fluency. There is, to my knowledge no published psycholinguistic research that has looked systematically at advance planning in written sentence processing by directly comparing data from speech and writing. A central aim of the two series of experiments that I report in this thesis was to fill this gap, making modality comparisons to tease apart speech-specific and language-general planning demands. I will introduce these in the following section.

1.4 Present research

This thesis takes up two discussions about the mechanisms underlying the language processor. The first series of experiments – presented in Chapter 2 – concerns the relationship between two sub-processes of grammatical encoding: the generation of lexical items and syntactic structure (V. S. Ferreira & Slevc, 2007; Wheeldon, 2011; Wheeldon, Smith, & Apperly, 2011). This series tested whether grammatical encoding in advance planning is lexically or structurally guided. The second series – presented in Chapter 3 – explores the impact of the conceptual representation on advance planning. This series tested the hypothesis that conceptual relationships between message elements determines the extent of advance grammatical encoding (see Konopka & Brown-Schmidt, 2014; Meyer, 1997).

Each of these empirical chapters includes three experiments. Image description tasks were used similar to those used in studies reported in the planning scope literature (Brown-Schmidt & Tanenhaus, 2006; Konopka, 2012; E.-K. Lee et al., 2013; Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Swets et al., 2014; Wagner et al., 2010). Participants had to describe the actions/arrangements of simple images – in separate typing and speech conditions – which were presented on a computer screen. To assess the extent of information that participants planned before the production onset was released, properties of the target utterance – syntactic structure and ease of lexical retrieval (Chapter 2) or conceptual contrast

(Chapter 3) – were manipulated by changing features of the stimulus display. Response differences resulting from these manipulations were assessed by virtue of two widely used tools: First, the time from stimulus onset to production onset (i.e. the onset latency) was used to measure changes in the general planning difficulty associated with the manipulation (e.g. Levelt & Maasen, 1981; Smith & Wheeldon, 1999; Martin et al., 2014; Wagner et al., 2010). Second, eye movements on the stimulus arrays were recorded as those provide detailed information as to whether a particular referent was relevant to advance planning and when planning of this referent must occur (e.g. Griffin, 2004a, 2004b; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer & Lethaus, 2004; Nottbusch, 2010; Swets et al., 2014). These extracted dependent variables were analysed using Bayesian modelling techniques which addressed the complexity of the nested data structure and provided a direct estimate of the strength of the evidence for a specific hypothesis (Gelman & Hill, 2007; Gelman et al., 2014; McElreath, 2016).

In Chapter 2 I will address a controversy between two models of language production. These models compete with respect to whether or not the generation of content and structure is independent (see e.g. Bock & Ferreira, 2014; V. S. Ferreira & Slevc, 2007; Wheeldon, 2011; Wheeldon et al., 2013). On the one hand there are *syntax-based* models which posit that lexical content and syntactic structure are planned independently (Chang et al., 2003, 2006, 2000; Costa & Caramazza, 2002; Dell, 1986; Dell & O'Seaghdha, 1992; V. S. Ferreira & Dell, 2000; V. S. Ferreira & Slevc, 2007; Garrett, 1975, 1982, 1988). A simplified illustration of this model can be found in Figure 1.2. Syntactic-frames are build on conceptual representations without access to lexically specified representations. Such a syntactic frame is “a virtual cognitive instantiation of hierarchical structure” (Bock & Ferreira, 2014, p. 22) that links between (distant) parts of the utterance. Evidence for syntax-based models comes from studies demonstrating that the obligatory unit of advance planning respects the syntactic structure of the sentence-initial phrase (Konopka, 2012; Martin et al., 2010, 2014; Smith & Wheeldon, 1999, 2004; Wheeldon, 2012; Wheeldon et al., 2013; Wagner et al., 2010). By asking participants to describe moving arrays of three images (e.g. a dog, a hat, and an apple moving in different directions) the authors repeatedly found that it takes longer to begin the production of sentences that start in a coordinated noun phrase such as *The dog and the hat moved above the apple* compared to sentences that start in a simple subject phrase such as *The dog moved above the hat and the apple.* The conclusion from this finding was that the extent of advance planning is dependent on the size of the sentence-initial phrase, suggesting a structurally guided planning strategy. Syntactic representations are planned first and guide lexical retrieval. Thus lexical representations of non-phrase initial items (e.g. *the hat*) might be incomplete and retrieved after production onset as shown by Wheeldon et al. (2013).

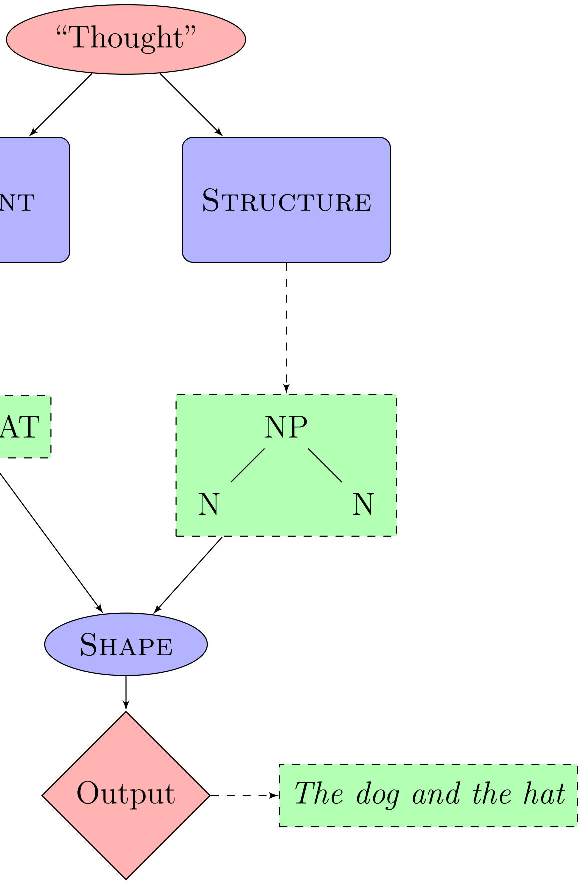


Figure 1.2: Syntax-based model of language production (see e.g. V. S. Ferreira & Slevc, 2007, p. 453). "Shape" of the intended output in the original model is the phonological and phonetic representation but in the present context it might only be an orthographic representation. Stages of grammatical encoding are shown as blue boxes and examples are given in green boxes. Illustrated is the composition of the phrase *The dog and the hat* as assumed by a syntax-based production model.

On the other hand there are *lexically-based* theories of language production which assume that syntactic properties are connected to representations stored as lexical units (Bock, 1982; Bock & Levelt, 1994; F. Ferreira, 2000; Griffin, 2001; Levelt, 1989, 2001). In these models syntactic properties and structure emerge following the activation of lexical representations. Figure 1.3 shows an illustration of such a lexically-based production model. These models assume that linguistic units are planned incrementally as minimal chunks and combined during the unfolding production of the sentence. Authors advocating a lexical perspective on language planning have argued that the effect introduced above – increased planning durations for coordinated noun phrase – is specific to noun phrases with two nominal heads which have an arbitrary structural order. Allum and Wheeldon (2007, 2009) showed that noun phrase with two lexical elements and an implicit structural hierarchy such as *The A with the B* show shorter onset latencies. Production onsets for coordinated noun phrases were longer as planning is required for two rather than just one functionally-determined noun. The authors favoured a non-syntactic interpretation for how the language processor determines which noun needs to be planned first, as this effect was independent of whether the first noun is the head noun as in English or a modifier as in Japanese. Allum and Wheeldon (2007) argued instead that planning embraces a minimal, thematically closed functional phrase; i.e. *the A* or *with the B* or *the A and the B*. Importantly this shows that advance planning does not necessarily embrace the entire sentence-initial phrase. An alternative possibility is that in both cases the entire phrase was preplanned but the absence of a structurally determined order required longer planning durations for coordinated noun phrases.

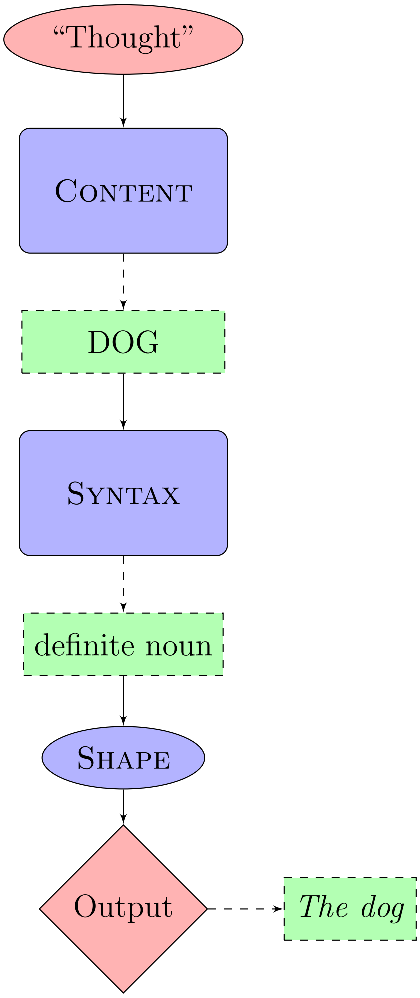


Figure 1.3: Lexically-based model of language production. This figure illustrates advance planning of the first noun *The dog* as in the phrase *The dog and the hat*. A second loop is required for the noun *the hat* which might be buffered before production onset. "Shape" refers to the phonological/phonetic or orthographic form of the output. Stages of grammatical encoding are coloured in blue and examples are given in green.

These authors (and Griffin, 2003) propose an alternative, lexicalist interpretation for the advance consideration of both nouns in a coordinated phrase. They argued that planning beyond the first noun was obligated by speech-specific communicative requirements: the second noun needs to be preplanned to maintain output fluency after production onset as there is not enough time available to retrieve the name of the second noun in parallel with the production of the preplanned unit. Thus advance planning for coordinated noun phrases can be explained by lexically-based theories without reference to the advance creation of syntactic frames. Instead coordinated noun phrases involve preplanning of both lexical items to avoid intra-sentential pausing which is imposed by modality-specific constraints (e.g. fluency) and is not obligated by the grammatical encoder as assumed by syntax-based theories. However there is, as I have argued before, no need to maintain output fluency in writing. Therefore lexically-based accounts would predict advance planning for coordinated noun phrases in speech but not in writing, as the latter does not need to take fluency requirements into account. Hence writing allows the reduction of the advance planning effort to the minimum extent required by the language processor.

Chapter 2 examines whether grammatical encoding beyond the phrase-initial noun is speech-specific. If grammatical encoding for coordinated noun phrases was found for speech but not for writing, we would conclude that this effect is modality-specific rather than language-general and thus the theory that syntactic frames guide advance linguistic processing can be ruled out. I found evidence that coordinated noun phrases are planned beyond the first noun in both speech and writing. This finding supports syntax-based models of language production and shows that the phrasal planning scope for coordinated noun phrases cannot simply be accounted for by fluency demands on speech.

More recent discussions on language production concluded that advance planning is not necessarily either hierarchically complex (e.g. Kuchinsky, Bock, & Irwin, 2011; Martin et al., 2010; E.-K. Lee et al., 2013; Smith & Wheeldon, 1999; Wheeldon et al., 2013) or lexically incremental (e.g. Allum & Wheeldon, 2007, 2009; Gleitman, January, Nappa, & Trueswell, 2007; Griffin, 2001) but there is strong evidence that planning can be both (Bock & Ferreira, 2014; Konopka & Meyer, 2014; Kuchinsky, 2009; Kuchinsky et al., 2011). In other words language production might be guided by a lexical incremental strategy while under other conditions the language processor might operate using syntactically complex units. Aspects influencing the choice of either route are known to be linked to extra-linguistic parameters (e.g. F. Ferreira & Swets, 2002; Martin et al., 2014; Swets et al., 2014; Wagner et al., 2010) and the pre-activation of linguistic information (e.g. Konopka, 2012; Konopka & Meyer, 2014; Swets, Jacovina, & Gerrig, 2013; Van de Velde, Meyer, & Konopka, 2014).

The results of the first series of experiments and those reported in the literature might be traced back to a conceptual processing stage that decides whether advance grammatical encoding merely requires the first lexical element (Konopka & Meyer, 2014). Although I found advance planning for coordinated noun phrase regardless of output modality, it is known from the literature that advance planning does not systematically embrace the sentence-initial phrase. Specifically advance planning of subordinated noun phrases tends to embrace the first noun only (Allum & Wheeldon, 2007, 2009). Allum and Wheeldon (2007) suggested that the lack of a syntactically determined hierarchy in coordinated noun phrases – as opposed to subordinated noun phrases – prohibits an incremental planning strategy. In order to facilitate advance planning, the grammatical encoder requires information on whether or not it is permitted to plan incrementally which is bound to be determined pre-syntactically.

This idea has already been pointed out by Meyer (1997) and by early research that had strong impact on modern psycholinguistics (Garrett, 1975; Lashley, 1951; Wundt, 1900). The idea is that the linearisation of language has to begin earlier in the planning process – at a high level of organisation. Konopka and Meyer (2014) suggested that this begins with the conceptualisation of the "thought" which Bock and Ferreira (2014) called "the heart of communication" (p. 22). The conceptualisation involves two possible routes – as illustrated in Figure 1.4. The thought needs to pass through either of these routes before it feeds into the grammatical encoder. One route is non-relational and concerns the encoding of characters or objects. The other route is relational and represents the proposition, the "comment" and any other type of dependency between message elements, e.g. the coordinating conjunction *and* in *the dog and the hat*. This conceptual relation is the foundation of syntactic structures. The non-relational and relational conceptual processing stage cascade into the lexical and structural planning stages of the grammatical encoder, respectively. It is therefore the conceptual stage that determines whether a lexical or syntactic planning strategy is used. There has been some discussion on whether language planning is driven by non-relational (Gleitman et al., 2007) or relational conceptual representations (Kuchinsky, 2009). However, as both planning strategies seem to be available to the language processor, the question is not which of these strategies guides advance planning but under which circumstances either strategy is used (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014; Konopka & Meyer, 2014; Kuchinsky, 2009). Therefore the route of advance planning required for complex noun phrases is not necessarily decided by the grammatical encoder but at a pre-syntactic, conceptual planning stage.

In Chapter 3 of this thesis I explore whether the extent of advance linguistic planning has been determined at a conceptual planning stage. The minimal unit of advance grammatical encoding might be determined by the conceptual relation between the phrase elements as discussed above. This would imply that the planning scope is not determined by the grammatical encoder as suggested by most authors (E.-K. Lee et al., 2013; Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Wagner et al., 2010; but see Allum & Wheeldon, 2007, 2009; Meyer, 1997) but at a pre-linguistic conceptual stage. For example in complex noun phrases advance planning might be induced by the conjunction (*and*) in coordinated phrases, while in subordinated noun phrases planning might either take the non-relational route (e.g. Allum & Wheeldon, 2007, 2009) or the relational route (E.-K. Lee et al., 2013; Nottbusch, 2010; Solomon & Pearlmutter, 2004). Preplanning conceptual properties of complex noun phrases has been explored with respect to semantic integration (Solomon & Pearlmutter, 2004), thematic status of phrase parts (Allum & Wheeldon, 2007, 2009; Zhao & Yang, 2013) and the referential uniqueness of the phrase (Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; Swets et al., 2014) but there is no evidence that the extent of advance planning in complex noun phrases is determined by the type of conceptual relation between message elements. In the experiments

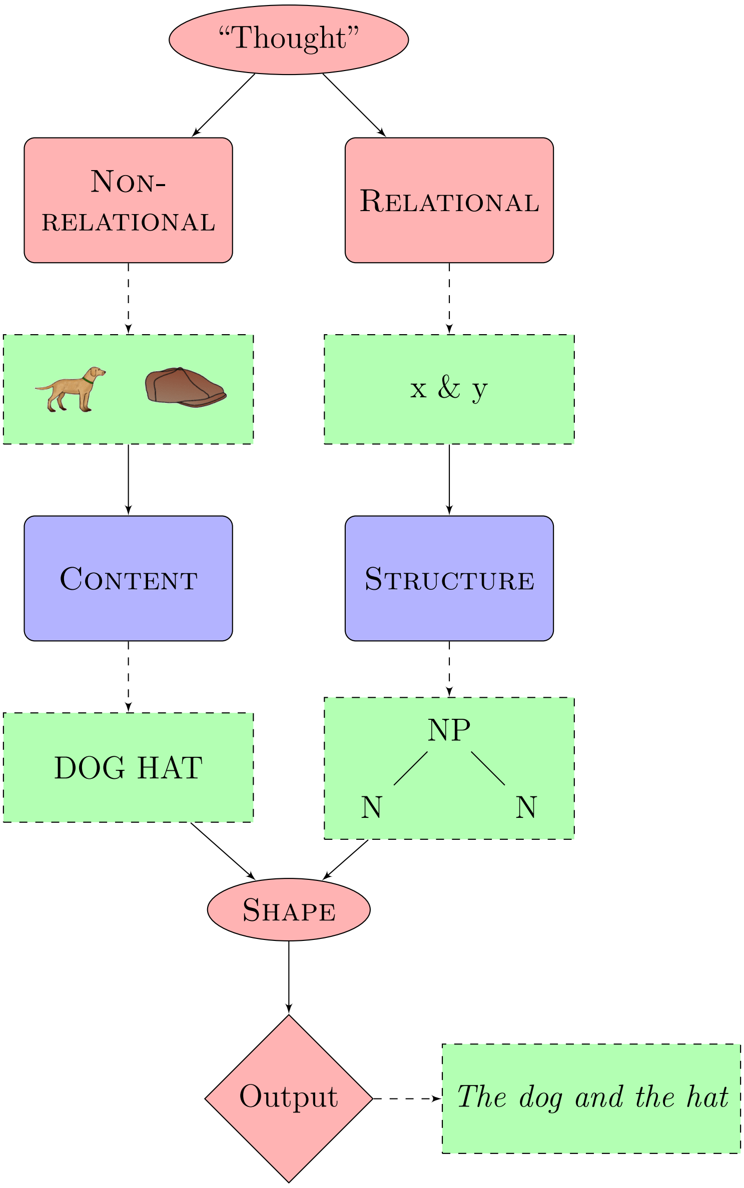


Figure 1.4: Model of conceptual representation prior and grammatical encoding (see Konopka & Brown-Schmidt, 2014) for the phrase *The dog and the hat*. This model is similar to the syntax-based model of language production shown in Figure 1.2 but a conceptual planning stage has been added. This stage is indicated by red boxes.. Grammatical encoding is shown in blue boxes. Examples are given in green boxes. "Shape" refers to the phonological/phonetic or orthographic form of the output.

reported in Chapter 3, I tested whether preplanning of complex – subordinated – noun phrases is determined by conceptual relations between nominal phrase elements. To be able to distinguish between planning requirements that are imposed by the relational part of the conceptual planning system from speech-specific factors, I took modality differences into account. I present evidence that conceptual relations determine whether advance planning has to scope beyond the phrase-initial noun. This effect replicated in both speech and writing showing that its source is language-general rather than speech-specific.

The present thesis is organised in the following way. The research questions that were introduced in this section will be described in detail in the respective chapters along with the empirical methods that were used to test each question. Chapter 2 tested whether grammatical encoding beyond the sentence-initial phrase requires lexical specification. Although there was evidence for planning beyond the first noun in both speech and writing, there was no indication that this obligates lexical retrieval. In Chapter 3, I tested whether the need for planning beyond the first noun is determined by the conceptual planning system. Support for this hypothesis was found across output modality. Chapter 4 will discuss the results of both series of experiments in light of modality-specificity of language production theory, some caveats, and a methodological implication. Finally this chapter will conclude this thesis by proposing that advance planning embraces coordinated noun phrases in sentence-initial position but does not require lexical specification beyond the first noun, while the extent of preplanning beyond the first noun is determined by a conceptual message-level representation. Reproducing these effects in both writing and speech demonstrates that these conclusions are language general rather than modality-specific. Future directions are indicated.

Chapter 2

Syntactic frames guide grammatical encoding

Abstract

Response onset latencies for spoken sentences that start with a conjoined noun phrase are typically longer than for sentences starting with a simple noun phrase. This is consistent with advance retrieval of syntactic frames independently of lexical retrieval. Alternatively, planning may be lexically driven and planning beyond the initial noun requires just to maintain spoken output fluency. Writing relaxes this output-fluency constraint. In three image-description experiments (Ns = 32) subjects produced sentences with simple and conjoined initial noun phrases in both speech and writing. Production onset latencies and participants' eye movements were recorded. Ease of lexical retrieval of sentences' second nouns was assessed by manipulating codability (Experiment 1) and by lexical 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.

2.1 Introduction

People can prepare entire sentences carefully in advance before they address an interlocutor or audience. However, typically sentences are planned in smaller units and planning of the sentence is incomplete at speech onset. Some advance planning is obligated prior to speech onset, 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. The present chapter 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; Konopka, 2012; Konopka & Bock, 2009; E.-K. Lee et al., 2013; Levelt & Maasen, 1981; Martin et al., 2010, 2014; Meyer, 1996; Smith & Wheeldon, 1999, 2001, 2004; Swets et al., 2014; Wagner et al., 2010; Wheeldon et al., 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 et al., 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). The research reported in this chapter addressed two issues that may account for this range of conclusions about the scope of obligatory planning. First, the scope of syntactic planning may or may not coincide exactly with lexical planning scope. Syntax may rely upon or emerge from lexical retrieval, and so share scope, or it may occur independently and in advance of lexical retrieval and therefore potentially have more extensive scope. Second, the advance plan may under certain conditions extend beyond the scope obligated by linguistic processing (Bock & Ferreira, 2014). Planning more than the obligatory planning unit might serve fluency requirements for spoken output. Hence, the observed planning scope may result in part from the pragmatic demands of specific experimental contexts rather than being a fundamental property of the language production system.

There is evidence that the extent of pre-sentence planning depends on the level of the structure that is being planned. Meyer (1996) found evidence that conceptual representations are planned for the entire clause. 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.

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

They found longer sentence onset duration for sentences with complex NPs. Similar effects were reported by Wheeldon et al. (2013) and Martin et al. (2010, 2014). 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) present evidence from event-related potentials showing semantic blocking effects for sentence-initial nouns only.

There appears, therefore, to be a planning-scope hierarchy, with conceptual planning scoping over the full clause, syntax planning extending over the sentence-initial verb-argument phrase, and lexical planning proceeding incrementally, on a word-by-word basis. This hierarchy is consistent with several *syntax-based* theories of language production (e.g. Bock, 1990; Bock, Irwin, Davidson, & Levelt, 2003; Bock & Ferreira, 2014; Chang et al., 2000, 2003, 2006; Costa & Caramazza, 2002; Dell, 1986; Dell & O'Seaghdha, 1992; V. S. Ferreira & Dell, 2000; Garrett, 1975). These theories assume that syntactic structures derive directly from conceptual representations but that lexical access is post-syntactic. Syntactic and lexical representations therefore have a degree of independence, with lexical planning filling a syntactic frame. In contrast, *lexically-based* theories (e.g. Bock, 1982; Bock & Levelt, 1994; F. Ferreira, 2000; Levelt, 1989, 2001) assume that syntactic structure is 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.

Several studies provide fairly direct evidence that syntactic structure does not rely on lexical specification (Allum & Wheeldon, 2007, 2009; Wheeldon, 2011, 2012; Wheeldon et al., 2011, 2013). 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).

While it has generally been assumed that the phrasal scope effect reflects advance grammatical planning that is independent of lexical processing and thus supports syntax-based accounts of language production, 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.

Alternatively sentence planning may be strictly lexical and incremental (Griffin, 2001, 2003, 2004b; Zhao & Yang, 2013, 2016). The additional planning effort for conjoined noun phrases reported by Smith and Wheeldon (1999) and Wheeldon et al. (2013) might not indicate the creation of syntactic frames, as the authors argued, but might result from non-linguistic factors acting to expand planning scope. There is a distinction between the minimum planning unit obligated by the language production system and the extent to which utterances are advance-planned in specific experimental (and non-experimental) contexts. It is, of course, possible to mentally prepare one or more clauses prior to output. The minimum planning unit obligated by the language production system may be a single word, or a determiner-noun pair, but in particular contexts speakers may plan further ahead. Therefore, planning beyond the first content word is also subject to context-specific non-linguistic factors. For instance, Griffin (2003) suggested that advance lexical retrieval serves fluency demands on the output. She provided data showing that speakers assess in advance the possibility of simultaneously producing the first and planning the second noun. In a naming task she found longer latencies when the first name in a pair of images was short. When there is insufficient time to prepare the second noun after production onset, the first name needs to be buffered and the second name requires advance planning to avoid within-sentence hesitation. Allum and Wheeldon (2007, p. 795) pointed out that conjoined noun phrases necessitate advance planning as there might not be enough time between the first and the second noun for parallel processing. Hence, the phrasal scope effect (e.g. Martin et al., 2010; Smith & Wheeldon, 1999; Wheeldon et al., 2013) may be the consequence of fluency demands on spoken utterances. If so, one would not expect phrasal scope effects in the absence of these fluency demands.

Spoken communication requires a high degree of output fluency because hesitation in speech has potential implications for listeners' understanding and interpretation of the message: Pauses in speech have a communicative effect (Clark & Fox Tree, 2002). In addition to specific language system demands, speakers' advance planning is affected by the need to minimise intra-sentence pausing once speaking commences (Levelt & Meyer, 2000; Meyer, 1997). By contrast, hesitation in written production, in most contexts, has no bearing on the text's eventual communicative effect. Writers may pause to plan what to write next at any point without any communicative effect. Arguably, therefore, in written production, but not in speech, advance planning can be reduced to the minimum unit required by the language production system to initiate the production of the sentence.

Despite the ubiquity of writing, planning mechanisms in written sentence production have been almost entirely neglected by researchers. Alario et al. (2006, p. 783–784) highlighted this modality bias and stressed the relevance of studies on writing (and sign language) to create a complete production model. There is some evidence from cross-modal syntactic priming showing that writing and speech employ the same syntactic processing system (Branigan et al., 1999; Cleland & Pickering, 2006; Hartsuiker & Westenberg, 2000; Hartsuiker et al., 2008). Research exploring planning scope in written sentence production is, to my knowledge, limited to three papers reporting preliminary findings (Nottbusch, 2010; Nottbusch et al., 2007; Torrance & Nottbusch, 2012). Nottbusch et al. (2007) and Nottbusch (2010) found evidence for increased sentence-initial planning time associated with producing noun phrase with a prepositional phrase modifier, compared to 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 of modality is, however, 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 that planning scope rarely extended beyond the first noun of the sentence. This effect was also present, but stronger, when sentences were written.

Studying written sentence production is therefore valuable in the present context because (a) this provides a more direct indication of obligatory planning scope without the possible confound imposed by the need for fluent (hesitation-free) output, and (b) incorporating written production in present language production models is, in itself, a worthwhile goal.

The present chapter reports three experiments in which participants generated short spoken and written sentences in response to image arrays. This research aimed (a) to confirm phrasal scope for advance planning in spoken production of simple sentences, (b) to determine whether this finding extends to the written modality, and (c) to determine the extent to which advance syntactic planning is dependent upon lexical retrieval. The comparison of spoken and written modalities is central to achieving this last aim. Summarising my argument: Syntax-based models imply that the minimally-obligated sentence-initial syntactic planning scope, which appears to be phrasal, is lexically independent and therefore can extend beyond minimally obligated lexical planning scope (e.g. Smith & Wheeldon, 1999; Wheeldon et al., 2013). Therefore non sentence-initial nouns within this scope do not need to be lexically specified in advance of output onset. Conversely, lexical accounts argue that the syntactic planning unit is based upon lexical retrieval. All previous studies have been in the spoken modality. It is possible that in speech planning scopes beyond the linguistic unit obligated by the language system. Phrasal scope might therefore result from the need to maintain fluency and may therefore not provide evidence against lexical accounts (Allum & Wheeldon, 2007; Griffin, 2003). Arguably, advance planning in a written context does not impose fluency constraints on the output. Examining sentence production in both speech and writing therefore controls for fluency demands providing a more direct test of whether the phrasal scope effect is independent of lexical processing.

2.2 Experiment 1

Experiment 1 aimed to confirm phrase-level scope of advance planning of simple sentences. It has been claimed previously that this scope results from a language production system constraint that requires the preplanning of the whole of the initial subject noun phrase. 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 that started 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. 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 only affect planning latencies if it is included in the subject noun phrase (i.e. just in the Complex NP condition). Syntax-based accounts hold that lexical preparation for non sentence-initial nouns is not obligatory. It is possible, however, that the pressure to be fluent in speech production results in processing of the second noun even when this is not required by the language-production system. This fluency explanation can be dismissed if codability effects in the Complex condition are present in both modalities. Therefore finding increased planning onset latencies for Complex NPs in the spoken modality provides robust evidence that initial planning has phrasal scope. Finding the same effect for written production is evidence against this effect resulting from the need for output fluency.

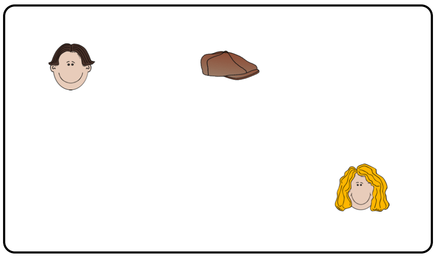
2.2.1 Method

2.2.1.1 Participants

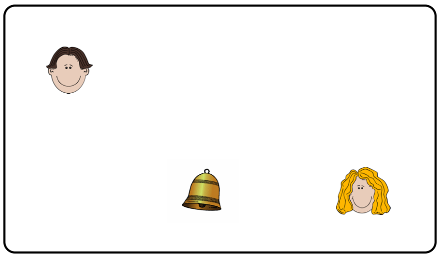
32 psychology students (26 female, mean age = 19.1 years, *SD* = 1.4, range: 18–25) participated as part of a research-reward scheme. All participants were self-reported as native speakers of British English, as free of linguistic impairments, and having normal or corrected-to-normal vision.

2.2.1.2 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 2.1). Images reached the target position after 100 ms and then stopped moving. Participants were asked to produce sentences of the form shown in Figure 2.1, with the order of the nouns in the sentence preserving the left-to-right order of images on the screen. In the array shown in Figure 2.1a, the leftmost image and the image in the centre of the screen moved up while the rightmost image moved down. In the other array shown in Figure 2.1b 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 2.1a is a conjoined noun phrase (*Peter and the hat*) and is, therefore, more complex than the subject phrase of the target sentence for Figure 2.1b 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 phrase.



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



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

Figure 2.1: Example stimulus screens. *N2* refers to the image in the centre.

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 2.1a), 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 2.1b).

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. Although all three images were areas of interest (henceforth, AOI) for the eye movement data, the critical variables were the proportion of eye samples to the image corresponding to N2 between stimulus and response onset, the time – relative to production onset – at which N2 was fixated subsequent to N1, and the proportion of trials for which this gaze shift occurred before production onset. Eye samples within the first 100 ms of each trial (the duration of image movement) were ignored.

2.2.1.3 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 years, $SD = 6.5$) 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.1.

The resulting 96 images were combined with images of *Peter* and *Tania*, the boy and the girl in Figure 2.1. First names were used as they 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. Forty-four filler trials were added to elicit 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).

2.2.1.4 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 phrase, examples of image arrays and the associated target sentences were intermingled with examples of fillers. Participants were also taught 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.

2.2.1.5 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 rate 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.

2.2.2 Results

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 were excluded from the analysis (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%) were and trials with durations longer than 40,000 ms (0.4%) were removed. For the analysis of eye data 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, 2015b). An adapted version of the code presented in D. Bates, Kliegl, Vasishth, and Baayen (2015) was used for analyses. All models were fitted with maximal random effects structures (Barr, Levy, Scheepers, & Tily, 2013; D. Bates et al., 2015). To assess effect size each predictor was sum coded (±1). The model was fitted with predictors for main effects of NP complexity, N2 codability, modality and their interactions. Posterior 95% credible intervals (henceforth, CrI) were calculated from the posterior samples. 95% CrIs that do not contain zero are evidence for an effect of the predictor variable. Also the proportion of posterior samples smaller than zero was reported (henceforth $P(\beta < 0)$). This proportion indicates the probability that the effect is negative, given the observed data. A proportion of posterior samples approaching zero is therefore strong support for a positive effect. In contrast, a proportion of posterior samples approaching one, is support for a negative effect. Inconclusive evidence includes large numbers of posterior samples of either polarity (see Kruschke, Aguinis, & Joo, 2012;

Nicenboim & Vasishth, 2016; Sorensen, Hohenstein, & Vasishth, 2016). The strength of support for a particular effect (i.e. the sensitivity of the data) was expressed in Bayes Factors calculated using the Savage-Dickey method (Dickey, Lientz, et al., 1970) (henceforth, BF_{10} signifying the evidence for the alternative hypothesis over the null hypothesis). A BF_{10} of 2, for example, means that the data are two times more likely under the alternative hypothesis than under the null hypothesis. A BF_{10} larger than 10 is considered strong support while extremely small BFs_{10} suggest evidence against the alternative hypothesis. Substantial evidence is indicated by a BF of, at least, 3–5 (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).[1] The proportion of eye samples was converted into empirical logits for analysis (see Agresti, 2002; Barr, 2008; Mirman, Dixon, & Magnuson, 2008). The conversion of the data results in multinomial distributions for smaller and larger proportions. Therefore a mixed effects model was fitted with four mixture components, i.e. combinations of four normal distributions, by varying the location of the intercept and the associated variance parameter. Model convergence was confirmed by visual inspection of traceplots of the Markov chain Monte Carlo chains and

---

[1] For onset latency, models were run with four chains with 2,000 iterations per chain, 1,000 iterations warm-up and no thinning. The models for proportion of eye samples were fitted with 6,000 iterations for each of four chains with 3,000 iterations warm-up and no thinning. The increased number of iterations was required to account for the multi-modal distribution of the eye data.

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/s/d91f0d35646b1dbf41d7.

2.2.2.1 Onset latency

The onset latency data are summarized in Table 2.1. For a visualisation of the distribution of the data see Appendix A.2.

Table 2.1: 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

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

Latencies were positively skewed and were therefore square-root transformed prior to analysis. The results of the Bayesian linear mixed model are presented in Table 2.2. The model gave strong evidence for longer onset latencies for Complex NPs compared to Simple NPs ($\text{BF}_{10} = 241$). Longer onset latencies were found for low codable N2 images compared to high codable images ($\text{BF}_{10} = 5$). The interaction of NP complexity and N2 codability was supported by the distribution of the posterior samples as shown in Table 2.2 but this effect was not substantial

(BF_{10} = 1.8). The interaction of N2 codability by NP complexity was inspected in nested comparisons within NP complexity. This was calculated from posterior samples contrasting low and high N2 codability. Strong evidence (BF_{10} = 125) 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], P($\beta < 0$) < .001) 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], P($\beta < 0$) = .259, $BF_{10} < 1$). The evidence for any other model predictor was negligible (all $BF_{10} < 1$).

Table 2.2: 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)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	0.59	0.31	0.86	< .001
N2 codability	0.41	0.11	0.72	.005
Modality	0.25	-0.60	1.09	.272
NP complexity:N2 codability	0.29	0.04	0.55	.011
NP complexity:Modality	0.00	-0.25	0.26	.487
N2 codability:Modality	-0.04	-0.29	0.21	.63
NP compl:N2 coda:Modality	0.07	-0.18	0.32	.281

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that effect β is smaller than 0, colon *":"* = interaction

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 complexity and N2 codability effects in writing, simple by-modality effects were tested. Similar NP complexity effects were found across modality with strong evidence ($\text{BF}_{10} = 67$) supporting longer latencies in Complex NPs in writing ($\hat{\mu} = 2.36$, 95% CrI[0.81, 3.86], $P(\beta < 0) < .001$) and speech ($\hat{\mu} = 2.32$, 95% CrI[0.83, 3.76], $P(\beta < 0) < .001$, $\text{BF}_{10} = 80$). Further the evidence for N2 codability effects showing longer latencies for low codable images for Complex NPs was found across modality. 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], $P(\beta < 0) = .004$, $\text{BF}_{10} = 19$) and speech ($\hat{\mu} = 1.35$, 95% CrI[0.29, 2.34], $P(\beta < 0) = .006$, $\text{BF}_{10} = 13$) but negligible evidence for N2 codability effects in Simple NPs, again, in both writing ($\hat{\mu} = 0.02$, 95% CrI[-1.08, 1.12], $P(\beta < 0) = .488$, $\text{BF}_{10} < 1$) and speech ($\hat{\mu} = .47$, 95% CrI[-0.55, 1.47], $P(\beta < 0) = .18$, $\text{BF}_{10} < 1$).

2.2.2.2 Eye movements

The first measure extracted from the eye movement data is the proportion of eye samples on the image representing the first noun (N1) and on the image representing the second noun (N2) as summarized in Table 2.3. These proportions span over the time before production (speech/writing) onset.

Table 2.3: Descriptive data of proportions of eye samples to the referent of the first noun (N1) and second noun (N2) of the target sentence prior to production onset. N2 is the critical noun bearing codability manipulation (Experiment 1)

			N1			N2			
Modality	NP complexity	N2 codability	*M*	*Mdn*	*IQR*	*M*	*Mdn*	*IQR*	*N*
speech	Complex	high	.73	.76	.37	.27	.24	.37	329
		low	.71	.75	.35	.28	.24	.34	319
	Simple	high	.89	1	.15	.10	.00	.15	322
		low	.87	1	.19	.12	.00	.18	315
writing	Complex	high	.77	.98	.45	.22	.00	.44	239
		low	.70	.79	.56	.28	.18	.53	249
	Simple	high	.85	1	.29	.13	.00	.25	236
		low	.81	.99	.37	.18	.00	.35	215

Note: M = mean, *Mdn* = median, *IQR* = interquartile range, N = number of observations

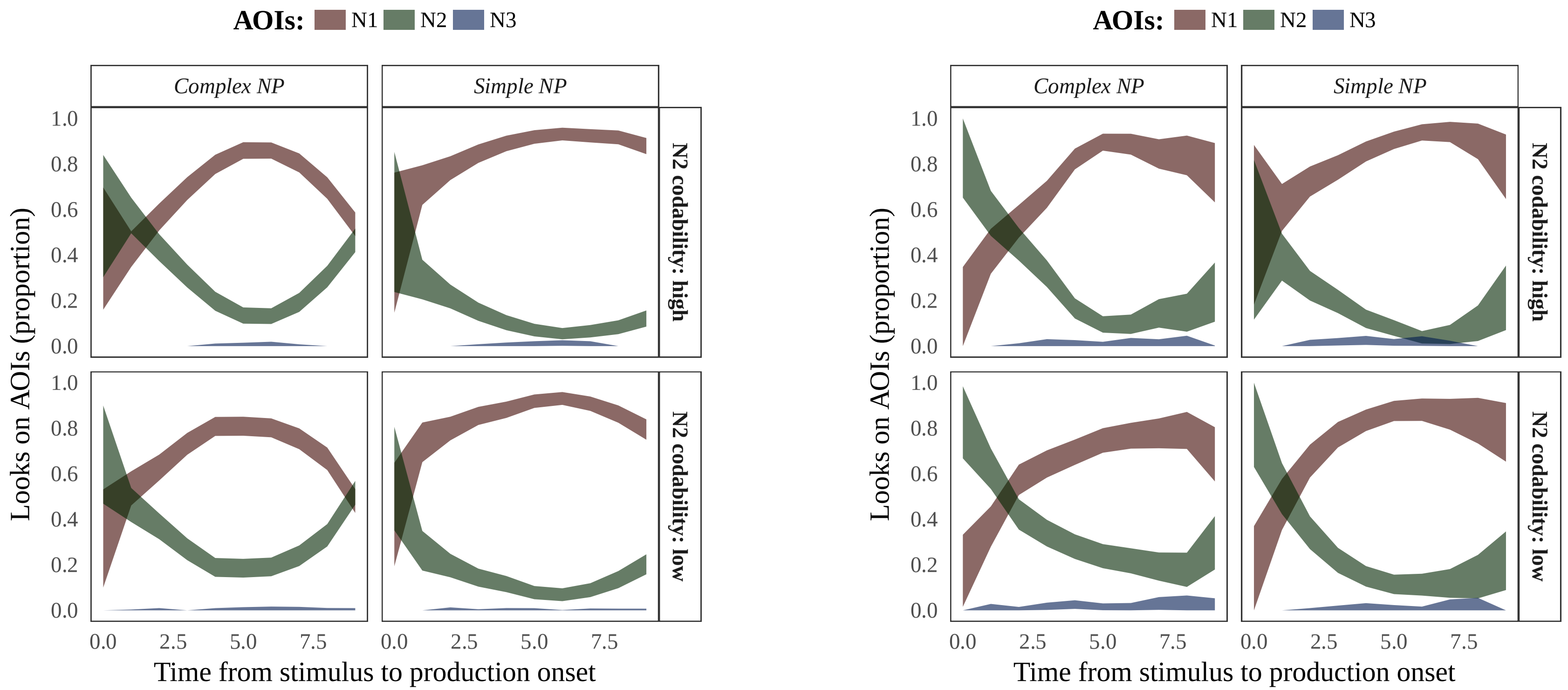
The results of the Bayesian linear mixed model are shown in Table 2.4. The analysis gave weak evidence ($\mathrm{BF}_{10} = 2.9$) supporting larger proportions of eye samples on N2 in Complex NPs compared to Simple NPs. The NP complexity effect was assessed within modality by testing simple main effects calculated from the model's posterior samples. There was moderate evidence ($\mathrm{BF}_{10} = 4$) for the NP complexity effect in speech ($\hat{\mu} = 0.31$, 95% CrI[0.08, 0.54], $P(\beta < 0) = .004$) but negligible evidence in writing ($\hat{\mu} = 0.06$, 95% CrI[-0.15, 0.28], $P(\beta < 0) = .312$, $\mathrm{BF}_{10} < 1$). Moreover the proportion of eye samples on N2 was larger in spoken trials than in written trials ($\mathrm{BF}_{10} > 1e8$).

Table 2.4: Main effects of first NP complexity, codability of N2, modality and their interactions inferred from the Bayesian linear mixed model on proportion of eye samples (in empirical logits) on N2 prior to production onset (Experiment 1)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	0.36	0.06	0.67	.008
N2 codability	0.07	-0.16	0.31	.270
Modality	2.24	1.75	2.74	$<$ .001
NP complexity:N2 codability	0.07	-0.15	0.29	.257
NP complexity:Modality	-0.25	-0.58	0.08	.936
N2 codability:Modality	0.10	-0.11	0.32	.164
NP compl:N2 coda:Modality	0.12	-0.16	0.42	.194

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper*=2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that effect β is smaller than 0, colon *":"* = interaction

The proportion of eye samples to each AOI is illustrated across time to production onset in Figure 2.2. These graphs illustrate that the proportion of looks to N2 increase while the looks to N1 decrease. The point at which the proportion of looks to N2 exceeds the proportion of looks to N1 – somewhere around production onset – indicates processing shift from the first noun to the preparation of the second noun (see e.g. Griffin, 2004b; Griffin & Bock, 2000; Konopka & Meyer, 2014; Meyer & Lethaus, 2004). A time course analysis – as often found in the literature for visual-world data in particular (e.g. Barr, 2008; Jaeger, 2008; Mirman et al., 2008) – would require approximately equal time windows over trials. The present eye data's time window is the stimulus-to-onset duration which varies by trial. A normalisation of the time axis would be a possible solution – as used for the visualisation in Figure 2.2. Such a normalisation, however, does not render meaningful time windows that would provide information about if and when advance planning was dedicated to N2. Instead two dependent variables were calculated from the eye data, namely (1) the time relative to production onset for when the gaze shift from N1 to N2 occurred and (2) the proportion of trials for which this gaze shift happened before production onset. Gaze shift from N1 to N2 was defined as the first, at least 100 ms long, fixation on N2 after the gaze left N1.



(a) Speech

(b) Writing

Figure 2.2: Proportion of eye samples to AOIs over time from stimulus onset to production onset illustrated by condition. AOIs are the referents of the three nouns (i.e. N1, N2, N3) as mentioned in the target sentence. The time axis was normalised within trial and binned. Bands indicate 95% confidence intervals (Experiment 1).

Gaze shift from N1 to N2 was detected in 87% of the data either before or after production onset. The time of gaze divergence relative to the production onset is shown in Figure 2.3.

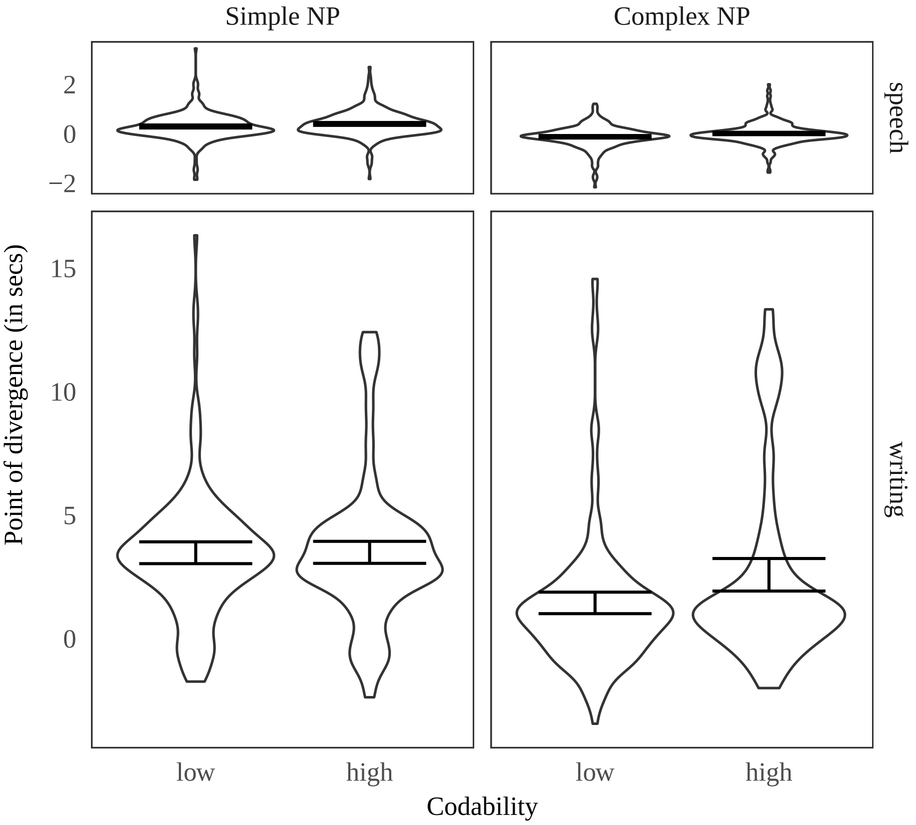


Figure 2.3: Bean plots of the point of gaze divergence from N1 to N2 relative to production onset. The beans illustrate the smoothed density of the distributions. The bands (boxes) illustrate the concentration of the data in 95% CIs. Null signifies the production onset (Experiment 1).

The time of divergence data were log transformed to account for positive skew. The data ware analysed in a Bayesian linear mixed model with main effects and interactions of NP complexity, N2 codability and modality. The results are summarized in Table 2.5. The model revealed substantial support for a main effect of NP complexity ($\text{BF}_{10} = 50261$)

supporting earlier gaze divergence in complex NPs. Further, there was strong support ($\text{BF}_{10} > 1\text{e}12$) for a main effect of modality showing earlier gaze divergence in speech compared to writing. The support for the modality by NP complexity interaction was moderate ($\text{BF}_{10} = 5$). This interaction was inspected in pairwise comparisons calculated from the model's posterior samples. These comparisons revealed strong support for modality differences in both Complex NPs ($\text{BF}_{10} > 1\text{e}5$) and Simple NPs ($\text{BF}_{10} > 6\text{e}10$) and strong support for NP complexity effects in speech ($\hat{\mu} = \text{-}0.16$, 95% CrI[-0.22, -0.1], $\text{P}(\beta < 0) > .999$, $\text{BF}_{10} = 210$) and writing ($\hat{\mu} = \text{-}0.47$, 95% CrI[-0.62, -0.32], $\text{P}(\beta < 0) > .999$, $\text{BF}_{10} = 14253$) with a larger effect magnitude in the latter. Both effects indicate earlier gaze divergence from N1 to N2 in Complex NPs. The effects of the other predictors were not supported by the model (all $\text{BF}_{10} < 1$).

Table 2.5: Main effects of NP complexity, N2 codability, modality and their interactions inferred by Bayesian linear mixed model on point of divergence data – time relative to production onset (Experiment 1)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	-0.08	-0.10	-0.06	$> .999$
N2 codability	-0.02	-0.05	-0.00	.978
Modality	0.18	0.15	0.21	$< .001$
NP complexity:N2 codability	-0.02	-0.04	-0.00	.994
NP complexity:Modality	-0.04	-0.06	-0.02	$> .999$
N2 codability:Modality	-0.01	-0.03	0.01	.853
NP compl:N2 coda:Modality	-0.02	-0.04	-0.00	.989

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability of effect β being smaller than 0, colon *":"* = interaction

The proportion of trials in which gaze divergence occurred before production onset is summarized by condition in Table 2.6.

Table 2.6: Descriptive summary for the proportion 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

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

A Bayesian linear mixed model with a Bernoulli distribution as sampling statement was fitted on the proportion of pre-onset gaze shift with main effects and interactions of NP complexity, N2 codability and modality as predictors. The model outcome is summarized in Table 2.7. The model revealed strong support ($\text{BF}_{10} > 2\text{e}5$) for a main effect of NP complexity showing a larger proportion of gaze shift before production onset in Complex NPs. Strong evidence ($\text{BF}_{10} = 1029$) was found for a main effect of modality indicating a larger proportion of pre-onset gaze shift to N2 for speech. There was strong support ($\text{BF}_{10} = 19$) for an interaction of these two main effects. This interaction was inspected in pairwise comparisons calculated from the posterior samples of the model. NP complexity effects were found in both speech ($\text{BF}_{10} > 2\text{e}7$) and writing ($\text{BF}_{10} = 312$). Within NP complexity comparisons showed strong evidence ($\mathrm{BF}_{10} > 4\mathrm{e}6$) for a larger probability of pre-onset gaze shift in Complex NPs for speech compared to writing ($\hat{\mu}$ = 4.03, 95% CrI[2.8, 5.34], P(β < 0) < .001). The same effect was seen in Simple NPs ($\hat{\mu}$ = 1.48, 95% CrI[0.18, 2.83], P(β < 0) = .012) which was found weaker by comparison but yet substantial (BF_{10} = 9). The main effect of N2 codability was found non-substantial (BF_{10} = 1.5). The evidence for all other model predictors was negligible (all BF_{10} < 1).

Table 2.7: Main effects of NP complexity, N2 codability, modality and their interactions inferred by Bayesian linear mixed model on the proportion of trials with gaze divergence from N1 to N2 before production onset (Experiment 1)

	Mean	*Lower*	*Upper*	*P(β < 0)*
NP complexity	0.85	0.63	1.07	< .001
N2 codability	0.23	0.03	0.43	.011
Modality	-0.69	-0.96	-0.43	> .999
NP complexity:N2 codability	0.11	-0.07	0.29	.110
NP complexity:Modality	-0.32	-0.51	-0.13	.999
N2 codability:Modality	-0.05	-0.24	0.11	.725
NP compl:N2 coda:Modality	0.11	-0.07	0.30	.113

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P(β < 0)* = probability of effect β being smaller than 0, colon *":"* = interaction

2.2.3 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 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 N2 in speech compared to writing before production began. Also looks to 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 fluency demands on the output. However fluency 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, where there is, arguably, pressure to produce fluent output, and in writing where this constraint is relaxed. 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 fluency 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 explanations. 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. Another, possibly conflating factor regards the structure of the elicit sentence. N2 was in the same clause from N1 for Complex NPs but in a different clauses for Simple NPs. One can therefore not rule out that planning obligated a clausal planning scope rather than just phrasal. This possibility is tested in Experiment 3.

2.3 Experiment 2

Findings from Experiment 1 are consistent with both syntactic and lexical planning embracing the whole of the subject noun phrase. It is possible, however, that how the referent images were presented artificially increased attention to N2 (the image bearing the codability manipulation) in the Complex condition. It is also possible that name agreement variation – the basis for the codability manipulation – is in part independent of ease of lexical retrieval.

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 et al., 2007; Kuchinsky et al., 2011) or by the prominence of a particular referent (Prat-Sala & Branigan, 2000). In the previous experiment N2 was the only novel image on the stimulus screen while the participant was familiar with the remaining images. This might have increased the prior attention to N2 and, hence, its incorporation into planning scope. To account for this problem, Experiment 2 used a novel image in the rightmost position on the stimulus screens.

Further, the position of the fixation target that appeared at the start of each trial broadly overlapped with the starting position of the image corresponding to the second noun. Although the first 100 ms of every trial were removed from the analysis this fixation target may have cued early lexical processing of N2. This methodological issue may also have influenced findings of previous studies in which the stimulus screen was preceded by a fixation crosses 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). Regardless of the presence or position of a target cross, participants may or may not have looked at the centre of the screen given the central fixation bias (see e.g. Parkhurst & Niebur, 2003). Therefore, advance sentence planning in previous studies might have been biased for any image located at the centre of the stimulus screen. 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.

Lastly, 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 that images with low agreement. 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.

One or more of these issues may have introduced early visual processing of N2 which triggered lexical processing that would not otherwise have occurred (i.e. that was not obligated by the language-production requirements of producing the sentence). Experiment 2 aimed at establishing whether evidence for advance lexical planning remained after the potential for experimental artifacts that push participants towards early preparation of N2 – arguably present in both Experiment 1 and a number of previous studies – had been removed.

2.3.1 Method

2.3.1.1 Participants

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

2.3.1.2 Design

Experiment 2 followed the same general design as Experiment 1. In a full factorial 2 × 2 × 2 design I 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:

(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.

The ease of lexical processing of N2 was manipulated by priming target names (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*).

2.3.1.3 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 if the most commonly used name is 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 to sample strings that matched the length of the image names (Medler & Binder, 2005). Stimulus items can be found in Appendix A.3.

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 A.4.

For the main experiment 96 images were shown in each condition (N2 prime by NP complexity by modality), counterbalanced in a Latin square design. Item sets were counterbalanced for whether *Peter* or *Tania* appeared in the leftmost 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).

2.3.1.4 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.

2.3.1.5 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.

2.3.2 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 trials with 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 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.

2.3.2.1 Onset latency

The onset latency data are shown in Table 2.8. A visualization of this can be found in Appendix A.5. To correct for positive skew the onset latency was logarithmically transformed for the analysis. The results of the Bayesian linear mixed model are presented in Table 2.9. The model revealed strong evidence ($BF_{10} > 5e5$) 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], $P(\beta < 0) < .001$, $BF_{10} = 516$) and speech ($\hat{\mu} = 0.22$, 95% CrI[0.15, 0.29], $P(\beta < 0) < .001$, $BF_{10} = 16822$). Moreover, longer onset latencies were found in writing compared to speech. This difference was not substantial ($BF_{10} = 2$). The data did not support an effect of lexical priming of N2, and did not support any interaction effects (all $BF_{10} < 1$).

Table 2.8: 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

Note: M = sample mean, *SE* = standard error, N = number of observations

Table 2.9: Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on onset latency (Experiment 2)

	Mean	*Lower*	*Upper*	*P(*$\beta < 0$*)*
NP complexity	0.05	0.04	0.06	$<$ .001
N2 prime	0.00	-0.01	0.02	.234
Modality	0.07	0.02	0.11	.001
NP complexity:N2 prime	-0.00	-0.01	0.01	.506
NP complexity:Modality	-0.00	-0.02	0.01	.752
N2 prime:Modality	-0.01	-0.02	0.00	.898
NP complexity:N2 prime:Modality	0.00	-0.01	0.02	.304

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P(*$\beta < 0$*)* = probability that effect β is smaller than 0, colon *":"* = interaction

2.3.2.2 Eye movements

Table 2.10 summarizes the proportion of eye samples to the referents of the first and second noun – N1 and N2 respectively – prior to production onset.

The proportion of eye samples was converted to empirical logits for statistical analysis. The results are shown in Table 2.11. The analysis revealed strong evidence ($\text{BF}_{10} = 827$) for an increased number of eye samples on N2 in Complex NPs compared to Simple NPs. Data also supported a main effect of priming N2 ($\text{BF}_{10} = 3$) with a larger proportion of eye samples in the non-word condition. This effect was necessarily driven by Complex NPs given the low proportion of eye samples on N2 in the Simple NP condition (see Table 2.10). Indeed, prime type contrasts nested within NP complexity revealed weak support ($\text{BF}_{10} = 2.5$) for priming effect in Complex NPs ($\hat{\mu}$ = -0.26, 95% CrI[-0.47, -0.06], $\text{P}(\beta < 0) = .994$) but no such effect in Simple NPs ($\hat{\mu}$ = -0.07, 95% CrI[-0.22, 0.09], $\text{P}(\beta < 0) = .822$, $\text{BF}_{10} < 1$). Further, a main effect of modality was found showing fewer eye samples on N2 in speech compared to writing ($\text{BF}_{10} = 50616$). An interaction was found for NP complexity by modality ($\text{BF}_{10} = 37$). This interaction was inspected in nested comparisons calculated from posterior samples of the Bayesian linear mixed model. Nested contrasts revealed strong support ($\text{BF}_{10} = 2592$) showing larger proportion of eye samples to N2 in Complex NPs in speech ($\hat{\mu}$ = 0.69, 95% CrI[0.42, 0.98], $\text{P}(\beta < 0) < .001$) but not in writing ($\hat{\mu}$ = 0.14, 95% CrI[-0.05, 0.33], $\text{P}(\beta < 0) = .082$, $\text{BF}_{10} < 1$). There was negligible evidence of other effects (all $\text{BF}_{10} < 1$).

Table 2.10: Descriptive data of proportion of eye samples to the referents (AOI) of the first and second noun – N1 and N2 respectively – prior to production onset. N2 is the referent of the critical noun bearing lexical priming manipulation (Experiment 2)

			N1			N2			
Modality	NP complexity	N2 prime	*M*	*Mdn*	*IQR*	*M*	*Mdn*	*IQR*	*N*
speech	Complex	image name	.69	.70	.34	.30	.29	.34	345
		non-word	.66	.68	.36	.33	.30	.36	339
	Simple	image name	.81	.88	.32	.18	.10	.31	342
		non-word	.82	.89	.33	.17	.09	.32	344
writing	Complex	image name	.74	.77	.45	.24	.22	.44	248
		non-word	.73	.73	.48	.25	.23	.46	259
	Simple	image name	.82	.98	.35	.15	.00	.31	226
		non-word	.81	.97	.36	.16	.00	.32	234

Note: *M* = mean, *Mdn* = median, *IQR* = interquartile range, *N* = number of observations

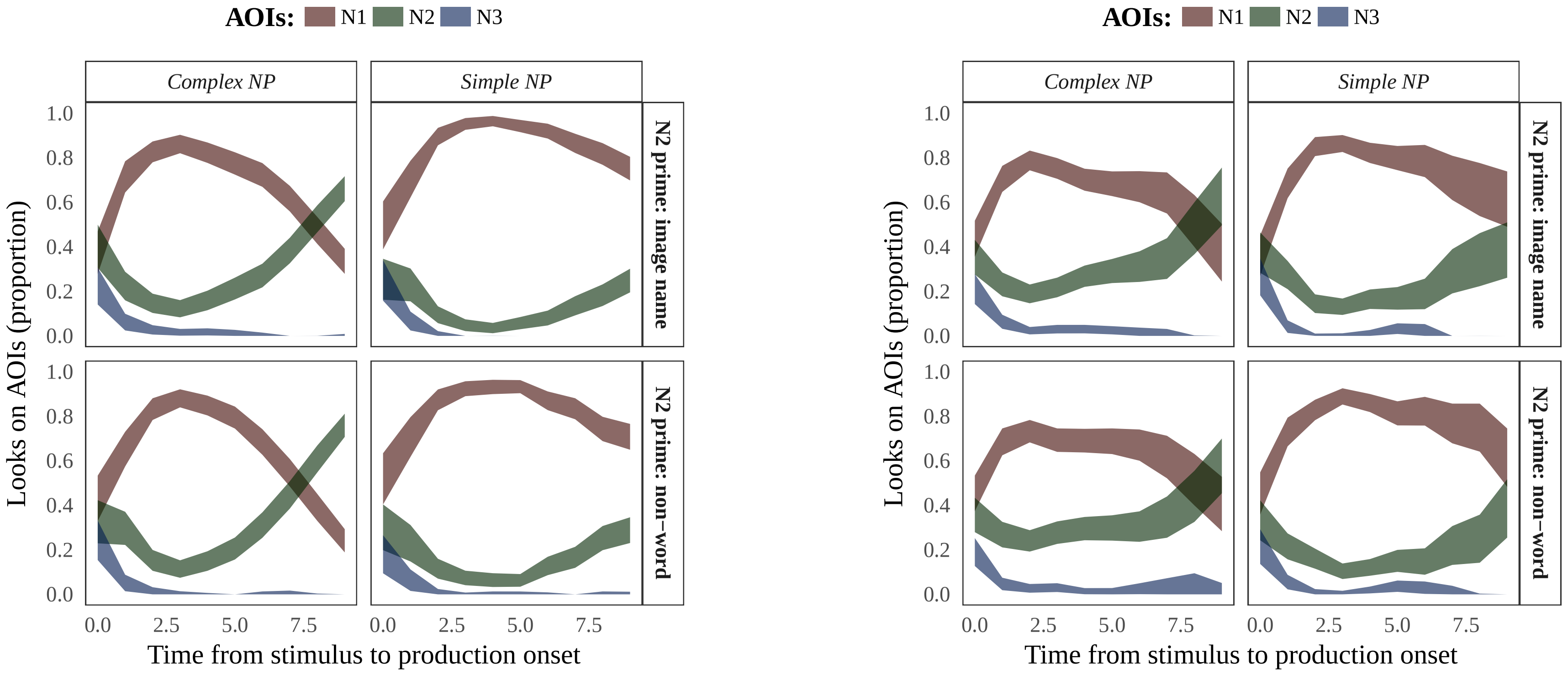
Table 2.11: Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on proportion of eye samples (in empirical logits) on N2 prior to production onset (Experiment 2)

	Mean	*Lower*	*Upper*	*P(*$\beta < 0$*)*
NP complexity	0.83	0.47	1.20	< .001
N2 prime	-0.33	-0.57	-0.07	.994
Modality	1.48	1.06	1.85	< .001
NP complexity:N2 prime	-0.19	-0.45	0.07	.928
NP complexity:Modality	-0.56	-0.87	-0.25	.999
N2 prime:Modality	0.25	0.00	0.49	.024
NP complexity:N2 prime:Modality	0.01	-0.26	0.28	.467

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P(*$\beta < 0$*)* = probability that effect β is smaller than 0, colon *":"* = interaction

The time course of the proportion of eye samples to each AOI is illustrated in Figure 2.4 for the time window before production onset summarized by condition and modality. These graphs illustrate the shift of attention away from N1 and towards N2 indicating processing shift from the first noun to the second noun (see e.g. Meyer & Lethaus, 2004).

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 N1 to N2 occurred and the proportion of trials for which this gaze shift happened before production onset. Gaze shift from N1 to N2 was defined as the first, at least 100 ms long, fixation on N2 after the gaze moved away from N1.



(a) Speech

(b) Writing

Figure 2.4: Proportion of eye samples to AOIs by condition over time from stimulus onset to production onset. AOIs are the referents of the three nouns (i.e. N1, N2, N3) as mentioned in the target sentence. The time axis was normalised within trial and binned. Bands indicate 95% confidence intervals (Experiment 2).

In 97% of the data gaze shift from N1 to N2 was detected either before or after production onset. The time of this shift relative to production onset is shown in Figure 2.5.

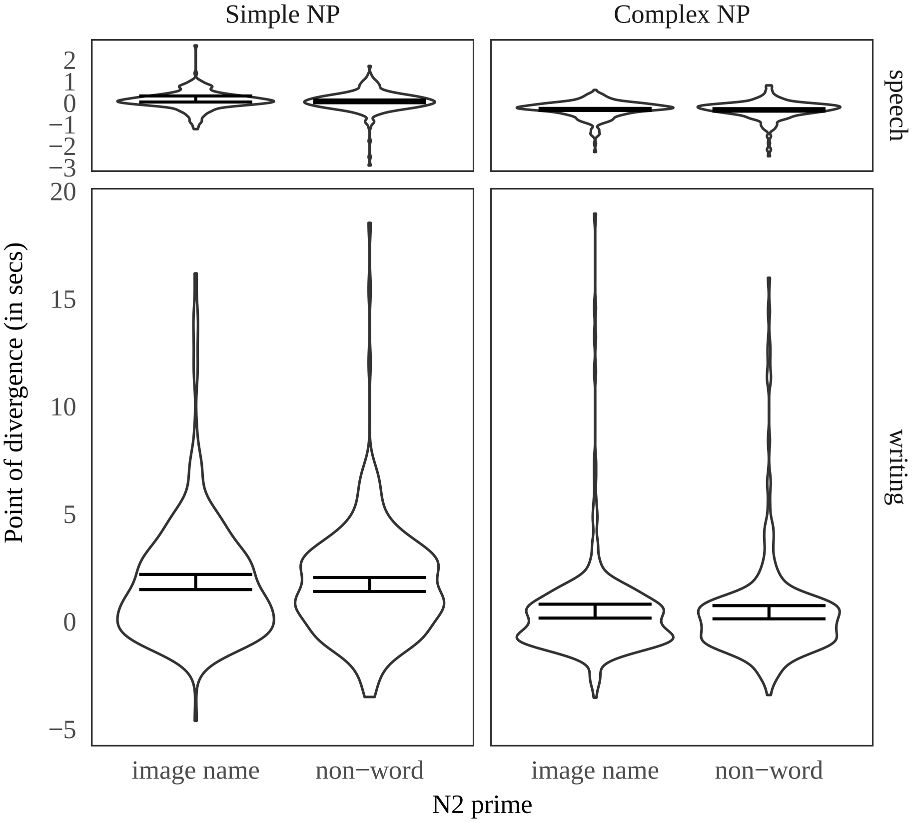


Figure 2.5: Bean plots of the point of gaze divergence from N1 to N2 relative to production onset. The beans illustrate the smoothed density of the distribution of the divergence data. The bands illustrate concentration in 95% CrIs. Null signifies the production onset (Experiment 2).

The data were log transformed to account for positive skew and analysed in a Bayesian linear mixed model with main effects and interactions of NP complexity, N2 Prime and modality. The results are summarized in Table 2.12. The model revealed substantial support for a main effect of NP complexity ($BF_{10} = 772$) supporting earlier gaze divergence 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], P($\beta < 0$) > .999, BF_{10} = 471) and speech ($\hat{\mu}$ = -0.17, 95% CrI[-0.26, -0.08], P($\beta < 0$) > .999, BF_{10} = 22). Further, there was strong support (BF_{10} = 49926) for a main effect of modality showing earlier gaze divergence in speech compared to writing. The support for all other predictors was negligible (all $BF_{10} < 1$).

Table 2.12: Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on point of divergence data (Experiment 2)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	-0.07	-0.09	-0.05	> .999
N2 prime	0.01	-0.01	0.02	.133
Modality	0.08	0.06	0.10	< .001
NP complexity:N2 prime	-0.00	-0.02	0.01	.675
NP complexity:Modality	-0.03	-0.05	-0.01	.997
N2 prime:Modality	0.00	-0.01	0.02	.352
NP complexity:N2 prime:Modality	-0.00	-0.01	0.01	.520

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability of effect β being smaller than 0, colon *":"* = interaction

The proportion of trials in which gaze divergence occurred before production onset is summarized by condition in Table 2.13. A logistic Bayesian mixed model was fitted on these proportions with the main effects and interactions of NP complexity, N2 codability and modality as predictors. The model outcome is summarized in Table 2.14. The model revealed strong support ($\mathrm{BF}_{10} > 8\mathrm{e}7$) for a main effect of NP complexity showing a larger proportion of gaze shifts before production onset in Complex NPs. Strong evidence ($\mathrm{BF}_{10} = 19107$) was found for a main effect of modality indicating larger proportions in speech compared to writing. There was strong support ($\mathrm{BF}_{10} = 181$) for the interaction of NP complexity and modality. This interaction was inspected in pairwise comparisons calculated from the posterior samples of the model. These comparisons revealed NP complexity effects in both speech ($\mathrm{BF}_{10} > 8\mathrm{e}6$) and writing ($\mathrm{BF}_{10} = 2239$). Within NP complexity comparisons showed strong evidence ($\mathrm{BF}_{10} = 137249$) for a larger probability of pre-onset gaze shift in Complex NPs for speech compared to writing ($\hat{\mu} = 5.34$, 95% CrI[3.91, 6.83], $P(\beta < 0) < .001$). The same effect was seen in Simple NPs ($\hat{\mu} = 1.55$, 95% CrI[0.28, 2.85], $P(\beta < 0) = .008$) which was weaker by comparison but yet substantial ($\mathrm{BF}_{10} = 13$). The evidence for all other model predictors was negligible (all $\mathrm{BF}_{10} < 1$).

Table 2.13: 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

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

Table 2.14: 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 divergence from first noun referent N1 to second noun referent N2 before production onset (Experiment 2)

	Mean	*Lower*	*Upper*	*P(*$\beta < 0$*)*
NP complexity	1.10	0.87	1.35	$<$ .001
N2 prime	-0.09	-0.25	0.07	.867
Modality	-0.86	-1.13	-0.60	$>$ .999
NP complexity:N2 prime	-0.01	-0.17	0.15	.575
NP complexity:Modality	-0.47	-0.70	-0.26	$>$ .999
N2 prime:Modality	0.09	-0.08	0.26	.145
NP complexity:N2 prime:Modality	-0.03	-0.20	0.13	.655

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P(*$\beta < 0$*)* = probability of effect β being smaller than 0, colon *":"* = interaction

2.3.3 Discussion

The aim of Experiment 2 was to establish whether results from Experiment 1 could be replicated after removing features of the method used in Experiment 1 (and previous similar studies) that might encourage greater attention to N2 prior to production onset. With regards to initial latencies, 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. However the lexical planning effect found on N2 in Experiment 1 – easily codable N2s gave shorter onset latencies in the Complex NP condition – was not replicated in Experiment 2. Pilot data from image naming indicated that the priming manipulation used in this experiment was effective in speeding lexical access. The absence of effects of N2 priming on initial latency therefore suggest that lexical retrieval beyond the first noun of the initial noun phrase is not obligated by the language production system. This finding is in line with syntax-based models of sentence production (e.g. Chang et al., 2006).

In the speech condition the proportion of eye samples on N2 was reduced for primed trials. The median proportion of looks to N2 was effectively zero for Simple NPs and there was no evidence for lexical priming in writing. This might be interpreted as evidence for an increased tendency to retrieve N2 prior to production onset where N2 was included in the sentence-initial phrase. However, because this effect was only present in speech, this is more readily attributed to the need for fluent spoken output than as a general feature of the language production system. Also note that there was no effect of N2 priming on any other eye measure or on the onset latency which may suggest that any lexical processing of N2 prior to production onset did not, in fact, play a role in preparing the subsequent utterance.

Experiment 2 therefore confirmed phrasal scope for sentence-initial planning but did not provide support for obligatory lexical retrieval beyond the first noun. The present experiment demonstrated that nouns within the sentence-initial phrase do not require lexical specification. The syntax of the first phrase is always planned but whether or not lexical retrieval exceeds the first noun is rather flexible (Wheeldon et al., 2013) and can be adjusted to serve the fluency demands of the output modality. These results suggest that attention dedicated to advance planning of N2 observed in the previous study was at least partially the consequence of increased prominence of N2 introduced by the experimental design.

However, in both Experiment 1 and Experiment 2 N2 was in the same clause from N1 for Complex NPs but in a different clauses for Simple NPs. These findings therefore do not rule out the possibility that the obligated planning scope is clausal rather than just phrasal. Experiment 3 tested this hypothesis by eliciting sentences in which N2 was always contained in the same clause from N1 in both Simple and Complex NPs.

2.4 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. Therefore, on the basis of the evidence from Experiments 1 and 2, the possibility cannot be dismissed that obligatory advance sentence planning scopes over the initial clause. This was suggested by some early studies (Bock & Cutting, 1992; Bock & Miller, 1991). Although several subsequent studies of spoken production have found evidence against clausal scope (see Smith & Wheeldon, 1999) this has yet to be tested in writing. Experiment 3 eliminated this possibility by eliciting transitive single-clause sentences (e.g. *N1 and the N2 moved above the N3*). If advance planning scopes over the sentence-initial clause rather than the sentence-initial phrase, one would not expect to observe the same pattern of results found in the previous two experiments. Instead, if sentence planning has a clausal scope, one would expect no NP complexity effects.

2.4.1 Method

2.4.1.1 Participants

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

2.4.1.2 Design & Material

This experiment used the same design and materials as Experiment 2. NP complexity was manipulated and crossed with the ease of N2 retrieval and output modality. In contrast to the previous experiments, participants were instructed to use descriptions as shown 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 (i.e. one clause).

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

2.4.1.3 Procedure & Apparatus

The procedure and the apparatus were the same as in Experiment 2.

2.4.2 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. Statistical analysis methods were the same as detailed for Experiment 1. For analysis of eye data further 12.5% were removed because proportion of total eye samples outside of defined AOIs was greater than .75. Statistical analysis followed the same methods as those described for Experiment 1.

2.4.2.1 Onset latency

Observed onset latencies are summarized in Table 2.15. For a visualisation of the entire distribution by-condition see Appendix A.6.

Table 2.15: 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

Note: M = sample mean, SE = standard error, N = number of observations

For statistical analysis the onset latency was square-root transformed to correct for positive skew. The results of the Bayesian linear mixed model are shown in Table 2.16. The model gave compelling evidence (BF_{10} = 3990) for longer onset latencies in Complex NPs compared to Simple NPs, and longer onset latencies in the written compared to the spoken output condition (BF_{10} = 1594). 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], P($\beta < 0$) < .001, BF_{10} = 197) and speech ($\hat{\mu}$ = 2.94, 95% CrI[1.58, 4.27], P($\beta < 0$) < .001, BF_{10} = 948). There was negligible support for a main effect of N2 prime (BF_{10} < 1). The posterior samples support a three-way interaction of NP complexity, N2 prime-type and output modality which was, however, not substantial (BF_{10} < 1). This suggests a follow-up inspection of N2 prime-type comparisons within NP complexity and output modality calculated from the posterior samples of the Bayesian linear mixed model. These nested comparisons revealed a tendency for shorter latencies in image name primes for Complex NPs when responses were spoken ($\hat{\mu}$ = -0.73, 95% CrI[-1.63, 0.13], P($\beta < 0$) = .946) which was not substantial (BF_{10} = 1.5). Negligible evidence of priming ($BF_{10} < 1$) was found in the remaining contrasts; either in Simple NPs in speech ($\hat{\mu}$ = 0.32, 95% CrI[-0.52, 1.17], P($\beta < 0$) = .23), or in Complex ($\hat{\mu}$ = -0.17, 95% CrI[-1.07, 0.72], P($\beta < 0$) = .647) or in Simple NPs in writing ($\hat{\mu}$ = 0.47, 95% CrI[-0.4, 1.36], P($\beta < 0$) = .151). No other interactions were supported by the data (all $BF_{10} < 1$).

Table 2.16: Main effects of first NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on onset latency (Experiment 3)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	0.66	0.43	0.91	< .001
N2 prime	-0.01	-0.22	0.20	.543
Modality	1.25	0.74	1.73	< .001
NP complexity:N2 prime	-0.05	-0.26	0.15	.685
NP complexity:Modality	-0.07	-0.30	0.15	.742
N2 prime:Modality	0.09	-0.16	0.34	.243
NP complexity:N2 prime:Modality	0.21	-0.01	0.43	.029

Note: *Mean* ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that β is smaller than 0, colon *":"* = interaction

2.4.2.2 Eye movements

The proportion of eye samples before production onset are summarized by condition in Table 2.17 for looks to both the referent of the first noun N1 and the referent of the second noun N2.

Table 2.17: Descriptive summary of proportion of eye samples to each referent (AOI) prior to production onset. N1 is the referent of first noun and N2 is the referent of the second noun – the critical noun bearing lexical priming manipulation (Experiment 3)

			N1			N2			
Modality	NP complexity	N2 prime	*M*	*Mdn*	*IQR*	*M*	*Mdn*	*IQR*	*N*
speech	Complex	image name	.67	.66	.31	.31	.31	.31	243
		non-word	.64	.63	.33	.35	.35	.34	228
	Simple	image name	.87	.99	.22	.12	.00	.21	239
		non-word	.85	.97	.27	.15	.01	.27	231
writing	Complex	image name	.70	.71	.50	.26	.25	.46	297
		non-word	.68	.68	.54	.29	.29	.51	272
	Simple	image name	.78	.90	.41	.18	.01	.36	263
		non-word	.81	.98	.37	.15	.00	.32	256

Note: M = mean, Mdn = median, IQR = interquartile range, N = number of observations

The Bayesian linear mixed model was fitted on the empirical logit of the proportion of eye samples. The results are shown in Table 2.18. The model revealed strong evidence ($\mathrm{BF}_{10} = 325$) supporting a main effect of NP complexity (showing enhanced proportions of eye samples on N2 for Complex compared to Simple NPs) and a main effect of modality (showing larger proportion of eye samples to N2 in speech compared to writing) ($\mathrm{BF}_{10} = 18$). There was further support ($\mathrm{BF}_{10} = 12$) for the modality by NP complexity interaction. This interaction was inspected in nested contrasts calculated from posterior samples. NP complexity type comparisons within modality revealed strong support ($\mathrm{BF}_{10} = 325$) for increased proportion of eye samples on N2 in Complex NPs for speech ($\hat{\mu} = 0.74$, 95% CrI[0.4, 1.08], $\mathrm{P}(\beta < 0) < .001$) but negligible evidence ($\mathrm{BF}_{10} < 1$) NP complexity difference in writing ($\hat{\mu} = 0.2$, 95% CrI[-0.01, 0.42], $\mathrm{P}(\beta < 0) = .029$). The evidence for all other model predictors was negligible (all $\mathrm{BF}_{10} < 1$).

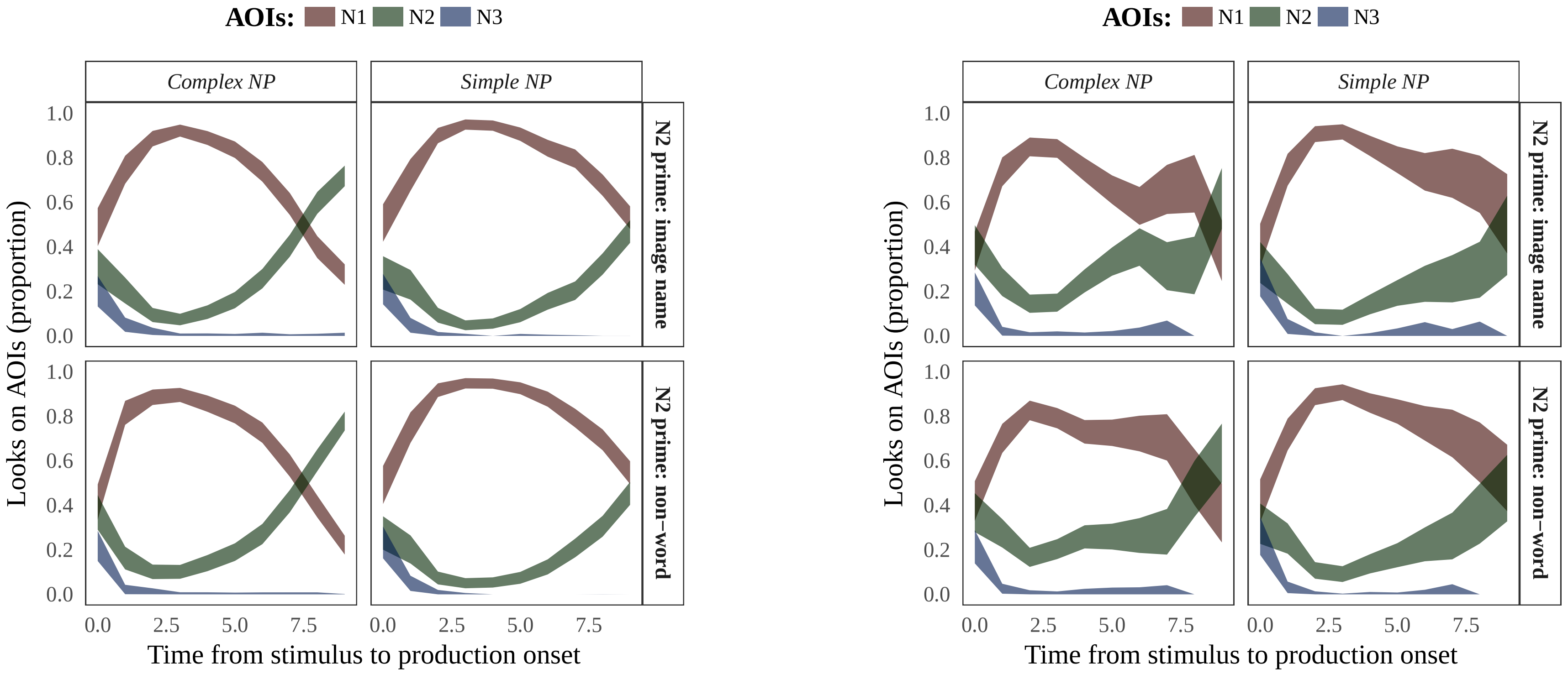
The time course of the proportion of eye samples to each AOI is illustrated by condition and modality in Figure 2.6 for the time period before production onset. These graphs illustrate the change of attention dedicated to N1 and N2 while there were only few eye samples on N3.

Table 2.18: Main effects of first NP complexity, prime on N2, modality and their interactions inferred from the Bayesian linear mixed model on proportion of eye samples on the referent of the image bearing the prime manipulation N2 prior to production onset (Experiment 3)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	0.94	0.51	1.39	$<$ .001
N2 prime	0.07	-0.15	0.29	.260
Modality	0.98	0.31	1.63	.003
NP complexity:N2 prime	-0.01	-0.25	0.22	.546
NP complexity:Modality	-0.54	-0.90	-0.17	.998
N2 prime:Modality	0.09	-0.13	0.31	.213
NP complexity:N2 prime:Modality	-0.15	-0.38	0.07	.911

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that effect β is smaller than 0, colon *":"* = interaction

As in the previous experiments, the time relative to production onset for when the gaze shift from N1 to N2 occurred was calculated as well as the proportion of trials for which this gaze shift happened before production onset. Gaze shift from N1 to N2 was defined as the first, at least 100 ms long, fixation on N2 after N1 was fixated. In 97% of the data this gaze shift from N1 to N2 was detected either before or after production onset. The time of divergence relative to production onset is shown in Figure 2.7.



(a) Speech

(b) Writing

Figure 2.6: Proportion of eye samples to AOIs by condition over time from stimulus onset to production onset. AOIs are the referents (i.e. N1, N2, N3) corresponding to the three nouns as mentioned in the target sentence. The time axis was normalised within trial and binned. Bands indicate 95% confidence intervals (Experiment 3).

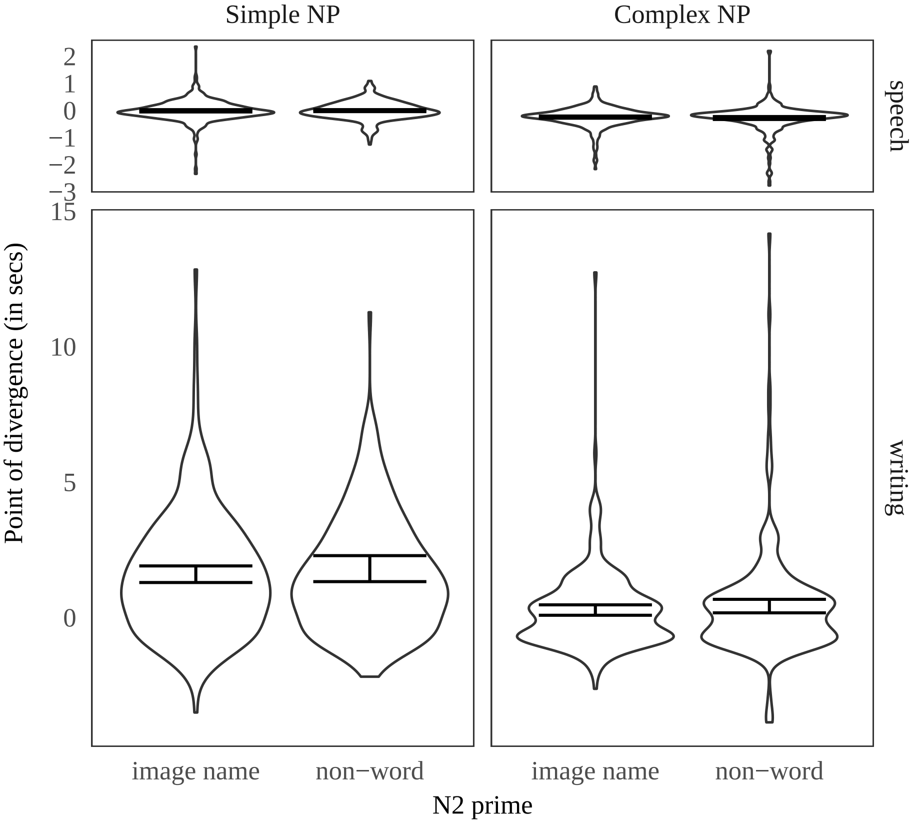


Figure 2.7: Bean plots of the point of gaze divergence from N1 to N2 relative to production onset. The beans illustrate the smoothed density of the distribution of the divergence data. The bands illustrate concentration in 95% CrIs. Null signifies the production onset (Experiment 3).

The data were log transformed to account for positive skew. The data were analysed in a Bayesian linear mixed model with main effects and interactions of NP complexity, N2 Prime and modality. The results are summarized in Table 2.19. The model revealed substantial support for a main effect of NP complexity ($BF_{10} = 3931$) supporting earlier gaze divergence in complex NPs. Further, there was strong support ($BF_{10} = 1063$) for a main effect of modality showing earlier gaze divergence in speech compared to writing. The model showed strong support ($BF_{10} = 23$)

for an interaction of these two main effects. The interaction was inspected in pairwise comparisons calculated from the posterior samples of the model. These comparisons revealed strong support (BF_{10} = 17) for earlier gaze divergence for Complex NPs in speech ($\hat{\mu}$ = -0.11, 95% CrI[-0.16, -0.06], $P(\beta < 0) > .999$). This effect was strongly supported in writing (BF_{10} = 4360) with a larger effect magnitude ($\hat{\mu}$ = -0.41, 95% CrI[-0.54, -0.29], $P(\beta < 0) > .999$). The support for all other predictors was negligible (all $BF_{10} < 1$).

The proportion of trials in which gaze divergence occurred before rather than after production onset is summarized by condition in Table 2.20.

Table 2.19: Main effects of NP complexity, prime on N2, modality and their interactions inferred by Bayesian linear mixed model on point of gaze divergence data – time relative to production onset (Experiment 3)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
NP complexity	-0.07	-0.08	-0.05	$> .999$
N2 prime	-0.00	-0.01	0.01	.816
Modality	0.09	0.07	0.12	$< .001$
NP complexity:N2 prime	0.00	-0.01	0.01	.297
NP complexity:Modality	-0.04	-0.05	-0.02	$> .999$
N2 prime:Modality	-0.01	-0.02	0.00	.921
NP complexity:N2 prime:Modality	-0.00	-0.01	0.01	.501

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability of β being smaller than 0, colon *":"* = interaction

Table 2.20: Descriptive summary for the proportion 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

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

A logistic Bayesian mixed model was fitted on the proportion data with the main effects and interactions of NP complexity, N2 codability and modality as predictors. The model outcome is summarized in Table 2.21. The analysis revealed strong support ($\text{BF}_{10} = 42932$) for the main effect of NP complexity showing a larger proportion of gaze shifts before production onset in Complex NPs. This effect was assessed in simple main effects within modality calculated from the model's posterior samples. These comparisons supported NP complexity effects for both writing ($\hat{\mu} = 2.66$, 95% CrI[1.61, 3.76], $P(\beta < 0) < .001$, $\text{BF}_{10} = 2237$) and speech ($\hat{\mu} = 3.09$, 95% CrI[2.08, 4.14], $P(\beta < 0) < .001$, $\text{BF}_{10} > 1e5$). Further, strong evidence ($\text{BF}_{10} = 2275$) 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 $\text{BF}_{10} < 1$).

Table 2.21: 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 divergence from N1 to N2 occurring before production onset (Experiment 3)

	Mean	*Lower*	*Upper*	*P(* $\beta < 0$ *)*
NP complexity	0.72	0.52	0.93	$<$ .001
N2 prime	0.04	-0.09	0.18	.279
Modality	-1.03	-1.43	-0.68	$>$ .999
NP complexity:N2 prime	-0.04	-0.16	0.08	.715
NP complexity:Modality	-0.05	-0.22	0.12	.733
N2 prime:Modality	0.08	-0.05	0.21	.100
NP complexity:N2 prime:Modality	0.05	-0.08	0.18	.231

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P(*$\beta < 0$*)* = probability of β being smaller than 0, colon *":"* = interaction

2.4.3 Discussion

The aim of Experiment 3 was to eliminate the possibility that the 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 borne out by the results of Experiment 3. The data provided 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 results obtained in Experiment 3 therefore exclude the clause as candidate for advance sentence planning and point towards the phrase as the fundamental planning unit in sentence production. 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 appears to be modality independent and is therefore plausibly a basic feature of the language production system.

This experiment provided evidence of shorter onset latencies for lexically primed referents in Complex NPs just when responses were spoken. This contrasts with production of the two clause sentences elicited in Experiment 2 for which no priming effect was observed in either modality. 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 speech is likely, in part, to be dependent on the requirement for fluent output. 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 required fluent output across the whole utterance. This contrasts with the two-clause structures elicited in Experiments 1 and 2 which will have had more relaxed fluency constraints:

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 judgement about the fluency 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.

2.5 General Discussion

The research presented here had two objectives: First, this study 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 dismiss the possibility that previously published findings were specific to the fluency demands imposed by spoken output. Second, this research aimed to establish whether phrase-level planning is lexically mediated, or whether planning of lexical items can potentially be delayed until after production onset. Again, previous research suggests that advance lexical planning might be the result of fluency demands of speech. Thus, crucially, establishing that results are common to both output modalities is quite strong 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 is, however not required, but may occur to meet output-fluency requirements. Before accepting this conclusion three possible alternative explanations for the presented findings are going to be discussed.

First, 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 (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 to ensure fluency after production onset. However, 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, Larochelle, & Grudin, 1988b; Olive, 2014). Therefore although in speech ease of retrieval of the sentence-initial noun might have encouraged more advance planning, 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.

Second, 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 lexical 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, lexical priming effects were in fact observed in Experiment 3. Second, pilot data (see Appendix A.4) indicated lexical priming using the same materials in an image naming experiment. Third, in analyses not reported in this chapter priming effects on production duration and of the attention to 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 N2 was reduced if primed. These three reasons suggest the lack of evidence for an effect of prime on production onset latency did not result simply from an ineffective priming manipulation.

A third 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 the presented results in two ways. It may be 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, however, 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 might be 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.

It remains possible, however, that syntactic planning scope is driven by variation in the conceptual representation of what needs to be expressed, i.e. scope effects are essentially a conceptual rather than a syntactic (or lexical) effect. In this study planning dedicated to the second noun in complex NPs 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 N2's 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/conceptual structure in sentence planning. This theory is being investigated in Chapter 3 showing that advance planning in complex noun phrases is the product of a pre-syntactic processing operation.

2.6 Conclusion

Advance processing of complex sentence-initial noun phrases has often been taken as evidence against lexical-incrementality, and in favour of syntax-based models of language production (e.g. Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Wagner et al., 2010; Wheeldon et al., 2013). The three experiments presented in this chapter addressed two possibilities, in line with lexically based models of sentence production, that explain this phrasal scope effect. First, advance processing beyond the first noun might be required to avoid intra-sentential hesitation (Allum & Wheeldon, 2007; Griffin, 2003). Complex NPs were found to be associated with longer production-onset latencies in both writing and speech. Second, advance planning may scope across the sentence-initial noun phrase, while only the sentence-initial content word requires lexical specification. This proposal assumes that syntactic planning is not mediated by lexical retrieval. Onset latencies were not affected by a manipulation designed to facilitate retrieval of the second noun in writing but found some evidence of this effect just in speech (Experiment 3). This research failed to find evidence that this advance grammatical planning requires (is driven by) lexical retrieval of the associated nouns. Lexical planning scope appears to be minimally incremental (Griffin, 2001) but may vary flexibly within the phrasal scope (see also Wheeldon et al., 2013). On the basis of these results the most parsimonious explanation is that sentence initial processing obligates advance planning of the syntactic structure of the sentence-initial phrase but permits lexical retrieval of non-initial items 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 a basic requirement of the language production system. The reported experiments are the most extensive (and possibly only) direct comparison of advance planning in speech and writing (and also the first systematic investigation of written sentence planning in general). This comparison provides strong evidence that phrasal planning scope in short sentence production is syntax-driven.

Chapter 3

Conceptual relations determine the extent of syntactic planning

Abstract

Producing syntactically complex phrases may or may not require planning beyond the initial determiner-noun pair. This was found to vary across syntactic phrase types. We hypothesised that the production system decides at a conceptual/semantic planning stage whether grammatical encoding has to span an incrementally simple noun or a hierarchically complex phrase structure as advance planning unit. In three image description experiments subordinated noun phrases were elicited by manipulating the contrasting meaning of the first and the second noun. Evidence from eye movement data revealed a larger proportion of looks and a pre-onset gaze shift towards the referent of contrastive nouns in non-phrase initial position while controlling for syntactic structure and lexical content. This shows that advance planning of complex noun phrases is determined by conceptual relational means.

3.1 Introduction

Reading, listening, speaking, and writing are intuitively perceived as processes that unfold in a linear fashion. The underlying structure of language, however, is hierarchically complex. Semantic and syntactic dependencies frequently exist between non-adjacent words. In the phrase *the blue hat*, for example, the determiner "the" is both semantically and syntactically dependent on "hat" and not "blue". These non-adjacencies affect how language is processed. In reading, longer distance between depending elements increases processing difficulty (e.g. Gibson, 2000). In language production the modifier "blue" in a noun phrase such as *the blue hat* needs to be preplanned before production onset in languages in which the adjective comes before the nominal head as in German and English, but not in languages in which it follows the noun as in French and Spanish (Brown-Schmidt & Konopka, 2008; Schriefers & Teruel, 1999; Schriefers, Teruel, & Meinshausen, 1998). While advance planning in sentence production might scope over hierarchically complex units (Allum & Wheeldon, 2007; E.-K. Lee et al., 2013; Martin et al., 2014; Smith & Wheeldon, 1999), lexical specification of the first lexical item is all that must be planned before output onset (F. Ferreira & Swets, 2002; Griffin, 2001; Zhao & Yang, 2016). Recent discussions in the literature on advance sentence processing have moved away from whether or not planning is incremental and towards the proposal that non-incremental advance planning is necessitated under certain conditions (Bock & Ferreira, 2014; Konopka & Kuchinsky, 2015; Konopka & Brown-Schmidt, 2014; Konopka & Meyer, 2014; Kuchinsky, 2009; Kuchinsky et al., 2011;

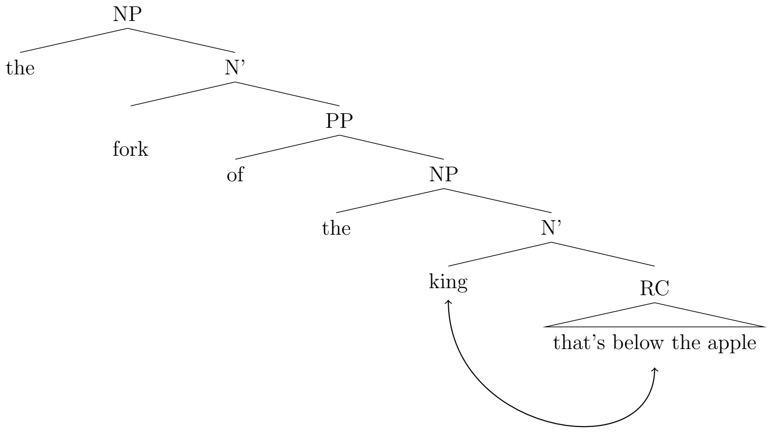
E.-K. Lee et al., 2013). The present chapter is concerned with the question of how the language production system decides whether or not to preplan units beyond the first determiner-noun pair in complex noun phrases.

In the language production model outlined by Bock and Ferreira (2014) processing syntactic dependencies starts with a rudimentary abstract representation or a "scaffold" that is created from a conceptual representation which feeds into the syntactic assembly. This theory builds on many other related models of language production (e.g. Bock, 1990; Chang et al., 2000, 2003, 2006; Costa & Caramazza, 2002; Dell, 1986; Dell & O'Seaghdha, 1992; V. S. Ferreira & Dell, 2000; Garrett, 1975). The increment that feeds into syntactic assembly requires at minimum the retrieval of the first lexical unit (e.g. Brown-Schmidt & Konopka, 2008; F. Ferreira & Swets, 2002; Griffin, 2001; Griffin & Bock, 2000; Wheeldon et al., 2013; Zhao & Yang, 2016). For example, Griffin (2001) elicited sentences with coordinated subject noun phrases. She manipulated correlates for the ease of lexical planning and found evidence for the lexical preparation of the sentence-initial noun but not further (see also Brown-Schmidt & Konopka, 2008; Zhao & Yang, 2016). However, under certain conditions advance planning can scope across and beyond the phrase (e.g. Bock & Miller, 1991; Meyer, 1996; Smith & Wheeldon, 1999). Evidence for such non-incremental, i.e. hierarchical, scaffolding comes from studies that predicted the duration required to release the production onset for a sentence from the structural complexity of the sentence-initial subject noun phrase (Allum & Wheeldon, 2007, 2009; F. Ferreira, 1991; Konopka, 2012; Levelt & Maasen, 1981; Martin et al., 2010, 2014; Nottbusch, 2010; Nottbusch et al., 2007; Smith & Wheeldon, 1999; Wagner et al., 2010; Wheeldon et al., 2013). These planning durations were found to be longer for coordinated subject phrases of the form *The A and the B* compared to simple determiner-noun pairs (e.g. *The A*) irrespectively of planning costs associated with lexical retrieval of the second noun (*the B*) (Griffin, 2001; Smith & Wheeldon, 1999; Wheeldon et al., 2013). Further Allum and Wheeldon (2007) showed that phrases with subordinated modification (e.g. *The A above the B*) require less advance planning than coordinated noun phrases regardless of the structural position of the nominal head noun. Conversely Nottbusch (2010) and Nottbusch et al. (2007) provided evidence for more advance planning for subordinated modifiers compared to coordinated noun phrases which was attributed to advance planning of syntactically embedded structures. In spite of authors' varying explanations these data suggest that some processing with respect to the organisation of the noun phrase – the scaffold – is required to determine whether or not advance planning has to extend beyond the first noun. Sentence planning can operate both incrementally and non-linearly hierarchically due to circumstances that are not yet well understood (Bock & Ferreira, 2014; Brown-Schmidt & Konopka, 2015; Kuchinsky, 2009; Kuchinsky et al., 2011; E.-K. Lee et al., 2013; Swets et al., 2014). The question that arises is not if but why the production system, in some instances, operates non-linearly using increments larger than the first determiner-noun pair (Konopka & Kuchinsky, 2015; Konopka & Bock, 2009; Konopka & Brown-Schmidt, 2014). In other words, under which conditions is structural assembly initiated before production onset?

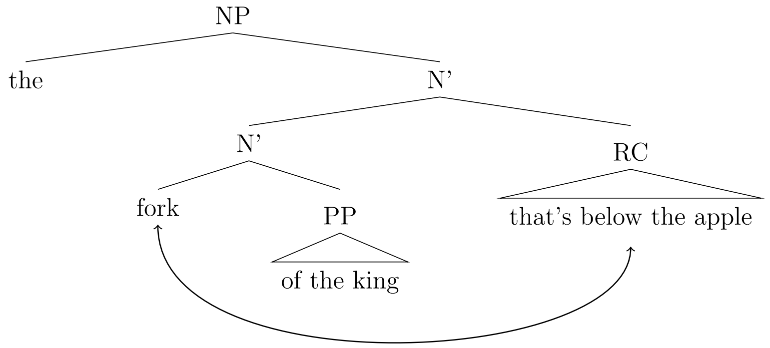
Speakers employ a minimal advance planning scope under time-pressure (F. Ferreira & Swets, 2002) and under increased processing load (Martin et al., 2014; Wagner et al., 2010). Further, speaker-specific characteristics such as a low working memory capacity (Swets et al., 2014; Van de Velde & Meyer, 2014; Wagner et al., 2010; Wheeldon et al., 2013) and older age (Griffin & Spieler, 2006) lead to reduced advance sentence preparation. Thus, planning smaller units in advance is generally understood to be less demanding for the linguistic processor (e.g. Christiansen & Chater, 2016; Levelt & Meyer, 2000). Non-linear planning, on the other hand, was observed when linear incremental processing is not easily available, for example, if events were difficult to encode linguistically (Kuchinsky, 2009; Van de Velde et al., 2014), when producing non-canonical verb-argument structures (Momma, Bowen, & Ferreira, 2017), for phrases with syntactically arbitrary noun order (Allum & Wheeldon, 2007, 2009), topicalised syntactic objects (Do, Kaiser, & Zhao, 2017), and perhaps to satisfy fluency demands on spoken utterances as discussed in Chapter 2.

To some extent the unique identifiability of the referent of a definite noun phrase may lead to advance non-linear planning (Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006; Konopka & Brown-Schmidt, 2014; E.-K. Lee et al., 2013; Swets et al., 2014). Speakers do not necessarily plan conceptual information in advance because new incoming information can easily be included into the sentence even after production onset (Brown-Schmidt & Konopka, 2015). Also the syntactic configuration of the utterance may influence advance planning. For example in a task where subjects were required to use short descriptions such as *the small butterfly* to direct a confederate to highlight an image, Brown-Schmidt and Konopka (2008) found longer fixations on a contrast depicting image (e.g. a larger butterfly) before the production onset. Spanish speakers, however, showed a diminished tendency to fixate on the contrast image before production onset as, unlike English, the colour adjective used as means of contrast disambiguation in Spanish is encoded postnominally (*la mariposa pequeña*).

E.-K. Lee et al. (2013) examined directly whether advance planning is mediated by syntactic dependencies between phrases elements. The authors elicited utterances of the form *(Click on) the fork of the king (that is) below the apple* in response to arrays of images. Reference to the correct target referent (i.e. the fork) required subjects to use two modifiers – a postnominal possessive (*of the king*) and locative relative clause (*that is below the apple*). The location of either the fork or the king in relation to the image of an apple changed the underlying syntactic structure of the utterance without affecting its surface form. The relative clause might either attach high or low in the syntactic tree (see Figure 3.1). Low attachment at the *king* node (Figure 3.1a) conveys that the king is below the apple; high attachment to the *fork* node (Figure 3.1b) expresses that the fork was below the apple. E.-K. Lee et al. (2013) predicted if phrase planning is mediated by hierarchical dependencies, longer onset latencies would be expected for high-attaching phrases as these require anticipatory planning of the relative clause modifier. This was what the authors found. They concluded that hierarchically organised phrases necessitate non-linear advance processing. Conversely, shorter onset latencies were observed for low attaching relative clauses because incremental planning allows the delay of some processing until after production onset.



(a) Low attachment: relative clause modifies the head of the propositional phrase (PP) *king*, i.e. the king is below the apple



(b) High attachment: relative clause modifies the head noun *fork*, i.e. the fork is below the apple

Figure 3.1: Syntactic tree structures for each relative clause (RC) attachment sites in the noun phrase (NP) (see E.-K. Lee et al., 2013). Arrows highlight dependencies between relative clause and noun.

E.-K. Lee et al. (2013) understood this effect as being driven by hierarchical syntactic dependencies. There are various sources of pressure on the production system leading to a minimal advance plan (F. Ferreira & Swets, 2002; Griffin, 2003; Griffin & Spieler, 2006; Levelt & Meyer, 2000). However, non-adjacent dependencies lead to utterance-initial structural assembly extending beyond the initial determiner-noun pair (Brown-Schmidt & Konopka, 2008; E.-K. Lee et al., 2013; Martin et al., 2014; Smith & Wheeldon, 1999). Thus incremental planning is not always going to be possible. If advance planning requires processing beyond the first increment, speakers typically require more time to release the production onset (E.-K. Lee et al., 2013; Martin et al., 2014; Smith & Wheeldon, 1999).

However, in order for incremental processing to facilitate advance planning, the production system must decide prior to the structure-building operation whether to operate incrementally or whether planning must scope beyond the first noun. In other words, before submitting a message to the language processor, the size of the message chunk needs to be determined. Kempen and Hoenkamp (1987, p. 205) pointed out that advance planning could, in principle, involve the parallel generation of temporary structures. Both the incremental and the hierarchical structure could be preplanned and the wrong structure would eventually be discarded. Alternatively, structures might be selected randomly. If the selected structure does not fit the message, it would need to be changed or modified. However, if it were indeed the case that alternative structures were preplanned at random or in parallel, one would not expect to observe processing facilitation when “grammatical encoding” – the process that translates a message into language by generating syntactic structure and lexical material – is permitted to unfold incrementally as shown by E.-K. Lee et al. (2013). Therefore, it is less plausible that the language processor decides from the generated syntactic structure whether or not incremental planning is an option.

Instead it seems to be the case that relations between message elements need to be determined pre-syntactically. Such a primary conceptual plan is typically understood as unordered (Konopka & Brown-Schmidt, 2014) and is guided by an interaction of both linguistic preferences and perceptual features (Gleitman et al., 2007; Kuchinsky, 2009; Kuchinsky et al., 2011; Tomlin, 1995). Rudimentary conceptual representations are built during swift visual apprehension of the stimulus screen (Griffin, 2004b; Griffin & Bock, 2000; Meyer & Lethaus, 2004). Konopka and Meyer (2014, p. 2) distinguish two processes that are required to map a message onto a linguistic structure: individual elements of the message need to be encoded which the authors dubbed the *non-relational* process; the relationship between these elements needs to be determined which constitutes the *relational process*. Some authors claim that advance planning gives priority to non-relational information (Gleitman et al., 2007). Thus message planning might scope over the first word only (Brown-Schmidt & Konopka, 2008, 2015). Others argued that the relationship between message elements guides planning (Bock, Irwin, & Davidson, 2004; Kuchinsky, 2009). Konopka and Meyer (2014) suggested that speakers can prioritise either process depending on contextual information.

Relational information is clearly important for preplanning a noun phrase. If the relationship between two nouns – sub- or coordinated – does not affect advance planning, it would be difficult to explain why sentence-initial noun phrases such as *The A and the B* systematically lead to longer planning durations compared to simple (*The A*) and subordinated (*The A above the B*) (e.g. E.-K. Lee et al., 2013; Martin et al., 2014; Wagner et al., 2010; Wheeldon et al., 2013 but see Allum & Wheeldon, 2007; Griffin, 2003). Most of these authors propose that advance planning for coordinated noun phrases is determined during grammatical encoding. On the other side Griffin (2003) suggested that planning beyond the first noun serves a modality-related function. Speech requires a certain extent of output fluency. Therefore preplanning might occur before production onset if there is no time to plan the second noun in parallel after production onset without intra-sentential pausing.

Furthermore, Allum and Wheeldon (2007) suggest that coordinated noun phrases require advance planning as there is no hierarchical dominance, i.e. the order of nouns in coordinated phrases is arbitrary. This explanation entails some form of pre-syntactic conceptual process that (a) determines whether or not the order is arbitrary or grammatically structured as in subordinated phrases (Allum & Wheeldon, 2007), and (b) determines the relation between the nouns which makes coordinated (i.e. conjunctive) noun phrases different from disjunctions (e.g. *The A* ***or*** *the B*). In the same way the syntactic structures elicited by E.-K. Lee et al. (2013) entail different semantic representations, i.e. the different stimulus arrays used to manipulate attachment ambiguity unavoidably not only elicit different syntactic structures but also convey different meanings. It is therefore not clear whether the observed effect is rooted in syntactic planning (E.-K. Lee et al., 2013) or in the generation of conceptual relations. The latter is plausible because both syntax and semantics have non-linear structures (Konopka & Brown-Schmidt, 2014). In order to allow the processor to plan incrementally the processor needs to decide prior to syntactic assembly whether or not hierarchical planning is required. In other words, the processor needs to decide the relationship between message elements before the language production system can encode the syntactic relation (i.e. linearity) between message elements. The hypothesis is that linearity in advance sentence planning is determined at the message level.

In sum, advance phrase planning may be guided both incrementally and hierarchically (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014; Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014). It follows that structural assembly may occur before or after production onset. Research has yet to determine which factors lead to advance structural planning.

The present research aimed to answer whether advance structural assembly of hierarchically complex noun phrases is determined by conceptual relations independently of syntactic dependencies. Specifically, it was tested whether relational information induces planning beyond the first increment in subordinated noun phrases. In this chapter I report three image description experiments in which subjects were required to produce modified noun phrases. All experiments manipulated semantic contrast to examine the influence of conceptual structure on phrase preplanning whilst keeping the syntactic structure and lexical content constant. Semantic contrast (often referred to as *focus*) distinguishes between a current information and potential alternatives (Jackendoff, 1972). In English contrast can be encoded by means of prosodic prominence (Selkirk, 1995). For instance, the phrase *The ball above the WINDOW* (capitals indicate prosodic prominence) refers to a ball that is located above a window as suppose to another ball above, say, a door. In this example the conceptual scope embraces both nouns disambiguating the reference of the first noun. On the other hand *The BALL above the window* distinguishes between different objects above the window and highlights the ball as opposed to, say, the racket above the window. In the latter example, the conceptual scope embraces the first noun only.

To test whether planning of conceptual relations exceeds the first increment, eye movements to the stimulus screen were monitored. Additionally both pre- and post-nominal modifier phrases were tested as the pre-syntactic nature of conceptual planning should induce advance planning across noun phrase modification type. The prediction is that if conceptual relations impact advance phrase planning, non-initial nouns with a contrastive function require anticipatory planning.

Experiment 1 used a written image description task to establish whether semantic contrast increases the planning scope. Experiment 2 and Experiment 3 used different designs with the same contrast manipulation. Experiment 2 eliminated a possible confound of Experiment 1. Experiment 3 used an interactive image description task and tested the semantic contrast manipulation in both writing and speech. The last experiment is a development of the first two experiments and addressed potential problems detected in the previous studies.

3.2 Experiment 1

Experiment 1 aimed to answer whether advance structural assembly of hierarchically complex noun phrases is determined by conceptual relations independently of syntactic dependencies. Semantic contrast was manipulated for the first and the second noun in subordinated noun phrases.

The prediction was that a contrastive second noun requires preplanning as it is relevant for the reference of the first noun. Crucially, advance planning under semantic contrast was tested by keeping the syntactic structure and the surface form constant.

3.2.1 Method

3.2.1.1 Participants

32 psychology students (5 male, mean age = 19.7 years, *SD* = 3.0, range: 18–32) participated as part of a research-reward scheme. All participants were self-reported native speakers of British English, free of linguistic impairments, and had normal or corrected-to-normal vision. Eight participants were replaced because they failed to produce a sufficient number of descriptions that matched the targeted structures.

3.2.1.2 Design

Participants were asked to write (i.e. type on a keyboard)[1] descriptions that unambiguously identify the coloured object in provided arrays of images as shown in Figure 3.2. The elicited description started either in a prenominal possessive modifier (e.g. *The cowboy's hamburger is green*) or a noun with a postnominal modifier (e.g. *The cowboy (that is) above the hamburger is green*). Participants were asked to use possessives when a line connected the target to another image (see Figure 3.2a and 3.2b) or postnominal modifiers indicating the vertical position of the target image otherwise (see Figure 3.2d and 3.2c).

---

[1] Writing as output modality allows to by-pass alternative explanations for advance structural planning such as preplanning to maintain fluency after production onset (Griffin, 2003).

(a) Possessive, N2 Contrast **(b)** Possessive, N1 Contrast

(c) Postnominal, N2 Contrast **(d)** Postnominal, N1 Contrast

Figure 3.2: Stimulus arrays (Experiment 1).

Modifier Type (possessive, postnominal) was crossed with Noun Contrast in a full factorial 2 × 2 design. Semantic Noun Contrast for the second noun (N2) was induced by virtue of a conceptually identical images for the first noun (N1) of the target phrase (henceforth, N2 Contrast). N2 Contrast refers to the contrastive function of N2 to specify the reference of N1; e.g. *cowboy's HAMBURGER* and not cowboy's brush, *cowboy above the HAMBURGER* and not the cowboy above the brush (see Figure 3.2a and 3.2c, respectively). On the other hand N1 Contrast refers to situations when N1 disambiguates the reference of N2; e.g. *COWBOY's hamburger* and not robot's hamburger or *COWBOY above the hamburger* not the robot above the hamburger (see Figure 3.2b and 3.2d, respectively).

Participants' eye movements and the stimulus-to-onset latency were recorded as indicators of advance planning. The time between the presentation onset of the stimulus array on the computer screen and the start of the sentence, i.e. the onset latency, reflects difficulty related to utterance planning process. Eye movements to the stimulus screen prior to response onset were recorded as indicator of advance sentence planning. The areas of interest (henceforth, AOI) are the images corresponding to each noun in the produced utterance. Eye movements to these referents, and in particular conditional changes of the gaze pattern, give an indication of how the stimulus screen was encoded.

3.2.1.3 Materials

32 items with arrays of four images were created with animate entities in one row and inanimate objects in the other row. Black and white drawing were taken from the database of the International Picture Naming Project (E. Bates et al., 2003; Székely et al., 2003, 2004, 2005). Items were distributed across four Latin square lists to counterbalance for Modifier Type and Noun Contrast. Also, the horizontal and vertical position of the target group were counterbalanced to make the position of the target (i.e. the coloured referent) unpredictable. The colour of the target was varied between items (i.e. green, red, blue, yellow). 48 fillers were added to each list. Filler arrays contained less than three images with no coloured image, one coloured image or more. Each list was presented in randomised order. A list of the stimulus material can be found in Appendix B.1.

3.2.1.4 Procedure

Every experimental session started with camera set-up and calibration and validation. Participants were familiarised with the intended constructions during instructions and a practice phase with 10 items. The experimenter monitored descriptions during the practice unit and corrected the structures indirectly, if necessary. A recalibration was performed before the experimenter left the lab. Each trial started with a centred fixation trigger that activated the image array. A recalibration was performed if the fixation trigger, an ellipsis (21 × 21 pixels) that required a fixation of 200 ms, failed. The images, size 200 × 200 pixels (including transparent margins), appeared equally spaced around the centre of the screen. A text box (896 × 50 pixels) was provided at the bottom of the screen in which the produced text was displayed. All images were shown on the screen until the participant finished the description by pressing return. Pauses were possible before/after each trial. The duration of the entire experiment was approximately 45 minutes.

3.2.1.5 Apparatus

Eye movements were recorded using an SR Research EyeLink 1000. The eye tracker was desk mounted and used in remote mode to ensure free head movements. Eye data were sampled monocular (right eye) on a frequency of 500 Hz. The participant was seated 55 to 60 cm away from the lens. The experiment was build in SR Research Experiment Builder. Stimuli were displayed on a 19” ViewSonic Graphic Series (G90fB) CRT monitor with a screen resolution of 1,280 × 1,024 pixels and 85 Hz refresh rate using an Intel Core 2 PC.

3.2.2 Results

All data were analysed using hierarchical Bayesian linear mixed effects models (Gelman et al., 2014; Kruschke, 2014). The probabilistic programming language Stan and the R interface Rstan (Carpenter et al., 2016; Hoffman & Gelman, 2014; Stan Development Team, 2015a, 2015b) was used along with the rstanarm package (Gabry & Goodrich, 2016) and an adapted version of the Stan code used in D. Bates et al. (2015). Models were fitted with maximal random effects structures if not stated otherwise (Barr et al., 2013; D. Bates et al., 2015). To assess the effect magnitude by the modelled slopes all predictors were sum coded. The 95% posterior probability mass – 95% credible intervals; henceforth, CrI – was calculated from the posterior samples. 95% CrIs that do not contain zero support the presence of an effect of the predictor onto the outcome variable. This probability mass will also be expressed as the proportion of posterior samples smaller than zero; henceforth $P(\beta < 0)$. This proportion indicates the probability, given the observed data, of observing a negative effect. If this proportion is approaching zero there is support for a positive effect. In contrast, if this proportion is approaching one, a negative effect effect is supported by the data. Inconclusive evidence would include large amounts of posterior samples of either polarity (see Kruschke et al., 2012; Sorensen et al., 2016). Finally, to assess the strength of support for a given effect of interest over the null hypothesis, Bayes Factors were calculated using the Savage-Dickey method for nested models (see Nicenboim & Vasishth, 2016; Dickey et al., 1970) (henceforth, BF_{10}). BF_{10} larger than 10 indicates strong support for a difference while extremely small BFs_{10} suggest evidence against the alternative hypothesis. Substantial evidence requires at least a BF between 3-5 (see e.g. Baguley, 2012; M. D. Lee & Wagenmakers, 2014; Wagenmakers et al., 2010). For example $\mathrm{BF}_{10} = 2$ means that the data are two times more likely under the alternative hypothesis than under the null hypothesis.

Models were fitted with weak, locally uniform priors and by-subject and by-item adjustments using LKJ priors on the correlation matrix of the variance-covariance matrix (see Nicenboim & Vasishth, 2016; Sorensen et al., 2016). If not specified differently, four chains with 2,000 iterations per chain were run with a warm-up of 1,000 iterations and no thinning. Model convergence was confirmed by Rubin-Gelman statistics ($\hat{R} = 1$) (Gelman & Rubin, 1992) and inspection of the Markov chain Monte Carlo chains.

The following data screening criteria were used prior to analysis. Trials with extremely long pausing or rephrasing of already produced text after response onset were removed (12.99%). Sentences that did not match the expected structures were excluded from the analysis (2.64%). Moreover, trials with onset latencies $>$ 14 secs (0.4%) and production durations $>$ 30 secs (1.17%) were discarded. In total, 15.4% of the data were removed. For the analysis of eye data further 0.7% were removed due to a proportion of eye samples larger than .75 outside of AOIs. Code and data are available on figshare.com/s/3dcaefc9082f0ab85ddf.

3.2.2.1 Onset latency

The onset latency is summarized in Table 3.1. For a visualisation of the distribution in bean plots (Kampstra, 2008; Phillips, 2016) see Appendix B.2.

Table 3.1: Descriptive summary of onset latency in ms (Experiment 1)

Modifier Type	Noun Contrast	*M*	*SE*	*N*
possessive	N1	2197	98	215
	N2	2471	133	220
postnominal	N1	1941	98	209
	N2	1987	104	222

Note: M = sample mean, SE = standard error, N = number of observations

For statistical analysis the onset latency was transformed to the reciprocal of its square root. The transformed onset latency was fitted in a Bayesian linear mixed model with main effects of Modifier Type, Noun Contrast and their interaction. The model output is summarized in Figure 3.3.

The analysis revealed evidence for a main effect of Modifier Type ($\text{BF}_{10} = 438$) showing shorter latencies for postnominal phrases. There was weak evidence ($\text{BF}_{10} = 2.4$) for a main effect of Noun Contrast indicating longer latencies for N2 Contrast. No substantial support ($\text{BF}_{10} = 1.1$) was found for an interaction. Further, Noun Contrast differences were assessed as simple effects within Modifier Type revealing longer onset latencies for N2 Contrast in possessives ($\hat{\mu} = 0.87$, 95% CrI[-

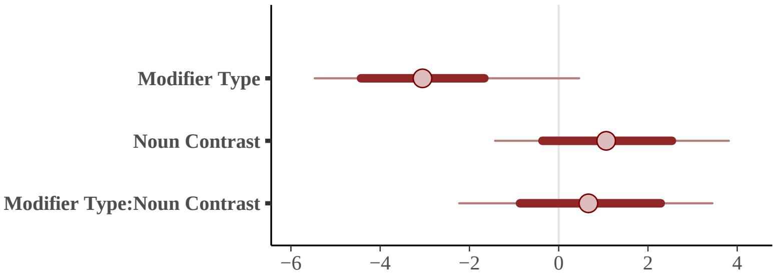


Figure 3.3: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the onset latency (Experiment 1). Dots indicate the posterior mean $\hat{\mu}$, the thick lines show the range of 95% of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons *":"* denote interactions.

0.17, 1.93], P($\beta < 0$) = .05) without substantial support ($\text{BF}_{10} = 2$). No Noun Contrast effect was found in postnominals ($\hat{\mu} = 0.2$, 95% CrI[-0.84, 1.26], P($\beta < 0$) = .36, $\text{BF}_{10} = 0.5$). Figure 3.4 summarizes these comparisons calculated from the posterior predicted values of the Bayesian model.

3.2.2.2 Eye data

Planning of the N2 referent was assessed using the divergence of looks from N1 to N2 (see e.g. Griffin & Bock, 2000; Konopka & Meyer, 2014), i.e. the time when participants stopped looking at the referent of N1 and started looking at the referent of N2. This point of divergence was estimated by trial as the onset of the first fixation on N2 (minimum duration 100 ms) after the first fixation on N1 ended. The time point of this gaze shift relative to production onset and the probability that this

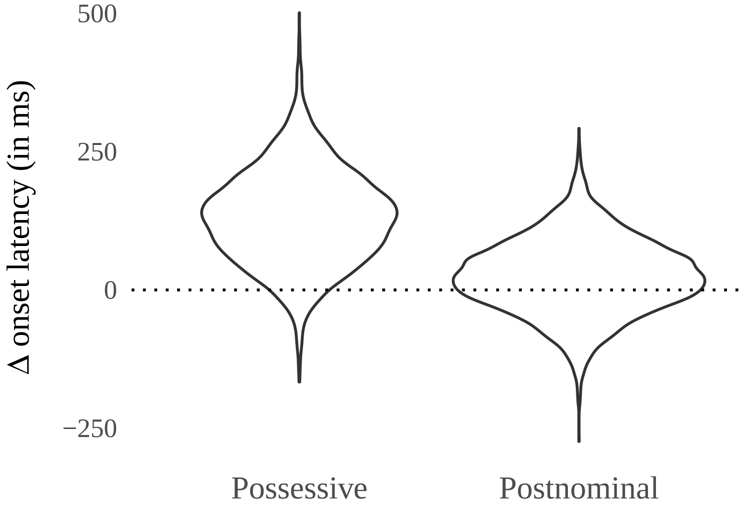


Figure 3.4: Summary of latency difference compared to 0 for Noun Contrast (Δ onset latency = N2 Contrast–N1 Contrast) by Modifier Type inferred from the Bayesian linear mixed model (Experiment 1).

gaze shift occurred before production onset was analysed to see whether N2 Contrast affects advance planning.

In 94% of the data the point of divergence was detected either before or after production onset. The time of divergence relative to production onset is shown in Table 3.2. The data were log transformed to account for positive skew. A Bayesian linear mixed model was fitted to assess whether Noun Contrast led to earlier gaze divergence. The model revealed negligible support (all $BF_{10} < 1$) for the main effect of Modifier Type ($\hat{\mu}$ = -0.03, 95% CrI[-0.23, 0.16], P($\beta < 0$) = .622), the main effect of Noun Contrast ($\hat{\mu}$ = 0.16, 95% CrI[-0.03, 0.35], P($\beta < 0$) = .054), and their interaction ($\hat{\mu}$ = 0.13, 95% CrI[-0.07, 0.34], P($\beta < 0$) = .101).

Table 3.2: Descriptive summary of the point when gaze shifted from N1 to N2 relative to production onset by Noun Contrast and by Modifier Type. Null is the production onset, positive values indicate gaze shift after production onset and negative values are gaze shift before production onset (Experiment 1)

Modifier	Contrast	*M*	*SE*	*N*
possessive	N1	577	230	194
	N2	149	239	203
postnominal	N1	239	179	202
	N2	50	150	212

Note: M = sample mean, SE = standard error, N = number of observations

Further, in 53% of the data divergence happened before production onset. These data are summarized by condition in Table 3.3. A logistic Bayesian mixed model was fitted on these proportions with the main effects and interaction of Modifier Type and Noun Contrast as predictors. The model revealed a main effect of Modifier Type ($\hat{\mu}$ = -0.87, 95% CrI[-1.55, -0.21], P($\beta < 0$) = .994, BF_{10} = 10) indicating a larger proportion of preonset divergence for postnominal modifiers, but negligible evidence ($\mathrm{BF}_{10} < 1$) for the main effect of Noun Contrast ($\hat{\mu}$ = -0.32, 95% CrI[-0.96, 0.31], P($\beta < 0$) = .852) and its by-Modifier Type interaction ($\hat{\mu}$ = 0.26, 95% CrI[-0.39, 0.9], P($\beta < 0$) = 0.22).

Table 3.3: Proportion of trials in which gaze shift from N1 to N2 was before production onset (Experiment 1)

Modifier	Contrast	*M*	*SE*	*N*
possessive	N1	.48	.03	214
	N2	.49	.03	219
postnominal	N1	.54	.04	208
	N2	.59	.03	219

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

Further, the proportion of eye samples on each of the target sentence's referents was calculated across the time before production onset (see Table 3.4).

Table 3.4: Proportion of eye samples to AOIs before production onset by condition (Experiment 1)

		AOI: N1			AOI: N2			
Modifier Type	Noun Contrast	*M*	*Mdn*	*IQR*	*M*	*Mdn*	*IQR*	*N*
possessive	N1	0.44	0.46	0.32	0.43	0.40	0.29	214
	N2	0.38	0.39	0.30	0.52	0.50	0.33	219
postnominal	N1	0.68	0.69	0.32	0.20	0.19	0.33	208
	N2	0.59	0.60	0.32	0.24	0.25	0.38	219

Note: M = sample mean, Mdn = sample median, IQR = interquartile range, N = number of observations

To assess contrast-related changes in the gaze pattern, the proportion of eye samples to the referent of N2 was compared across condition. For statistical analysis the proportion of eye samples was converted to empirical logits (see Jaeger, 2008; Mirman et al., 2008). A Bayesian linear mixed effects model (4 chains, 4000 iterations) was fitted with main effects and interaction of Modifier Type and Noun Contrast. To account for the multi-modal distribution of the data, the models were specified with three mixture components, a combination of three normal distributions, varying the location of each intercept and its variance parameter. The model results can be found in Figure 3.5.

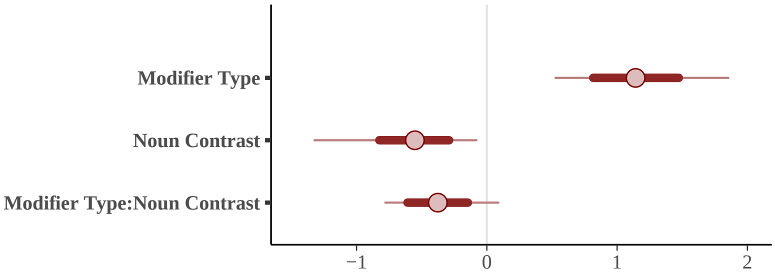


Figure 3.5: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the proportion of eye samples on the referent of the target sentences' second noun (Experiment 1). Dots indicate the posterior mean $\hat{\mu}$, thick lines show the range of 95% of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons *":"* denote interactions.

The analysis revealed a main effect of Modifier Type ($\mathrm{BF}_{10} > 4\mathrm{e}5$) showing a larger proportion of eye samples on N2 for possessives, and strong evidence for a Noun Contrast effect ($\mathrm{BF}_{10} = 199$) supporting a larger proportion of eye samples on N2 if N2 was contrastive. The Noun

Contrast effect varied by Modifier Type as indicated by the interaction (BF_{10} = 16). From the posterior samples of the model, nested Noun Contrast differences within each Modifier Type were calculated. These comparisons revealed strong evidence (BF_{10} = 1278) for a larger proportion of eye samples to the referent of N2, if N2 was contrastive, for possessives ($\hat{\mu}$ = -0.46, 95% CrI[-0.65, -0.27], P($\beta < 0$) > .999) but no such effect for postnominals ($\hat{\mu}$ = -0.09, 95% CrI[-0.26, 0.07], P($\beta < 0$) = .865, BF_{10} = 0.14). Figure 3.6 summarizes the Noun Contrast differences by Modifier Type as calculated from the posterior predicted values of the Bayesian model.

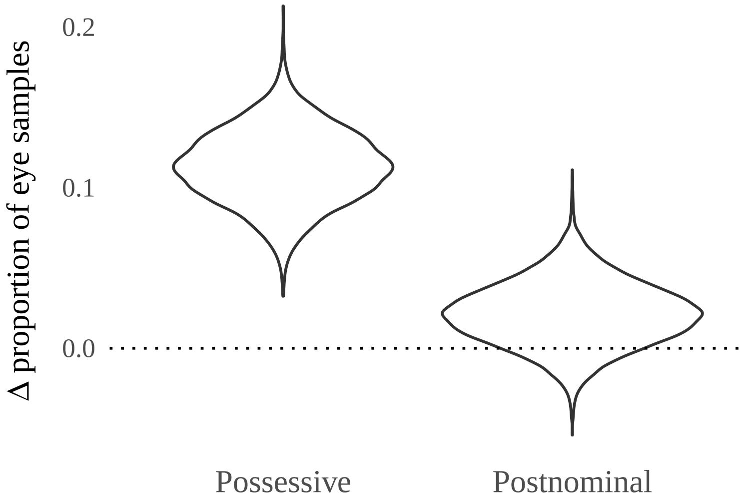


Figure 3.6: Summary of Noun Contrast difference compared to 0 for proportions of eye samples on N2 (Δ proportion = N2 Contrast–N1 Contrast) inferred from the Bayesian linear mixed model (Experiment 1).

3.2.3 Discussion

This study explored whether Noun Contrast leads to advance conceptual phrase planning. Conceptual contrast led to increased planning of non phrase-initial nouns. The data revealed that the proportion of eye samples on the N2 referent increased if N2 had a contrastive function. This effect was associated with tentatively longer production onset latencies. This supports that advance planning is affected by the conceptual structure of the phrase.

This contrast effect was found for possessives but not for postnominal modifier phrases. If this contrast effect were observed for postnominal modifiers, this would be evidence that advance conceptual planning is induced independently of the phrase's syntactic structure. However, increased planning dedicated to the N2 referent for N2 Contrast was seen in possessives only. This effect seems to be contingent on the apprehension of both noun referents. Conceptual planning might take into account the structural function of the phrase-initial noun (Brown-Schmidt & Konopka, 2015 but see E.-K. Lee et al., 2013). Head-initial phrases may allow to postpone planning of the modifier until after production onset. In contrast a phrase-initial modifier requires a head-noun as part of the conceptual plan. More trivially, advance planning of N2 might be due to the visual attraction induced by the colour of the target referent. During screen apprehension the first saccade will target the coloured image which is N2 in possessives but N1 in postnominal phrases. Either way, the Noun Contrast effect seems dependent on initial apprehension of both referents. Thus, the conceptual structure could only vary if more than one referent was conceptualised before production onset.

Problems with the present design were addressed in Experiment 2.

3.3 Experiment 2

Experiment 2 aimed to further examine whether conceptual contrast affects advance planning accounting for problems identified in Experiment 1. Experiment 1 showed strong evidence for N2 Contrast effects in the proportion of eye movements on the N2 referent but not in any of the other data. This effect might be explained by the presence of a conceptually identical comparator of the target image. If the target was in the phrase position N2 (i.e. in possessives) and had an identical comparator, participants looked to both images as they became aware of the potential ambiguity. Experiment 2 addressed this concern by removing the conceptually identical possessor noun and illustrated possession of more items by adding lines. In addition, conceptually identical images for the possessed objects were mirrored to make the identification of the referential ambiguity more difficult.

3.3.1 Method

3.3.1.1 Participants

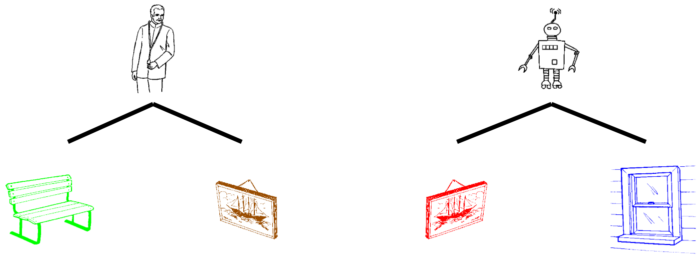
64 psychology students (17 male, mean age = 19.8 years, *SD* = 2.2, range: 18–30) participated in this experiment as part of a research-reward scheme. All participants were self-reported native speakers of British English, had no linguistic impairments and normal or corrected-to-normal vision. Four participants were replaced as they failed to provide sufficient utterances matching the structures described in the design section.

3.3.1.2 Design

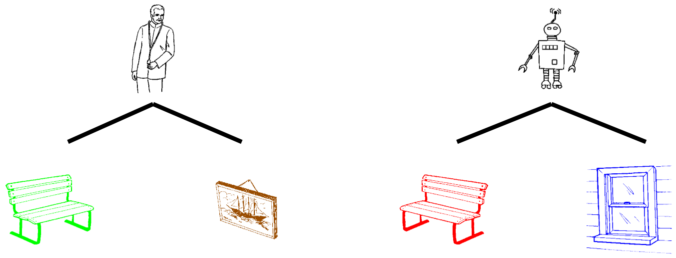
Similar to Experiment 1 the elicited modifier types were possessives and postnominals crossed with semantic Noun Contrast. These structures were produced in response to stimulus arrays as shown in Figure 3.7. An auditory presented trigger question (e.g. *What is brown?*) was presented. One group of participants was instructed to use prenominal possessive modifiers (e.g. *the man's picture is brown*) and a second group was instructed to use postnominal possessives (e.g. *the picture of the man is brown*). Importantly the Noun Contrast induced by the stimulus arrays varies by Modifier Type. The conceptually identical comparator of the coloured target image rendered N1 Contrast for possessives and N2 Contrast for postnominal phrases (e.g. Figure 3.7a). N2 Contrast for possessives and N1 Contrast for postnominals (Figure 3.7b) was created using an alternative possessee (i.e. the green bench) of the mutual possessor (i.e. the man) as indicated by two connecting lines rather than two conceptually identical images (see Experiment 1).

3.3.1.3 Material

32 items were created from the images used in the previous experiment. Items were divided into two Latin square lists one for each Noun Contrast. Modifier Types was alternated between-subjects. The location of the left and the right triplet of images and the location of the target image within the triplet were counterbalanced. The latter was varied such that maximally one image intruded the two conceptually identical images. Images that appeared twice were mirrored. Each of the possessed images was coloured in either green, brown, red or blue. The target colour



(a) N1 Contrast for possessives; N2 Contrast for postnominals



(b) N2 Contrast for possessives; N1 Contrast for postnominals

Figure 3.7: Stimulus arrays (Experiment 2) for the trigger question *What image is brown?*

was manipulated between items. To avoid strategic responses each list of items was filled with 48 filler arrays targeting constructions that did not involve possession (e.g. *the left ring is blue*). Stimuli and fillers were presented in random order. The stimulus material is presented in Appendix B.3.

3.3.1.4 Procedure & apparatus

Every experiment started with a camera set-up, calibration and validation. Participants were familiarised with the intended Modifier Type and ambiguous stimulus arrays including and excluding possession during instructions and 12 practice items. The experimenter monitored the produced utterances during the practice phase and corrected the structures indirectly if necessary. Recalibration was performed before the experimenter left lab. Each trial started with a centred fixation trigger that activated the image array. A recalibration was performed if the fixation trigger failed. The images, size 150 × 150 pixels, appeared in upper two-third of the centre of the screen and a text box (896 × 50 pixels) was provided at the bottom third were the produced text appeared. The participant heard the trigger question via a Logitec headset on an ASIO audio driver supported by a Creative SB X-Fi sound card. The time between stimulus onset and question varied as a function of the number of images (400 ms per image, thus 2,400 ms for conditional items). All images were shown on the screen until the participant ended the trial by pressing return. Pauses were possible after each trial. The duration of the experiment was approximately 35 minutes. The equipment used was the same as in Experiment 1.

3.3.2 Results

Prior to analysis data points were removed that exhibited extremely long pausing during production or structural rephrasing after response onset (11.14%). Moreover, sentences that did not match the expected structure were excluded from the analysis (0.98%). Trials with stimulus-to-onset latencies $<$ 100 ms (0.78%) and $>$ 10 secs (0.2%) and production durations $>$ 40 secs (0.15%) were discarded. In total, 12.6% of the data were removed. For the analysis of eye data further 2.5% were removed due to a proportion of eye samples larger than .75 outside of AOIs. Statistical analysis followed the same methods as those described for Experiment 1.

3.3.2.1 Onset latency

The onset latency, i.e. the duration from the onset of colour word in the trigger question to response onset, is summarized in Table 3.5. For a visualisation of the data see Appendix B.4.

Table 3.5: Descriptive data of onset latency in ms (Experiment 2)

Modifier Type	Noun Contrast	*M*	*SE*	*N*
possessive	N1	1667	43	459
	N2	1668	42	467
postnominal	N1	1533	42	438
	N2	1533	50	425

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

The log onset latency was fitted in a Bayesian linear mixed effects model with main effects of Modifier Type, Noun Contrast and their interaction. Random by-subject slopes for Modifier Type and the Noun Contrast by Modifier Type interaction were not included as subjects contributed data only one modifier level. The model output is shown in Figure 3.8. There was no substantive evidence for any model predictor ($\text{BF}_{10} < 1$).

3.3.2.2 Eye data

The analysis of the eye data followed the description in Experiment 1. First the time point and proportion of preonset gaze shift from N1 to N2 was determined. In 88% of the data this gaze shift was detected

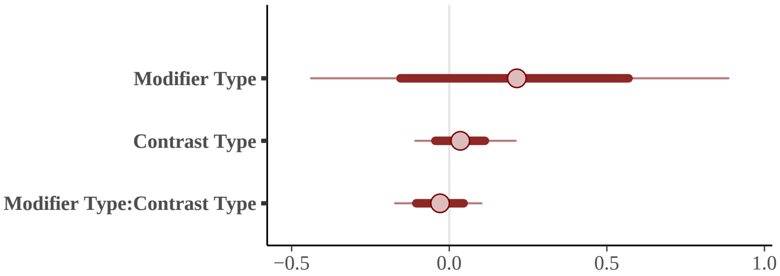


Figure 3.8: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the onset latency (Experiment 2). Dots indicate the posterior mean $\hat{\mu}$, the thick lines show the range of 95% of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons *":"* denote interactions.

either before or after production onset. The time of gaze shift relative to production onset can found in Table 3.6. A Bayesian linear mixed model was fitting including main effects and interaction of Modifier Type and Noun Contrast. There was negligible evidence (all $\text{BF}_{10} < 1$) for the main effect of Noun Contrast ($\hat{\mu} = 0$, 95% CrI[-0.04, 0.05], $P(\beta < 0) = .478$), the main effect of Modifier Type ($\hat{\mu} =$ -.05, 95% CrI[-0.16, 0.06], $P(\beta < 0) = .821$), and the interaction ($\hat{\mu} = 0.02$, 95% CrI[-0.02, 0.07], $P(\beta < 0) = .176$).

Table 3.6: Descriptive summary of the point of gaze shift from N1 to N2 relative to production onset. Null is the production onset, positive values indicate gaze shift after production onset and negative values are gaze shift before production onset (Experiment 2)

Modifier	Contrast	*M*	*SE*	*N*
possessive	N1	1535	164	379
	N2	1367	146	365
postnominal	N1	1544	108	399
	N2	1684	120	387

Note: M = sample mean, SE = standard error, N = number of observations

Further, in 26% of the data this gaze shift occurred before production onset. These data are summarized by condition in Table 3.7. A logistic Bayesian mixed model was fitted with main effects and interaction of Noun Contrast and Modifier Type. The model revealed negligible evidence ($\text{BF}_{10} < 1$) for the main effect of Noun Contrast ($\hat{\mu}$ = 0.28, 95% CrI[-0.27, 0.85], P($\beta < 0$) = .17) and the by-Modifier Type interaction ($\hat{\mu}$ = -0.2, 95% CrI[-0.76, 0.34], P($\beta < 0$) = .768). The main effect of Modifier Type ($\hat{\mu} = 1.7$, 95% CrI[0.38, 3.07], P($\beta < 0$) = .006, $\text{BF}_{10} = 17$) indicated a larger probability of gaze shift to N2 before production onset appearing in possessives.

Table 3.7: Proportion of trials in which the point of divergence of looks from N1 to N2 was before production onset (Experiment 2)

Modifier	Contrast	*M*	*SE*	*N*
possessive	N1	.30	.02	449
	N2	.29	.02	458
postnominal	N1	.22	.02	425
	N2	.20	.02	412

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

The proportion of eye samples on AOIs were aggregated across time from onset of the colour word to production onset. AOIs are the image corresponding to the head and the modifier noun. The head noun referent is N1 in postnominals and N2 in possessives. The modifier referent is N2 in postnominals and N1 in possessives. These proportions are summarized in Table 3.8.

Table 3.8: Proportion of eye samples to AOI before production onset by condition (Experiment 2)

		AOI: modifier			AOI: head			
Modifier Type	Noun Contrast	*M*	*Mdn*	*IQR*	*M*	*Mdn*	*IQR*	*N*
possessive	N1	0.38	0.39	0.30	0.32	0.30	0.29	449
	N2	0.37	0.37	0.35	0.38	0.34	0.29	458
postnominal	N1	0.08	0	0.13	0.59	0.58	0.30	425
	N2	0.08	0	0.10	0.58	0.57	0.28	412

Note: AOI: modifier = referent of N1 in possessives and N2 in postnominals (e.g. *man*), AOI: head = referent of N2 in possessives and N1 in postnominals (e.g. *painting*). M = sample mean, Mdn = sample median, IQR = interquartile range, N = number of observations

For inferential analysis, these data were converted to empirical logits. To account for the multi-modal distribution of the data, Bayesian linear mixed models were fitted with three mixture components (4 chains, 6000 iterations). Predictors were main effects and interactions of Modifier Type, Noun Contrast, and AOI. The results are shown in Figure 3.9.

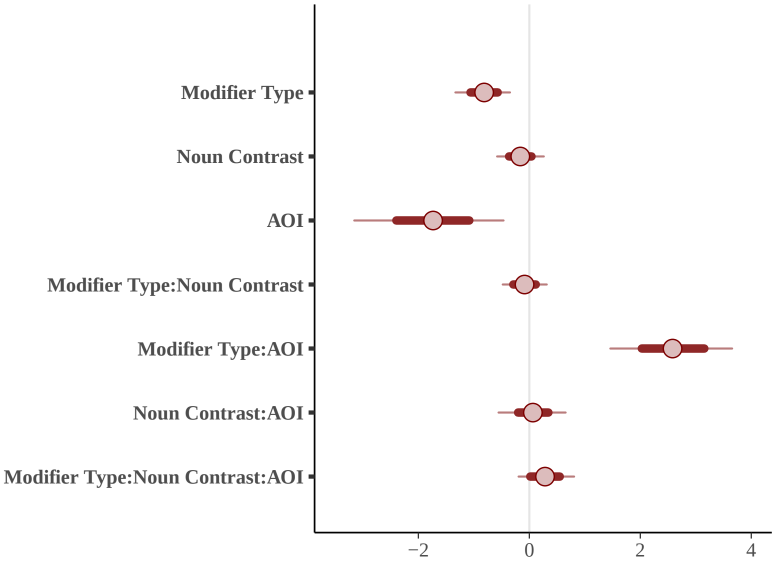


Figure 3.9: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the proportion of eye samples to each referent of the target sentence (Experiment 2). Dots indicate the posterior mean $\hat{\mu}$, thick lines show the range of 95% of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons ":" denote interactions.

The model revealed negligible support for the main effect of Noun Contrast and for by-Noun Contrast interactions ($\mathrm{BF}_{10} < 1$). Simple effects of Noun Contrast were assessed within Modifier Type and AOI revealing larger proportions of eye samples to the referent of N2 for N2 Contrast in possessives ($\hat{\mu}$ = -0.15, 95% CrI[-0.27, -0.02], $P(\beta < 0) >$

.999) but the evidence for this difference was negligible ($BF_{10} = 1$). The Noun Contrast comparisons revealed no difference for N2 in postnominals ($\hat{\mu}$ = -0.07, 95% CrI[-0.17, 0.03], $P(\beta < 0) = .927$, $BF_{10} < 1$). For N1 no difference ($BF_{10} < 1$) was found in possessives ($\hat{\mu} = 0.02$, 95% CrI[-0.09, 0.14], $P(\beta < 0) = .351$) and postnominals ($\hat{\mu} = 0.04$, 95% CrI[-0.09, 0.16], $P(\beta < 0) = .291$). Figure 3.10 summarizes the Noun Contrast differences calculated from the posterior samples of the Bayesian model.

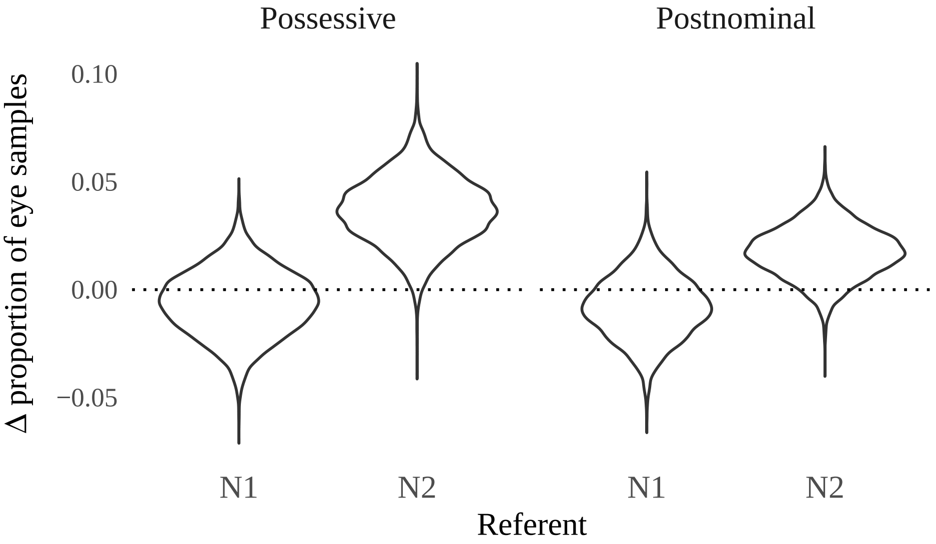


Figure 3.10: Summary of Noun Contrast effect for proportion of eye samples on each noun's referent inferred from the Bayesian linear mixed model. The y-axis shows the Noun Contrast difference (Δ proportion = N2 Contrast–N1 Contrast) with 0 indicated by a dotted line (Experiment 2). *Referent* refers to the appearance of the image's name in the target phrase and not the actual AOI; i.e. N1 in possessive phrases is the same AOI as N2 in postnominal phrases and *vice versa.*

Further, the model provided strong support for main effect of Modifier Type ($BF_{10} > 2e5$), AOI ($BF_{10} = 7333$) and the Modifier Type by-AOI interaction ($BF_{10} > 3e11$). Nested comparisons within AOI were used to inspect this interaction. We found strong evidence for larger proportions of eye samples to N1 in possessives compared to postnominal phrases ($\hat{\mu} = 0.88$, 95% CrI[0.58, 1.19], $P(\beta < 0) < .001$, $BF_{10} > 6e4$) and smaller proportions of eye samples to N2 in possessives compared to postnominal phrases ($\hat{\mu} = -1.7$, 95% CrI[-2.01, -1.39], $P(\beta < 0) > .999$, $BF_{10} > 4e9$).

3.3.3 Discussion

The aim of this experiment was to provide evidence for conceptual structure determining advance phrase planning after accounting for possible problems encountered in Experiment 1. However, this experiment failed to provide conclusive evidence.

Experiment 1 revealed larger proportions of eye samples to N2 in possessives. There was some indication for this effect in Experiment 2 although it was not substantial. One possibility is that this effect could be explained by the mere presence of a conceptually identical comparator. The absence of Noun Contrast differences in the eye data on N1 for postnominals ($BF_{10} < 1$) suggests that this is not necessarily the case as one would have expected a similar pattern as observed for N2 in possessives.

As descriptions in Experiment 1 and 2 were planned in the absence of an interlocutor, conceptual structuring might have been by-passed in a majority of the trial. There was no need to encode a conceptual contrast in the absence of a decoder. This concern was addressed in Experiment 3.

3.4 Experiment 3

Experiment 3 aims to further investigate whether advance conceptual structuring subserves advance phrase planning. Experiment 1 and 2 showed little or no converging support that conceptual contrast affects the advance planning process in noun phrase modification. This lack of evidence might be due to the experimental set-up used – notably the output modality that was employed in the production task or the lack of communicative interactivity.

The output modality used in Experiment 1 and 2 might have led to the absence of noun contrast effects for two reasons. First, writers have to plan just enough information to onset the sentence while more structural planning can be postponed until after production onset. Spoken utterances require enough advance processing to include new incoming information into the utterance after production onset without interruptions of the speech stream (Brown-Schmidt & Konopka, 2015). Further speech requires minimising the extent of preplanned information and avoidance of intra-sentential pausing (Levelt & Meyer, 2000; Meyer, 1997) by, for example, assessing the possibility of parallel planning subsequently to the production onset (Griffin, 2003). Hesitations in writing, however, have no implication on the text's communicative effect. Second, conceptual contrast in English is typically encoded by virtue of prosodic prominence (e.g. *the man's BEARD*). Although implicit prosody – "inner speech" – is known to influence comprehension in silent reading (e.g. McCurdy, Kentner, & Vasishth, 2013; Thomson & Jarmulowicz, 2016; Wade-Woolley & Heggie, 2015), it is entirely unknown whether prosody is planned in the production of written text. Assuming there is no stage at which writers plan the prosody of their "inner speech", the preparation of contrast (i.e. noun stress) might have been dismissed altogether. To account for the possibility that output modality resulted in the failure to observe noun contrast effects in the previous experiments, both written and spoken descriptions were recorded in Experiment 3.

Another explanation for the absence of noun contrast effects is the following. Broadly speaking there was no communicative need to plan (and produce) a different semantic structure for N1 and N2 Contrast in Experiment 1 and 2. As there was no addressee present participants may have strategically produced phrases as schemas without being aware of conceptual contrasts. Findings from spoken discourse show that people tend to produce fewer precise referential phrases in the absence of an interlocutor (Van Der Wege, 2009). Language (i.e. speech) is typically used in communication although most of sentence planning research elicited utterances in the absence of an interlocutor (e.g. Allum & Wheeldon, 2007; Martin et al., 2010, 2014; Meyer, 1996; Smith & Wheeldon, 1999, 2004; Wagner et al., 2010). Experiment 3 used an interactive image description task (e.g. Brown-Schmidt & Konopka, 2008, 2015; Brown-Schmidt & Tanenhaus, 2006; E.-K. Lee et al., 2013; Swets et al., 2013, 2014) in combination with a prime-target design to tease out the contrast structure of the noun phrase.

3.4.1 Method

3.4.1.1 Participants

64 psychology students (7 male, mean age = 20.4 years, *SD* = 5.4, range: 18–50) participated as part of a research-reward scheme. All participants were self-reported native speakers of British English, free of linguistic impairments, and had normal or corrected-to-normal vision. Five participants were replaced because of difficulty to follow the task or because of problems to record their eye movements.

3.4.1.2 Design

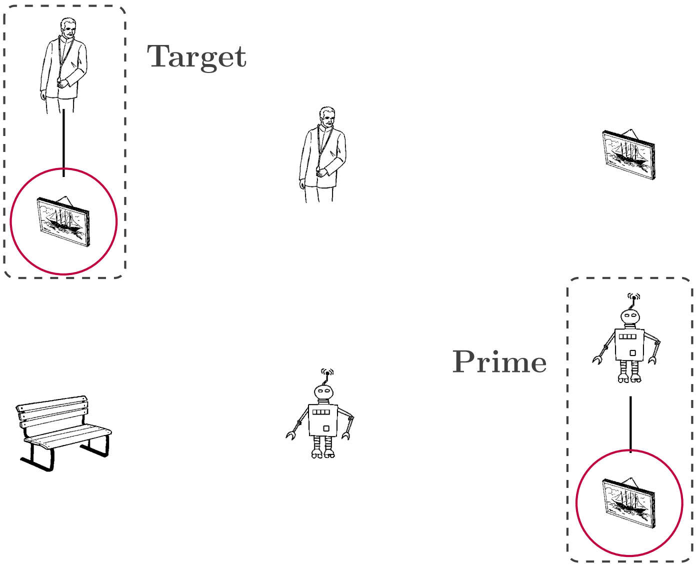
Participants were presented with arrays of images containing two pairs connected by a vertical line and four distractors (Figure 3.11). Distractors were included to require modification for all target images. The lower image of each image pair was highlighted one at a time. Participants had to instruct the experimenter to click on the highlighted image using possessives or postnominal structures. The Noun Contrast manipulation is similar to the previous experiments and depends on the noun-modifier structure used by the participant. Figure 3.11a requires N2 Contrast for postnominal phrases (*the painting with the MAN*) but N1 Contrast for possessives (*the MAN's painting*). Figure 3.11b, on the other hand, requires N1 Contrast for postnominal phrases (*the PAINTING with the man*) and N2 Contrast for possessives (*the man's PAINTING*). Noun Contrast was stressed using a prime-target design: the noun referent of the comparator image pair (henceforth, the prime) was targeted before the referent of the target pair. For example, in Figure 3.11a the painting bottom-right was highlighted first and then the painting top-left. Thus, two descriptions in the same Noun Contrast condition were elicited per screen. Each participant completed a written and a spoken session. This manipulation rendered a 2 × 2 × 2 × 2 design.

3.4.1.3 Material

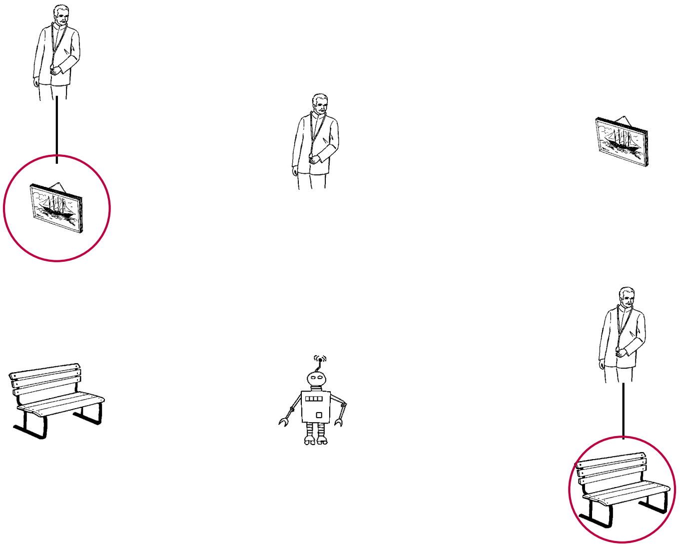
Each stimulus screen comprised six cells. Two pairs of the 32 target image pairs and their respective comparators from Experiment 2 were included (see Appendix B.3). All unique four images of each item were added to the remaining cells. Items were counterbalanced over four Latin Square lists (Noun Contrast, Modality Type). Order of modality session was counterbalanced between subjects. Two filler lists containing 16 trials each were created. One filler list was used in the first, the other in the second session. Filler lists and stimuli were presented randomised within session. Fillers contained arrays with targets different form the stimuli to avoid strategic use of descriptions and anticipation of the target. For example, in prime and target trials images different from the bottom image of pairs were prompted, and colour (in combination with modifiers) had to be used to disambiguate the target image (e.g. *the cat's green ball*). The location of prime and target image was randomised within each Latin square list. In total, each subject saw 64 arrays of images, 32 per modality with 50% fillers.

3.4.1.4 Procedure

Every experiment started with a camera set-up, calibration and validation. Participant and experimenter were seated on different screens



(a) N1 Contrast for possessives; N2 Contrast for postnominals



(b) N2 Contrast for possessives; N1 Contrast for postnominals

Figure 3.11: Stimulus arrays (Experiment 3). Frames highlight the prime and target group. The circle in the target group (man and painting) was displayed after the subject finished responding to the prime trial and feedback was provided.

unable to see the each other's screen. The participant's task was to unambiguously instruct the experimenter to click on the image that is highlighted by a circle. A purple circle highlighting the target image appeared simultaneously to trial onset. The participants were told to press enter at the end of every description (i.e. to send the written message to the experimenter's screen). This activated the mouse input which permitted the experimenter to click on the image according to the participant's instruction. Mouse clicks on the correct image prompted a green circle as feedback. If the description was ambiguous, the experimenter clicked on a comparator image and a red circled appeared. Feedback circles disappeared after 250 ms. The second target image was highlighted immediately after feedback disappeared. The mouse cursor was only visible after the participant pressed enter. Participants were familiarised with the experimental task during a practice session with ten trials. A recalibration was performed before the experiment started. Each trial began with a centred fixation trigger that activated the image array. Recalibration was performed if the fixation trigger failed to prompt the next screen. The images, size 150 × 150 pixels, appeared around the centre of the screen. In the written session, a text box (896 × 50 pixels) was provided in the middle of the screen. Pauses were possible after each trial. The duration of the experiment was approximately 45 minutes.

3.4.1.5 Apparatus

The apparatus was similar to the previous experiments. Spoken responses were recorded with a Logitec headset on ASIO audio driver supported by the Creative SB X-Fi sound card.

3.4.2 Results

Prior to analysis trials with the following properties had to be removed: Extremely long pausing, ambiguous descriptions and self-corrections of the phrase structure were removed (9.9%). Trials in which either or both the prime and target response are ambiguous were excluded such as prime-target pairs that did not contain the same modifier type (5.6%). Moreover, trial with long onset latencies > 10,000 ms were removed (0.3%). For the analysis of eye data further 2.4% of the data were removed due to a proportion of eye samples larger than .75 outside of AOIs. Statistical analysis followed the same methods as those described for Experiment 1.

3.4.2.1 Modifier choice

Participants were allowed to chose noun phrase modification freely. Responses were divided into prenomial/possessive modifiers and postnominal modifiers (e.g. prepositional phrase, relative clause). The proportion of postnominal modifiers produced is summarized in Table 3.9.

A Bayesian generalized linear mixed model with binomial link function was fitted on choice of Modifier Type as dependent variable with main effects and interaction of Noun Contrast and modality to assess whether modifier choice was independent of Noun Contrast. The model revealed negligible evidence for the main effect of Noun Contrast ($\hat{\mu}$ = -0.69, 95% CrI[-3, 1.35], P($\beta < 0$) = .74, BF_{10} = 1.2) and the Noun Contrast by modality interaction ($\hat{\mu}$ = -0.17, 95% CrI[-2.11, 1.65], P($\beta < 0$) = .578, $\text{BF}_{10} < 1$). The model revealed evidence for the main effect of modality showing larger proportions of postnominal phrases for speech ($\hat{\mu} = 9.33$, 95% CrI[1.87, 16.9], $P(\beta < 0) = .005$, $BF_{10} = 101$).

Table 3.9: Descriptive summary of proportion of postnominal modifier structures – as opposed to possessive modifier structures – produced by Modality and Noun Contrast condition (Experiment 3)

Modality	Noun Contrast	*M*	*SE*	*N*
speech	N1	.69	.01	924
	N2	.71	.01	956
writing	N1	.55	.02	827
	N2	.58	.02	863

Note: *M* = sample mean, *SE* = standard error, *N* = number of observations

3.4.2.2 Onset latency

The onset latency, the duration from appearance of the circle highlighting the target to production onset, is summarized by condition in Table 3.10. For a visualisation see Appendix B.5.

Table 3.10: Descriptive summary of onset latency (in ms) for prime and target responses by condition (Experiment 3)

			Prime			Target		
Modality	Modifier Type	Noun Contrast	*M*	*SE*	*N*	*M*	*SE*	*N*
speech	possessive	N1	1638	70	143	1187	36	143
		N2	1608	77	137	1147	43	137
	postnominal	N1	1733	49	319	1258	30	319
		N2	1646	40	341	1037	19	341
writing	possessive	N1	2559	93	173	1713	52	198
		N2	2666	98	174	1712	55	187
	postnominal	N1	2315	86	214	1776	50	242
		N2	2093	68	244	1522	45	258

Note: M = sample mean, SE = standard error, N = number of observations

The onset latency was log transformed to correct for positive skew and analysed in a Bayesian linear mixed model with main effects of modality, prime, Modifier Type, Noun Contrast and their interactions. The model output is summarized in Figure 3.12. The analysis revealed strong evidence for a main effect of Noun Contrast ($\text{BF}_{10} = 68$) showing longer latencies for N1 Contrast and the Modifier Type by Noun Contrast interaction ($\text{BF}_{10} = 3.5$). Pairwise comparisons within Modifier Type calculated from the model's posterior samples of the Bayesian model revealed strong evidence ($\text{BF}_{10} = 1563$) indicating a Noun Contrast effect in postnominal phrases ($\hat{\mu} = 0.45$, 95% CrI[0.27, 0.62], $P(\beta < 0) < .001$) but not in possessives ($\hat{\mu} = 0.09$, 95% CrI[-0.13, 0.31], $P(\beta < 0) = .218$, $\text{BF}_{10} = 0.15$). These differences as derived from the posterior samples of the Bayesian model are visualised in Figure 3.13.

Further, there was strong support for a modality by prime by modifier interaction ($\text{BF}_{10} = 63$). This interaction showed varying modifier type differences within prime/target and modality. In the written data there was strong evidence ($\text{BF}_{10} = 187$) showing longer latencies for possessives in prime trials ($\hat{\mu} = 0.44$, 95% CrI[0.22, 0.65], $P(\beta < 0) < .001$) but no modifier difference in target trials ($\hat{\mu} = 0.18$, 95% CrI[-0.02, 0.36], $P(\beta < 0) = .034$, $\text{BF}_{10} = 0.5$). There was no Modifier Type difference in speech, neither in prime ($\hat{\mu} = -0.22$, 95% CrI[-0.47, 0.02], $P(\beta < 0) = .962$, $\text{BF}_{10} = 0.7$) nor in target trials ($\hat{\mu} = -0.02$, 95% CrI[-0.22, 0.18], $P(\beta < 0) = .569$, $\text{BF}_{10} = 0.1$).

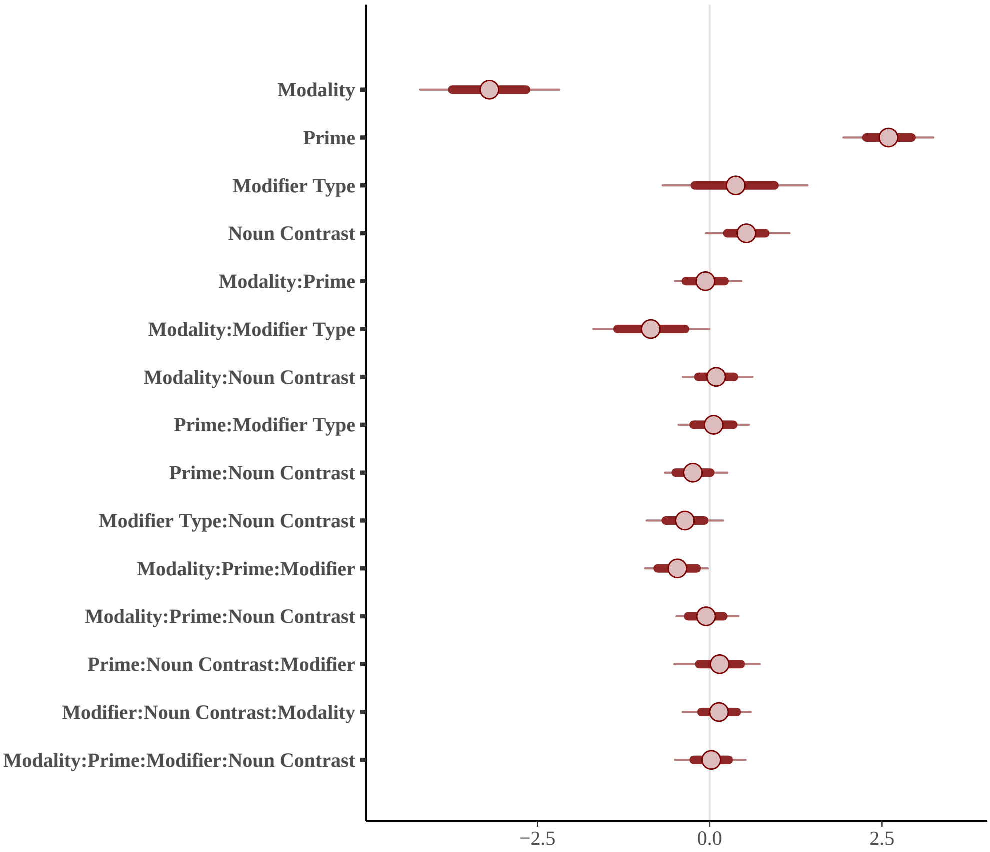


Figure 3.12: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the onset latency (Experiment 3). Dots indicate the posterior mean $\hat{\mu}$, the thick lines show the 95% range of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons *":"* denote interactions.

3.4.2.3 Eye data

The proportion of eye samples to N1 and N2 from the onset of the presentation of the circle to production onset is illustrated in Figure 3.14. The displayed figures for prime and target trials and each modality illustrate the divergence of AOI looks (see e.g. Griffin & Bock, 2000; Konopka

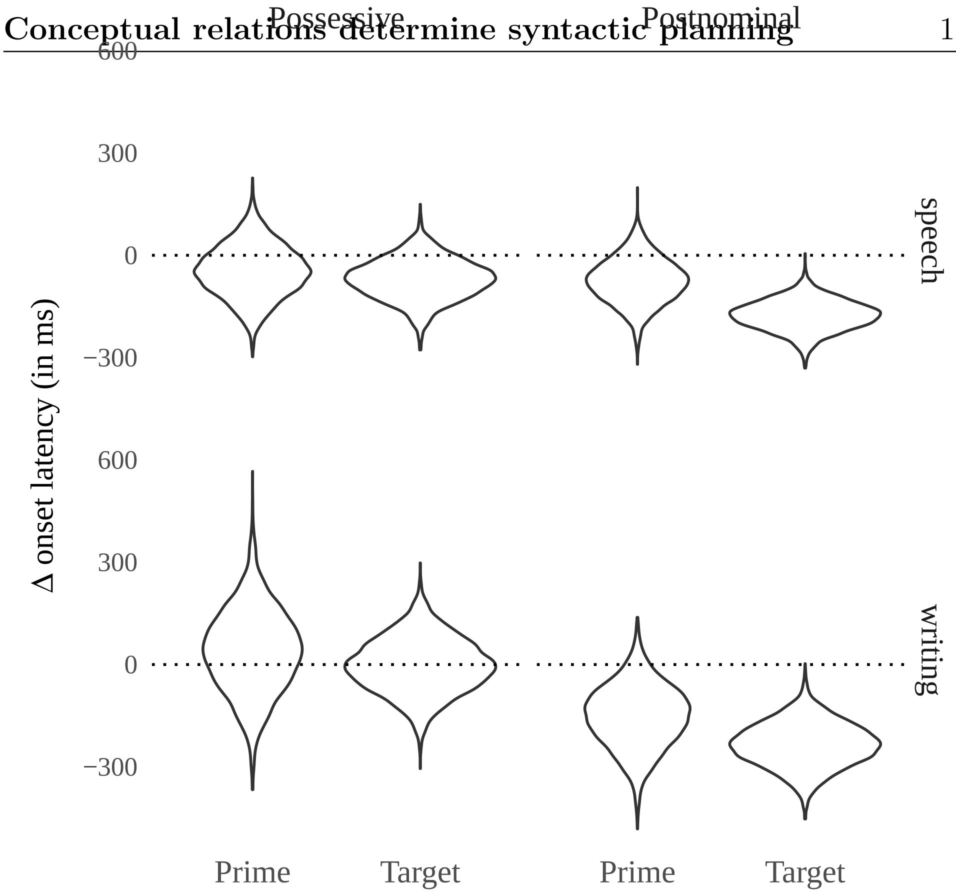
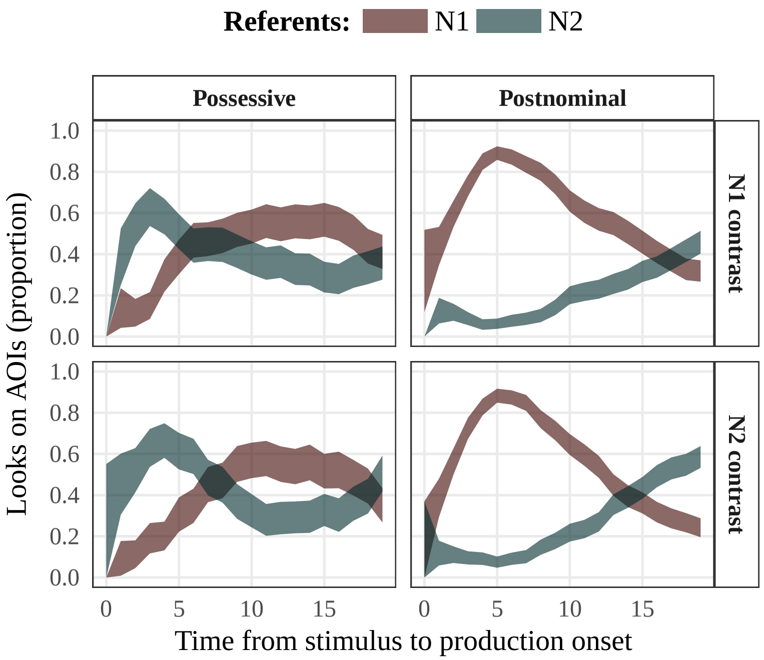
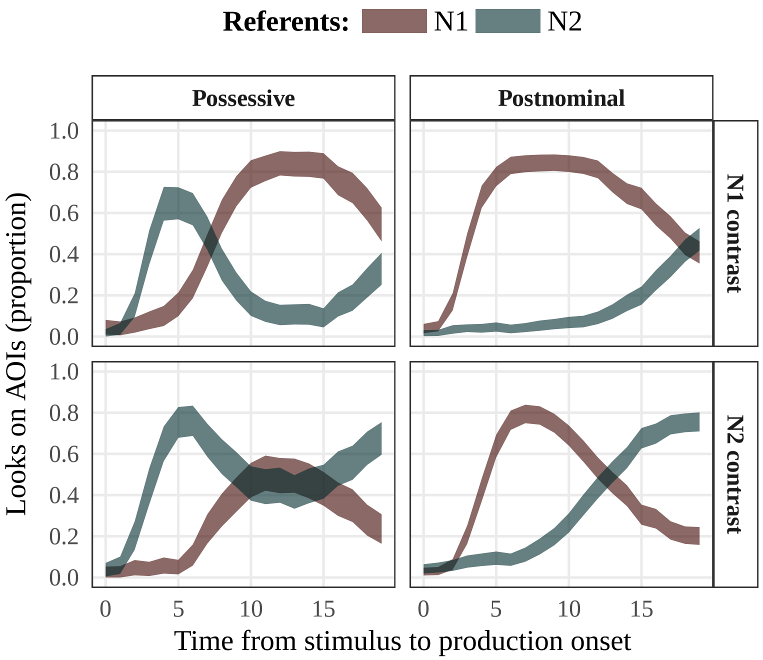


Figure 3.13: Summary of Noun Contrast effects on onset latency inferred from the Bayesian linear mixed model. The Noun Contrast difference (Δ onset latency = N2 Contrast–N1 Contrast) is shown on the y-axis (Experiment 3). Zero indicates no difference as indicated by the dotted line.

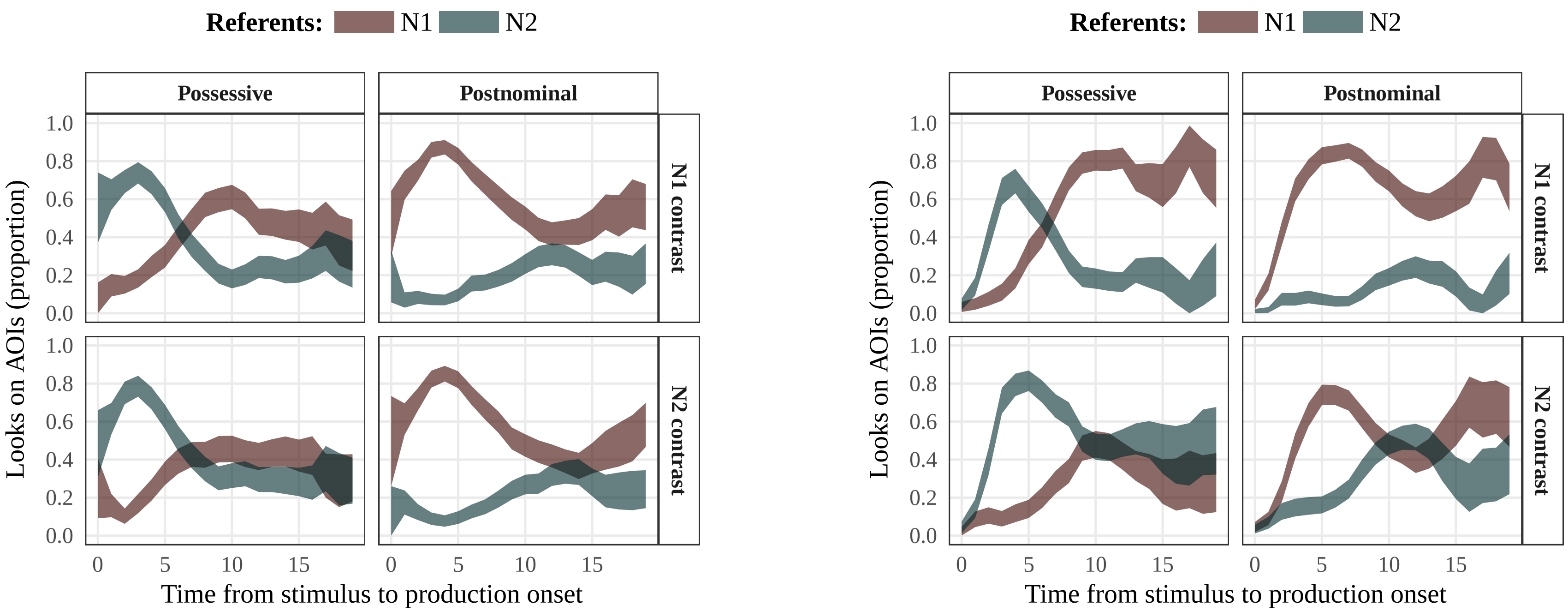
& Meyer, 2014). Figure 3.14a and Figure 3.14b show that, if the target referent was N1 – as for postnominal phrases – gaze divergence happened before production onset only if N2 was contrastive, not if N1 was contrastive. Further, this can be seen in speech for both prime and target trials but not as clearly in writing (see Figure 3.14c and Figure 3.14d).



(a) Speech, prime trial



(b) Speech, target trial



(c) Writing, prime trial

(d) Writing, target trial

Figure 3.14: Proportion of eye samples from stimulus onset to production onset. Referents are images in the order mentioned in phrase; possessives: *the N1's N2*, postnominals: *the N1 with the N2*. Time axis was normalised by trial and binned. Bands indicate 95% confidence intervals (Experiment 3).

The time – relative the production onset – when gaze divergence from N1 to N2 occurred was calculated from the eye data such as the proportion of trials in which gaze shift happened before production onset (see Results section, Experiment 1). In 97% of the data divergence was detected either before or after production onset. The data showing the time at which gaze diverged from N1 to N2 are shown in Figure 3.15.

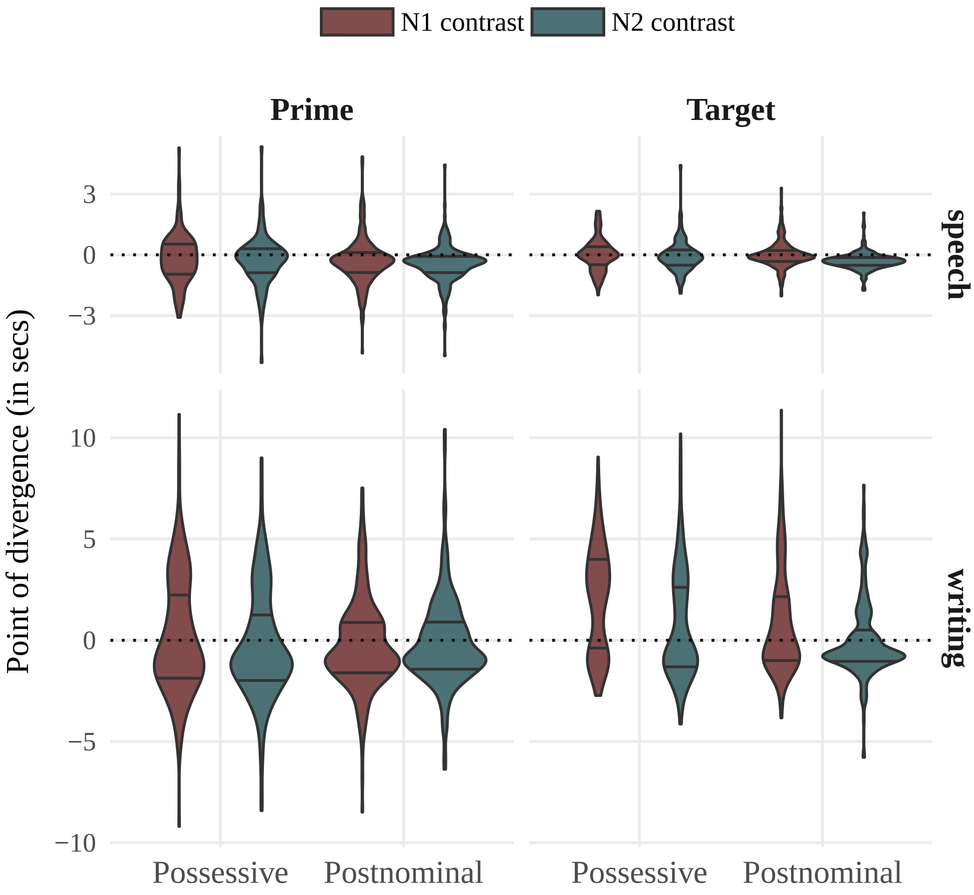


Figure 3.15: Bean plots of the point of gaze divergence from N1 to N2 relative to production onset. The beans illustrate the smoothed density of their distribution. The bands illustrate the concentration of the data between the first and third quantile. Zero signifies the production onset indicated by a dotted line (Experiment 3).

For statistical analysis the data were shifted above zero to allow log transformed correcting for positive skew. The data were analysed using Bayesian linear mixed models. Model predictors were main effects and interactions of modality, prime/target trial, Modifier Type and Noun Contrast. The model outcome can be found in Figure 3.16.

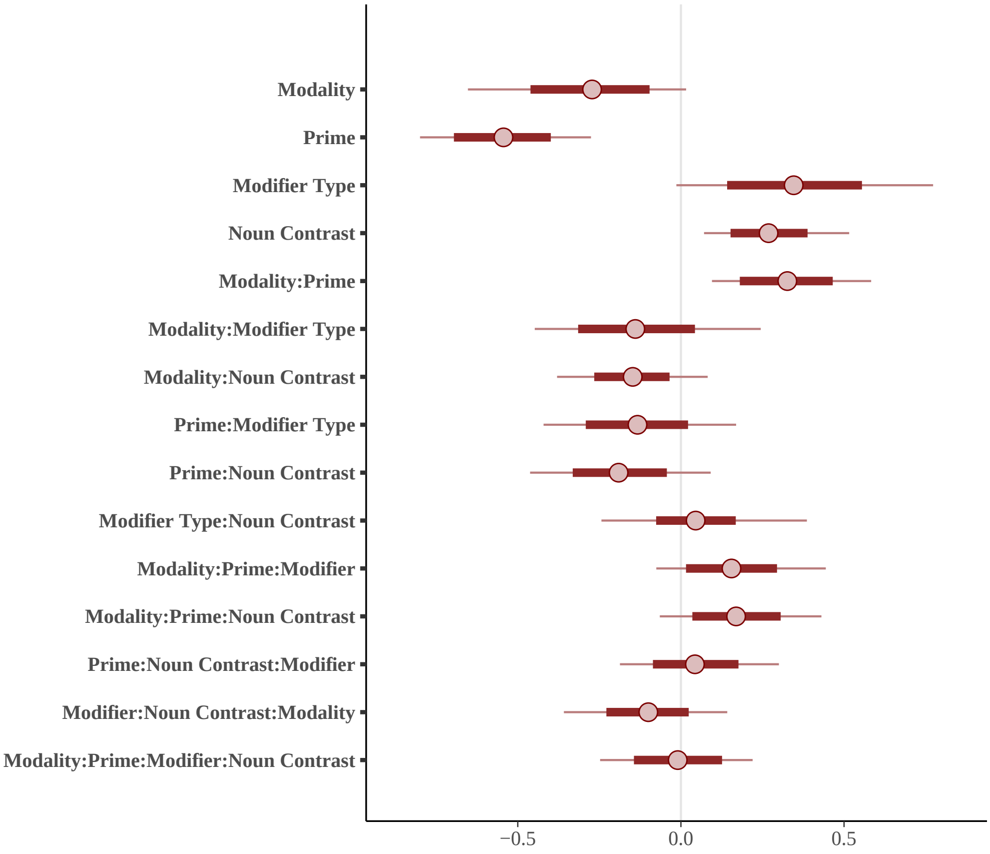


Figure 3.16: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the time of gaze divergence from N1 to N2 relative to production onset (Experiment 3). Dots indicate the posterior mean $\hat{\mu}$, the thick lines show the 95% range of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons ":" denote interactions.

We found strong evidence for a main effect of Noun Contrast ($BF_{10} = 228$) showing that N2 Contrast compared to N1 Contrast exhibited earlier gaze divergence from N1 to N2. There was negligible evidence for any of the by-Noun Contrast interactions ($BF_{10} < 2$). The Noun Contrast effects calculated from posterior samples of the Bayesian model are summarized in Figure 3.17.

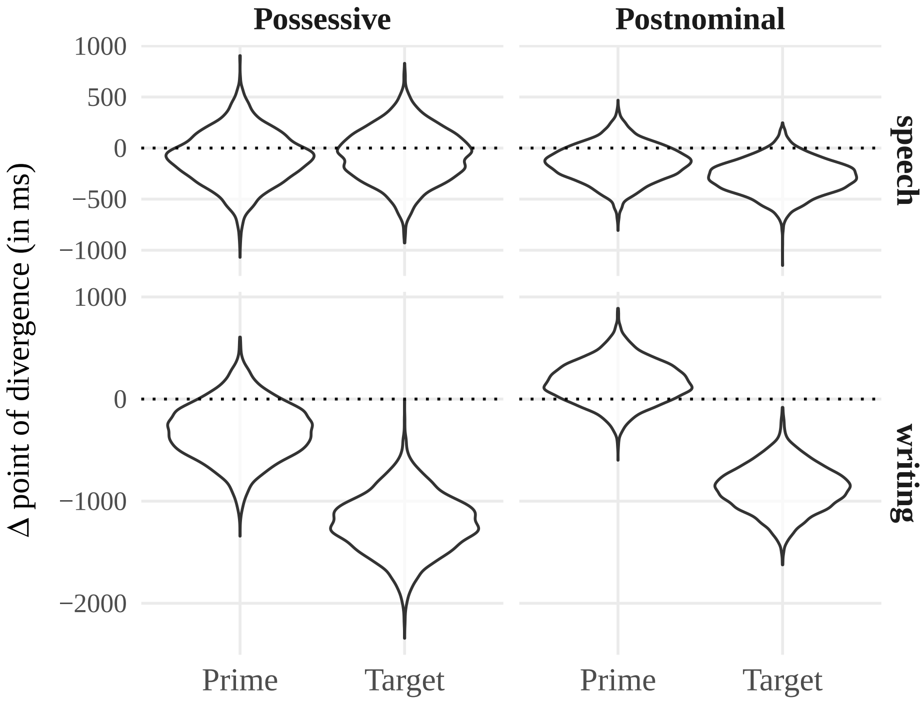


Figure 3.17: Summary of Noun Contrast effect for time of gaze divergence inferred from Bayesian linear mixed model. Noun Contrast effect (Δ point of divergence = N2 Contrast–N1 Contrast) is shown on the y-axis (Experiment 3). Null indicates no difference.

In 61% of these data gaze divergence happened before production onset. These data – the proportion of trials in which gaze shift to N2 occurred before production onset – are summarized by condition in Table 3.11.

Table 3.11: Proportion of trials in which the point of divergence for looks from N1 to N2 occurred before production onset (Experiment 3)

			Prime			Target		
Modality	Modifier Type	Contrast	*M*	*SE*	*N*	*M*	*SE*	*N*
speech	possessive	N1	.58	.04	139	.52	.04	143
		N2	.62	.04	130	.61	.04	136
	postnominal	N1	.70	.03	313	.59	.03	316
		N2	.80	.02	336	.84	.02	340
writing	possessive	N1	.56	.03	212	.30	.03	204
		N2	.63	.03	197	.52	.04	199
	postnominal	N1	.61	.03	270	.40	.03	276
		N2	.59	.03	281	.63	.03	270

Note: M = conditional mean, SE = standard error, N = number of observations

The proportion of pre-onset gaze shift was analysed in a logistic Bayesian linear mixed model. Model predictors were main effects and interactions of modality, prime/target trial, Modifier Type and Noun Contrast. The model outcome is summarized in Figure 3.18.

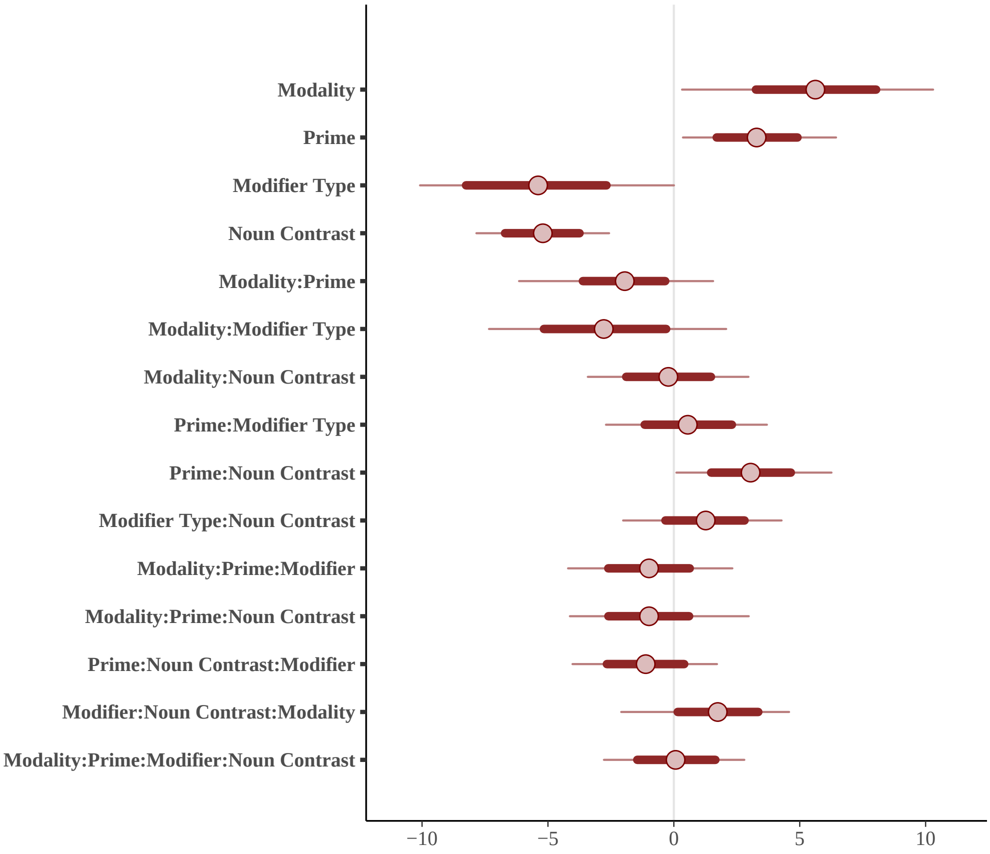


Figure 3.18: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the proportion of gaze divergence from N1 to N2 before production onset (Experiment 3). Dots indicate the posterior mean $\hat{\mu}$, the thick lines show the 95% range of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons ":" denote interactions.

There was strong evidence ($\text{BF}_{10} > 7\text{e}6$) for a main effect of Noun Contrast indicating that gaze divergence from N1 to N2 was more likely to occur before production onset for N2 Contrast. The Noun Contrast effect varied by Modifier Type and modality as suggested by the moderate support (BF_{10} = 8.5) for the three way interaction. Pairwise comparisons revealed strong evidence for a Noun Contrast effect for postnominal phrases in speech ($\hat{\mu}$ = -2.11, 95% CrI[-2.85, -1.38], P($\beta < 0$) > .999, BF_{10} = 17857) and writing ($\hat{\mu}$ = -1.13, 95% CrI[-1.85, -0.41], P($\beta < 0$) > .999, BF_{10} = 43), and for possessives in writing ($\hat{\mu}$ = -1.37, 95% CrI[-2.19, -0.58], P($\beta < 0$) > .999, BF_{10} = 120) but not in speech ($\hat{\mu}$ = -0.6, 95% CrI[-1.53, 0.34], P($\beta < 0$) = .898, BF_{10} = 1.1). Further, the Noun Contrast effect interacted with prime (BF_{10} = 550). Gaze divergence before production onset was more likely for N2 Contrasts in target trials ($\hat{\mu}$ = -4.13, 95% CrI[-5.2, -3.04], P($\beta < 0$) > .999, BF_{10} > 1e7) but the evidence for a Noun Contrast effect in prime trials was rather weak ($\hat{\mu}$ = -1.08, 95% CrI[-2.18, 0.02], P($\beta < 0$) = .97, BF_{10} = 3). The Noun Contrast effects calculated from posterior samples of the Bayesian model can be found summarized in Figure 3.19.

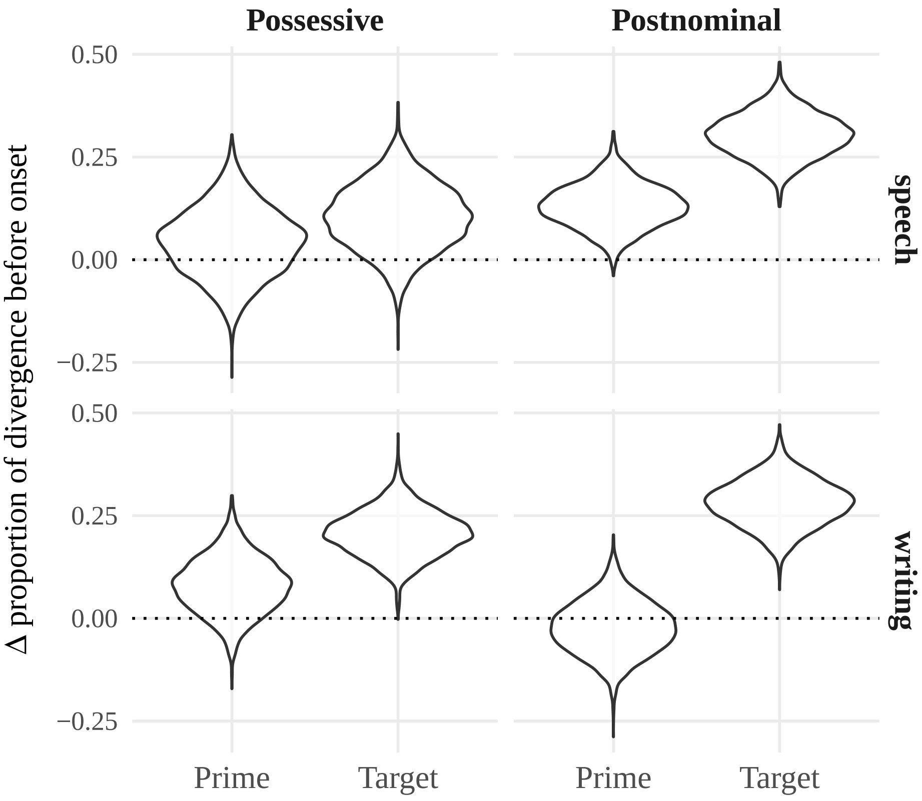


Figure 3.19: Summary of Noun Contrast effect on proportion of gaze divergence from N1 to N2 before production onset inferred from the Bayesian linear mixed model. The Noun Contrast difference (Δ proportion of pre-onset gaze shift = N2 Contrast–N1 Contrast) is shown on the y-axis (Experiment 3). The dotted line indicates a difference of 0.

Finally, the proportion of eye samples to the modifier and head noun referent was calculated across the time from stimulus (i.e. appearance of target marker) to production onset. These proportions are summarized by condition in Table 3.12.

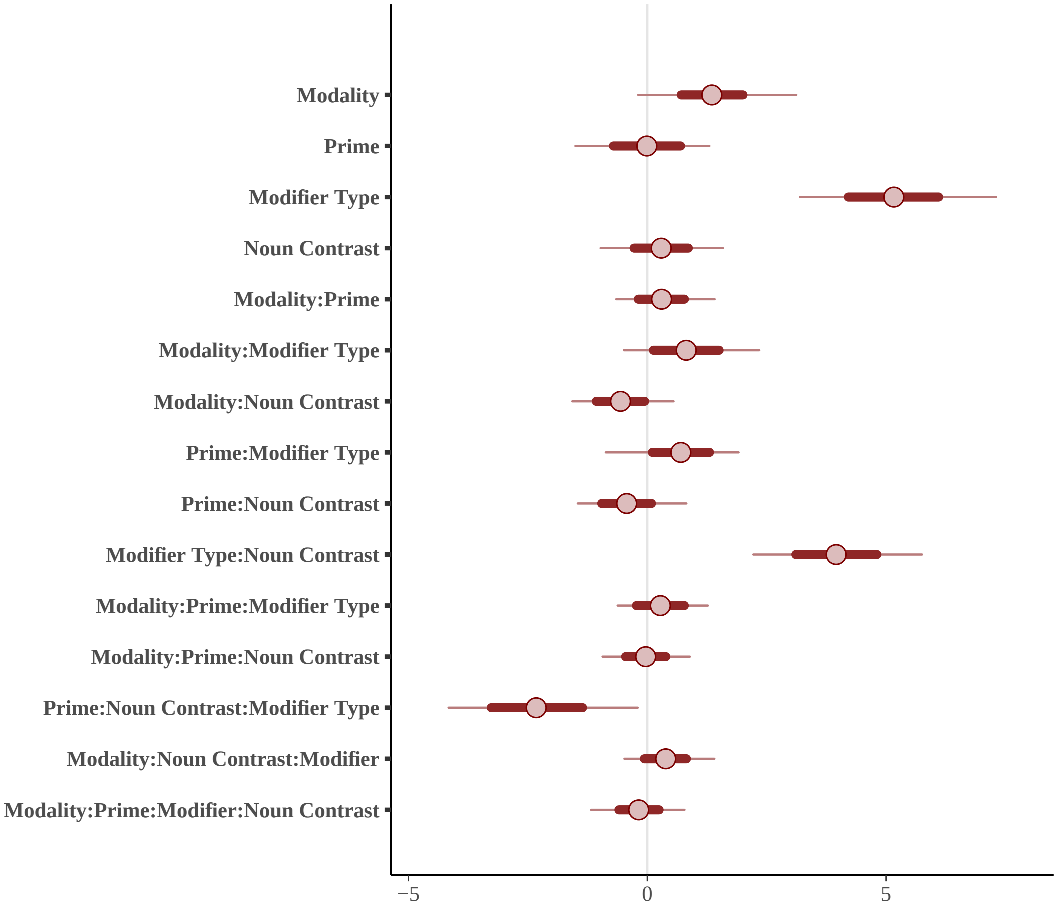
Table 3.12: Descriptive summary of aggregated proportion of eye samples to AOIs before production onset (Experiment 3)

				Prime				Target			
Modality	Modifier Type	AOI	Contrast	*M*	*Mdn*	*IQR*	*N*	*M*	*Mdn*	*IQR*	*N*
speech	possessive	modifier	N1	.41	.40	.26	139	.25	.25	.16	143
			N2	.44	.41	.27	130	.49	.50	.28	136
		head	N1	.46	.46	.30	139	.52	.54	.25	143
			N2	.44	.45	.27	130	.28	.27	.29	136
	postnominal	modifier	N1	.63	.60	.38	313	.63	.65	.29	316
			N2	.59	.56	.38	336	.42	.42	.23	340
		head	N1	.21	.19	.33	313	.13	.10	.24	316
			N2	.27	.27	.29	336	.36	.36	.27	340
writing	possessive	modifier	N1	.40	.38	.20	212	.32	.31	.19	204
			N2	.47	.42	.29	197	.54	.53	.24	199
		head	N1	.43	.43	.23	212	.50	.49	.28	204
			N2	.37	.35	.29	197	.26	.28	.20	199
	postnominal	modifier	N1	.66	.65	.37	270	.65	.67	.24	276
			N2	.64	.63	.38	281	.51	.50	.25	270
		head	N1	.18	.16	.30	270	.11	.02	.20	276
			N2	.20	.19	.32	281	.27	.27	.40	270

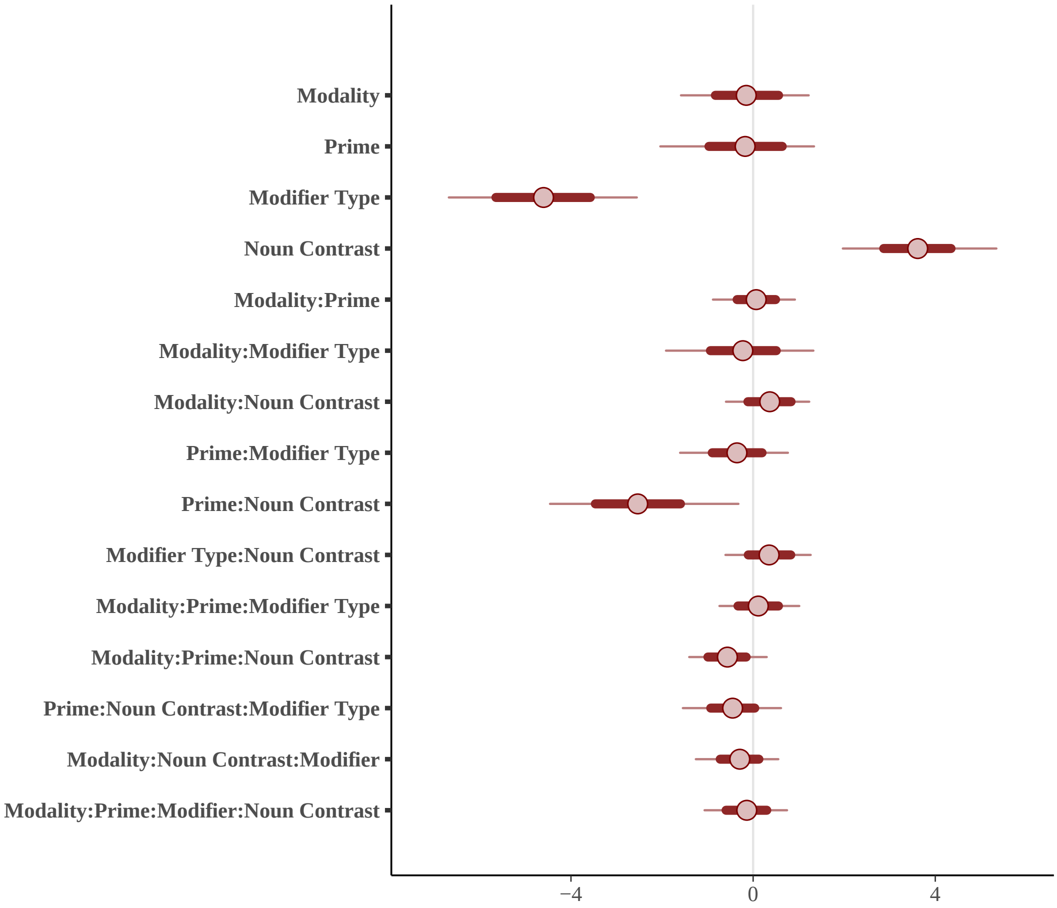
Note: AOI: modifier = *man*, head = *painting*; modifier is N1 and head N2 in possessives, head is N1 and modifier N2 in postnominal phrases. M = sample mean, Mdn = sample median, IQR = interquartile range, N = number of observations

For statistical analysis the proportion of eye samples was converted to empirical logits (see Jaeger, 2008; Mirman et al., 2008). To account for the multi-modal distribution of the data, Bayesian linear mixed models were fitted with three mixture components; a combination of three normal distributions (4 chains, 6000 iterations). The proportion of eye samples was analysed separately for each of the critical areas of interest; i.e. the modifier noun (e.g. the man) and the head noun (e.g. the painting). For contrastive N2 increased proportions of eye samples on N2 were predicted, i.e. the modifier noun in postnominals and the head noun in possessives.

The outcome of both models is summarized in Figure 3.20. Figure 3.20a shows the model for the proportion of eye samples on the modifier referent, i.e. N2 in postnominal phrases and N1 in possessive phrases. There was strong evidence for a Modifier Type by prime by Noun Contrast interaction ($BF_{10} = 5155$). This interaction was inspected as between Noun Contrast comparisons within prime and within Modifier Type. These comparisons revealed strong evidence ($BF_{10} > 4e9$) for a larger proportion of eye samples to the contrastive modifier N2 in postnominal phrases of target trials ($\hat{\mu}$ = -1.39, 95% CrI[-1.67, -1.11], $P(\beta < 0) > .999$) but weak support ($BF_{10} = 2.6$) for this difference in prime trials ($\hat{\mu}$ = -0.44, 95% CrI[-0.83, -0.05], $P(\beta < 0) = .988$). Smaller proportions of eye samples were found for N2 Contrast on phrase-initial modifier referents ($\hat{\mu}$ = 1.75, 95% CrI[1.4, 2.09], $P(\beta < 0) < .001$, BF_{10} > 1e8) for target trials but not in prime trials ($\hat{\mu}$ = 0.37, 95% CrI[-0.08, 0.83], $P(\beta < 0) = .053$, $BF_{10} < 1$).



(a) AOI: Modifier referent – N2 in postnominals, N1 in possessives



(b) AOI: Head referent – N1 in postnominals, N2 in possessives

Figure 3.20: Summary of posterior probability mass for each predictor inferred from the Bayesian linear mixed model on the proportion of eye samples (Experiment 3). Models were fitted for each AOI independently. Dots indicate the posterior mean $\hat{\mu}$, thick lines show the range of 95% of the probability mass (95% CrI) and the thin lines show the entire range of the posterior probability mass. Colons ":" denote interactions.

Figure 3.20b shows the model for the proportion of eye samples on the head referent, i.e. N2 in possessives and N1 in postnominal phrases. Strong evidence was found for a main effect of Noun Contrast ($\text{BF}_{10} >$ 1e8) showing larger proportions of eye samples for N1 Contrast. This effect varied by prime condition indicated by the two-way interaction ($\mathrm{BF}_{10} = 6472$). This interaction was inspected by between Noun Contrast comparisons within prime condition showing strong evidence ($\mathrm{BF}_{10} >$ 3e14) for a larger proportion of looks to the head noun referent – N1 in postnominals and N2 in possessives – for N1 Contrast in target trials ($\hat{\mu} = 3.07$, 95% CrI[2.58, 3.57], $\mathrm{P}(\beta < 0) < .001$) but not in prime trials ($\hat{\mu} = 0.54$, 95% CrI[-0.15, 1.24], $\mathrm{P}(\beta < 0) = .061$, $\mathrm{BF}_{10} = 1.2$). The posterior predicted Noun Contrast effects of both models are summarized in Figure 3.21. This figure illustrates that the proportion of eye samples on the referent of N2 is larger for N2 Contrast in postnominals but smaller in possessives.

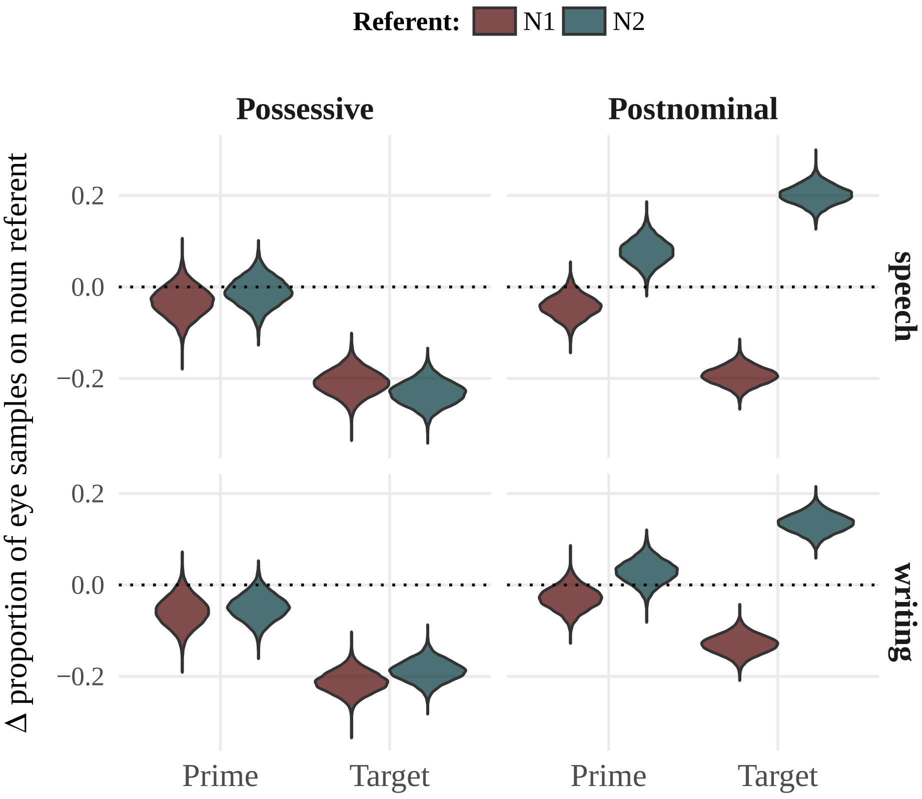


Figure 3.21: Noun Contrast differences (Δ proportion = N2 Contrast– N1 Contrast) for proportion of eye samples on each referent prior to production onset inferred from the respective Bayesian linear mixed models (Experiment 3). *Referent* refers to the image in the order mentioned in the target phrase – possessives: *the N1's N2*; postnominals: *the N1 with the N2*. No difference is indicated by 0, signified by the dotted line.

3.4.3 Discussion

The aim of Experiment 3 was to further investigate whether a conceptual structure underlies advance phrase planning. An interactive production task was used to elicit both written and spoken utterances. The analysis revealed strong evidence showing that noun contrast predicts advance planning of non phrase-initial referents. This shows that phrase planing involves the generation of conceptual structure independently of syntactic form and lexically content.

This conclusion was supported by contrastive non-initial nouns leading to a larger proportion of preonset gaze shifts from N1 to N2 across modifier type in writing and for postnominals but not for possessives in speech. The timing data for this gaze shift supports this contrast effect across all conditions suggesting that noun contrast for possessives was planned after production onset in speech. Also the proportion of eye samples increased for contrastive N2s in postnominal phrases but not in possessives.

An alternative explanation for the Noun Contrast effect relates to priming of event referents (Bock, 1986; Konopka, 2012; Konopka & Meyer, 2014; Van de Velde & Meyer, 2014). To tease out the contrastive relation of the noun pairs, a priming design was used in the present study in which subjects produced the same structure twice with one repeated noun and contrasting the other; e.g. repeating N1 in N2 Contrast. Language planning gives priority to information that is “easy” to encode over information that is “hard” to encode (Konopka & Meyer, 2014). Therefore, having already encoded N1 might facilitate processing of N2. Further under N1 Contrast the primed referent is N2 which would need to be buffered until N1 was encoded to onset the phrase. Such a priming effect for referent images might explain both onset latency and eye data without taking relational information (i.e. contrast) into account. This explanation provides a non-relational account for the assumed relational contrast effect in terms of N2 focused advance planning. The prediction of this account would be that priming N1 under N2 Contrast reduced the onset latency and increases the proportion of looks to the unprimed referent N2. However, contrast effects were not just observed in unprimed but also in primed trials. If priming alone explained these results, one would not have expected these effects in the unprimed trials.

Although the eye data strongly indicate advance planning for N2 Contrast, associated planning durations were found to be systematically shorter for postnominal modifier phrases. In line with E.-K. Lee et al. (2013) and other data from the sentence planning literature (e.g. Martin et al., 2010; Wheeldon et al., 2013), it was predicted that an increased planning scope would lead to longer onset latencies as syntactic and lexical processing is generally costly. The eye data revealed indeed an early apprehension of the non-initial noun referent. However, shorter onset latencies indicated reduced planning effort for N2 Contrast in postnominal phrases. As both lexical content and syntactic structure were held constant, while minimally changing the presentation context, all observed differences must reflect variations in the conceptualisation process.

It is uncertain whether the effect observed in the onset latency is linked to the same underlying cognitive process as the effect observed in the eye movements. The Noun Contrast effect in the onset latency was observed in both prime and target trials while the proportion of N1 to N2 gaze shifts and the proportion of eye samples to N2 was only affected by Noun Contrast in target trials. This suggests that the difference observed in the onset latency is related to a more general processing advantage. For instance if participants aimed at avoiding the use of ambiguous descriptions they would carefully search for images that were conceptually identical to the target. For N2 Contrast in postnominals there were two images identical to the target image while there was just one for N1 Contrast. This might have facilitated the release of production onset. However, this does not explain the Noun Contrast effects observed in the eye data for three reasons: First the eye data show pre-onset Noun Contrast effects in target but not in prime trials (whereas the onset latency effect was observed in both). Second the eye data show Noun Contrast effects in both Modifier Types while the onset latency was affected in postnominals only. Third the contrast structure is not only dependent on the stimulus screen but co-varies with the produced phrase type. Thus if the number of comparators would explain the reported effects one would have expected to see the reversed Noun Contrast effect for possessives instead.

In addition to the tested hypothesis, there was evidence that the release of the production onset in speech is more likely to require planning beyond the minimal linguistic planning unit (i.e. the first noun) than in writing. This was supported by modifier type differences varying across prime/target and output modality revealing longer planning durations for possessives than for postnominal phrases in writing but not in speech. Taken together this suggests that the extent of conceptual planning is uniform across output modality while the scope of grammatical encoding is subject to modality specific constraints. This observation can be explained in terms of avoidance of intra-sentential hesitations. As there was insufficient time for parallel lexical retrieval after production onset (see Allum & Wheeldon, 2007; Griffin, 2003), the second noun required advance preparation to avoid interruption of the speech stream.

Conversely, in writing, lexical retrieval for the second noun in postnominal phrases could be postponed until after production as intra-sentential hesitations do not affect the communicative product. In possessives, however, both nouns required advance planning due to the phrase-initial position of the modifier noun. This phrase type difference for writing was found in prime but not in target trials. The absence of this difference in target trials might be due to the preactivation of syntactic structure and lexical material which may have increased the planning scope (see Konopka, 2012; Van de Velde & Meyer, 2014) as processing demands were relieved (Levelt & Meyer, 2000; Wagner et al., 2010). Hence, the output medium affects the scope of advance linguistic encoding. Spoken output requires preonset planning beyond the first noun, while preparing the first determiner-noun pair suffices to onset writing. However, sentence planning may scope beyond the first determiner-noun pair regardless of output modality if linguistic processing demands are reduced.

3.5 General Discussion

The aim of this chapter was to determine whether conceptual relations influence the planning scope in the production of complex noun phrases. In particular these experiments tested whether the contrastive function of a noun affects the preparation of a complex syntactic phrase irrespective of head-modifier order (see e.g. Allum & Wheeldon, 2007, 2009; Smith & Wheeldon, 1999; Wheeldon et al., 2013) and attachment hierarchies (see E.-K. Lee et al., 2013). Indeed evidence showed that the conceptual structure affects the preparation of non phrase-initial nouns even when syntactic structure and lexical content of the elicited phrases were held constant.

Advance planning beyond the first increment is guided by the semantic relationship between the phrase-initial and later nouns. This conclusion was supported by the finding that the contrastive function of non phrase-initial nouns affect its advance planning. The following evidence was provided: Experiment 1 showed a larger amount of looks to the referent of the second noun associated with longer onset latencies in possessives. Experiment 2 showed weak evidence for the same effect for eye data (but not in onset latencies). The strongest evidence comes from Experiment 3. This experiment revealed larger proportions of eye samples to the referent of non-initial contrastive nouns before production onset, earlier gaze divergence from N1 to N2, and an increased probability to observe gaze divergence from N1 to N2 before production onset. In the reminder of this section, three questions will be discussed regarding the presented evidence, the role of conceptual planning (Konopka & Meyer, 2014), and the contribution of evidence from two output modalities. The section concludes with the interpretation and implications of the present data.

One question that arises from these data is whether advance planning of conceptual relations affects the onset latency at all. Our evidence is mainly based on the dependent variables extracted from the eye movement data of Experiment 3. This experiment provided clear evidence showing that conceptual contrast leads to advance planning beyond the phrase-initial noun. However conceptual contrast did not increase the onset latency which one would typically expect for enhanced planning demands going beyond the phrase-initial noun (Allum & Wheeldon, 2007; Levelt & Maasen, 1981; Martin et al., 2014; Smith & Wheeldon, 1999). The opposite was found showing shorter duration for contrastive non-phrase initial nouns. The latency effect was attributed to a cognitive process different from the one reflected in the eye data. These shorter onset latencies might reflect a facilitated detection of conceptually identical comparators. There are at least three alternative accounts of why the predicted noun contrast effect was not observed in the onset latency data but was found in the eye data.

One possibility is that the time participants spent scanning the image array for the presence of images identical to the target referent covered the noun contrast effect. Another possibility is that advance planning of conceptual relations does not affect onset latencies at all. The generation of conceptual relations is typically characterised as being "very rapid" (Griffin & Bock, 2000; Konopka & Brown-Schmidt, 2014) and involves an efficient allocation of processing resources (Swets et al., 2014) which would not be reflected in onset latencies. The latter assumes that planning dedicated to the non-phrase initial noun is purely conceptual and does not involve the activation of lexical or syntactic representations as those would be costly in terms of processing time (Martin et al., 2014; Wheeldon et al., 2013). A third possibility is that conceptual contrast does not add to the processing difficulty if noun phrase preplanning scoped beyond the first noun anyway. As discussed in Experiment 3 there is some indication that planning scoped, at least to some extent, beyond the first noun. As eye movements provide more fine-grained information about the preparation of each referent, it is possible to separate the message-level planning process from more general processing difficulty as expressed in onset latencies.

Another important question that has received attention recently is whether message planning is incremental building structure around conceptual entities or hierarchical using conceptual relations to guide planning (Konopka & Brown-Schmidt, 2014; Kuchinsky, 2009; E.-K. Lee et al., 2013). The generation of conceptual representations might either give priority to conceptual entities (Gleitman et al., 2007) or to relationships between those entities (Bock et al., 2004) depending on the context (Konopka & Meyer, 2014). The present research provided evidence that conceptual relations lead to planning beyond the phrase-initial noun. However, this does not suggest that message planning is generally guided by conceptual relations. In the present context the generation of conceptual relations is triggered by the need to uniquely identify a particular target image by contrasting it to available alternatives. Hence, every conceptual relation would need to be proceeded by the encoding of a conceptual entity. The attention shift to the next entity is, then, initiated by the need for establishing a contrastive relation between the current and another entity.

Lastly, Experiment 3 reported evidence from spoken and written data showing evidence for conceptual contrast which is a semantic relation that, at least in English, is typically encoded by means of prosodic stress (Jackendoff, 1972; Selkirk, 1995). Both modalities were tested as the absence of prosodic stress for noun contrast might have eliminated contrast effects in Experiment 1 and 2. The literature on advance planning in language production has largely neglected other modalities than speech.

This is problematic as the scope of advance planning is know to be flexible and context specific (F. Ferreira & Swets, 2002; Wagner et al., 2010). Therefore, all available results might be specific to the tested modality and would not necessarily generalise to other modalities. As the present evidence largely reproduced across output modality one can be confident that the proposed results concern a fundamental modality-independent process underlying sentence planning.

The data of Experiment 3 suggest that the absence of contrast effects in the Experiment 1 and 2 was not due to the tested modality. More importantly the replication of this effect in writing and speech has two implications: On the one hand the reported effect must be purely conceptual and cannot be explained in terms of prosodic fluency. A possible concern for spoken responses is that advance planning in speech, in contrast to writers, needs to address fluency requirements on the output (Griffin, 2003). The addition of prosodic stress to a later word might require an early anticipation of the phrase prosody (see e.g. Fuchs, Petrone, Krivokapić, & Hoole, 2013) which does not allow intra-sentential pausing. Therefore planning would be more likely to scope beyond the first noun. However this seems unlikely as the same effects were observed in writing which arguably does neither involve the same fluency constraints nor the generation of prosody. The first point assumed that advance planning in writing involves implicit speech. On the other hand, advance planning in writing might involve the generation of implicit prosody similar as in silent reading (e.g. McCurdy et al., 2013; Thomson & Jarmulowicz, 2016; Wade-Woolley & Heggie, 2015). There is some evidence that phonological representations are activated in orthography and vice versa at least on a segmental level (Qu & Damian, 2017; Vernon, Torrance, & Baguley, 2017; Weingarten et al., 2004). If writing involves the generation of implicit intonation the present results might not reflect a purely conceptual planning process but rather the response to a conceptual representation in a supra-segmental phonological planning process. As there is (to my knowledge) no evidence for the generation of implicit prosody in writing, this interesting possibility may be explored in future research.

In sum, taking into account the alternative explanations and caveats discussed above, the following explanation is being proposed as the most parsimonious account of these findings. The data presented in this research provide evidence that conceptual relations underlying the advance planning of complex noun phrase guide structural assembly. These relations influence advance structural planning beyond the first increment to prepare hierarchically complex relationships between phrase elements. This was found across phrase type. This result shows that structural relations are planned at a conceptual level prior to the generation of syntactic dependencies. This finding is in line with the idea that phrase planning involves hierarchical message representations (Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014; Kuchinsky et al., 2011). Furthermore, this conclusion indicates that syntactic frames (E.-K. Lee et al., 2013; Martin et al., 2014; Smith & Wheeldon, 1999; Wheeldon et al., 2013) and thematic functional units (Allum & Wheeldon, 2007, 2009; Zhao et al., 2015) might play a secondary role for determining the planning scope in sentence production. While conceptual relations must be decided before a message is being submitted to the syntactic assembly process, grammatical encoding might operate purely incremental. The process that creates conceptual relations might, therefore, be the most fundamental operation in the generation of complex noun phrases.

3.6 Conclusion

Research on sentence planning concluded that the syntactic structure of the noun phrase affects pre-onset planning of non phrase-initial items one way or another (Allum & Wheeldon, 2007, 2009; E.-K. Lee et al., 2013; Martin et al., 2010, 2014; Nottbusch, 2010; Smith & Wheeldon, 1999; Wagner et al., 2010). The process of “scaffolding” syntactic configurations, however, must rely on a representation of semantic relations (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014; Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014; Kuchinsky, 2009). This research provided evidence that the generation of hierarchical syntactic dependencies in complex noun phrases is grounded in the generation of semantic relations. Syntactic assembly is the product of the semantic organisation of the message. Therefore it is not necessarily the transition from thought to language that underlies hierarchical dependencies (E.-K. Lee et al., 2013) but rather the relational organisation of the thought itself. This suggests that processing demands attributed to preplanning of hierarchical syntactic structures (E.-K. Lee et al., 2013; Martin et al., 2010; Smith & Wheeldon, 1999) originate from a pre-syntactic planning stage that determines conceptual relations between message elements.

Chapter 4

General Discussion

4.1 Summary

In this thesis I have explored the cognitive mechanisms underlying advance planning in sentence production. I tested specific hypotheses about two stages of advance planning: First, I examined the hypothesis that advance planning is mediated by lexical representations rather than guided by syntactic frames (Chapter 2). Second, I tested whether the extent of advance syntactic planning is determined on a conceptual level of representation (Chapter 3). Importantly, in order to draw conclusion about the language production system these hypotheses have to be tested in more than one modality. Language planning research has focused almost exclusively on data from spoken production and has ignored other modalities. Speech, however, involves modality-specific environmental processing factors that might not apply for other modalities – keyboard typing in the present studies. Hence, the theory that has been derived from existing empirical findings might be partly or largely speech-specific rather than language-general. This bias has warranted the investigation of the two major stages of advance planning in both speech and writing. The reminder of this section summarises the results of this investigation.

Chapter 2 explored the mechanisms subserving grammatical encoding by testing whether the advance preparation of coordinated noun phrases

(Levelt & Maasen, 1981; Martin et al., 2010, 2014; Smith & Wheeldon, 1999; Wagner et al., 2010) is speech-specific. The prediction was that if planning beyond the first noun addresses fluency demands for spoken utterances, one would not expect writers to preplan the entire coordinated noun phrase but only the first noun. This prediction would be in line with lexically-based theories of language production (Allum & Wheeldon, 2007; Griffin, 2003). Further these theories assume that if both nouns in a coordinated noun phrase are planned before onset they must have been fully lexically retrieved. This observation would constitute evidence against explanations for advance planning of coordinated noun phrases that were related to syntax-based theories of language production (Chang et al., 2000, 2003, 2006; V. S. Ferreira & Slevc, 2007; Garrett, 1975, 1980). I found evidence for planning across the entire phrase with similar effects in both speech and writing. Lexical retrieval beyond the first noun was not consistently observed. This supports syntax-based theories of language planning. Coordinated noun phrases require advance structural planning while lexical retrieval for non-initial nouns can be postponed until after production onset. These findings were not merely a by-product of speech-specific processing factors as they reproduced in the written output modality.

Advance planning was consistently found – here and in previous research – for coordinated phrases. Our findings show that increased planning dedicated to coordinated noun phrases cannot simply be explained by speech-specific processing demands as this effect was found regardless of output modality. Noun phrases with subordinated modification, on the other hand, exhibited varying results present in the literature. Some authors found that in subordinated noun phrases only the phrase-initial noun is planned regardless of its syntactic function (Allum & Wheeldon, 2007, 2009; Brown-Schmidt & Konopka, 2008) while other authors found increased planning difficulty related to the phrase's syntactic complexity (F. Ferreira, 1991; E.-K. Lee et al., 2013; Nottbusch, 2010; Nottbusch et al., 2007). This shows that the noun phrase is not necessarily a fundamental unit of advance planning.

However, the varying results for advance planning of coordinated and subordinated noun phrases indicate that some processing stage has to take the relation between nominal elements into consideration. Otherwise it would not be possible for the linguistic processor to know whether or not planning beyond the first noun is required. The question, then, is how does the grammatical encoder know whether it is permitted to plan incrementally or whether hierarchically complex planning is required.

I propose that the scope of advance planning is determined at a pre-syntactic processing stage rather than during grammatical encoding. More recent research suggested that advance planning is not either incremental or hierarchical but might be both under conditions that are yet to be determined (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014; Konopka & Meyer, 2014; Kuchinsky et al., 2011). Konopka and Meyer (2014) argued that the production system gives priority to a lexical incremental or hierarchical strategy depending on the ease of conceptual accessibility provided by the context. Therefore, advance planning of complex noun phrases beyond the phrase-initial noun is possibly determined at a conceptual stage rather than during grammatical encoding. I tested this hypothesis in the experiments described in Chapter 3.

Chapter 3 tested the hypothesis that the scope of advance planning exceeds the first noun if the generated conceptual representation of the message requires the planning processor to do so. Generally speaking this would mean that advance planning is not determined by the grammatical encoder but before; during the conceptualisation process. I tested whether preplanning of subordinated noun phrases can be induced by manipulating a conceptual property of the phrase; the referential contrast of the phrase-initial and the subsequent noun. Crucially, the syntax (and in fact the entire surface form) of the phrase was held constant. This effect was tested in both speech and writing. The data presented in Experiment 3 show advance planning beyond the first phrase element if the second noun is referentially contrastive – the data of Experiment 1 and 2 were rather inconclusive. Thus, subordinated noun phrases are planned hierarchically rather than incrementally if features of the conceptual representation require planning beyond the first noun. This suggests that the unit of advance planning is not pre-determined by the grammatical encoder but is dependent on factors that are part of the conceptual processing stage. The conceptually determined planning scope feeds, then, into the grammatical encoder. The same pattern of effects was found in both speech and writing which demonstrates that the observed difference cannot be explained by modality-specific factors and must be a property of a more fundamental stage of cognitive/language processing.

Taken together the two series of experiments presented in this thesis show that advance planning of complex noun phrases minimally scopes over just the first noun, but that advance planning beyond the first noun is obligated under certain conditions. These conditions are modality independent and therefore fundamental to the language production system: The data suggest that the extended planning scope for coordinated noun phrases cannot solely be explained by communicative pressure to maintain output fluency – as suggested by Allum and Wheeldon (2007) and Griffin (2003). Furthermore I argued that the production system cannot decided on grounds of syntactic properties whether or not advance planning beyond the first noun is required. Instead I provided evidence that the conceptual planning stage determines whether the minimal planning unit requires hierarchical processing beyond the first noun. In other words, the processing system decides pre-syntactically whether to use a non-relational lexical route or a relational hierarchical route (Bock & Ferreira, 2014; Konopka & Brown-Schmidt, 2014). Advance planning beyond the sentence-initial noun is not imposed by the grammatical encoder but is subject to a pre-linguistic processing stage. This research contributes to the planning literature as a first step for developing a modality general model of language production. In the reminder of this chapter I will consider several possible challenges of this investigation, five limitations, and one methodological implication.

4.2 Differences between speech and writing

The conclusions that I drew in the previous section need to be qualified in two ways. First, the absence of an audience, in the majority of the presented experiments, did not influence the present results. Second, writing does not involve fluency constraints and does not impose any new constraints that encourage advance planning. I will discuss these claims in this section.

One criticism of the experiments that I conducted is that most of them – except Experiment 3 in Chapter 3 – depend on an assumed audience. Specifically, it might depend on an imagined audience whether fluency demands are taken into account in speech as for Chapter 2 and whether utterance planning involve ambiguity and contrast considerations in Chapter 3. In other words, there was, in fact, no audience in most of the presented experiments that could have misinterpreted disfluencies or ambiguous utterances. Also there were no contextual factors that invoke pressure on the output. Instead speakers may have tried to initiate utterances as quickly as possible. In Experiments 1–3 of Chapter 2, the absence of an audience may have removed the need to produce fluent utterances in order to onset sentences as quickly as possible.

Arguably, however, the lack of an audience did not have this effect. This is for several reasons. First, the prediction of this factor is against hypothesis. Thus, even if the absence of an audience influences advance planning, I still observed the predicted effects in the experiments of Chapter 2. Second, speech is an automatised skill that involves habitual practice to keep utterances fluent. Utterances that happened to involve disfluencies were removed from the data analysis. Also if speakers initiated utterances as soon as possible – rather than avoiding hesitations by preplanning utterances – we would not observe longer onset latencies for coordinated noun phrases at all. This shows that fluency in spoken utterances is not dependent on the presence of an audience. In the experiments presented in Chapter 3 it was crucial that participants produce unambiguous descriptions. In the absence of an addressee it might have been less important to avoid misunderstandings – to produce sufficiently explicit utterances. While in the presence of an addressee (Experiment 3) there was strong evidence for the Noun Contrast effect, there was only weak or no evidence for this effect in Experiments 1 and 2, respectively. Although there were several other differences between those experiments, the presence of an addressee may have indeed influenced advance planning of the semantic content of the elicited utterances.

Another possibility is that, although writing might be used to isolate obligatory advance planning by removing fluency pressure on the output, writers may have tried to produce sentences as fluently as possibles. One reason for maintaining fluency in the production of written sentences could be that messages were first formulated as inner speech before translating them into written text. This would mean that phonological representations were necessarily activated during writing. There is indeed evidence from image naming studies showing that orthographic representations are being activated via the phonological route (e.g. Bonin & Fayol, 2000; Nottbusch et al., 2005; Qu & Damian, 2017; Weingarten et al., 2004; Zhang & Damian, 2010) but this is not necessarily the case (e.g. Bonin et al., 1998; Rapp et al., 1997; Sahel et al., 2005). For advance planning in sentence production the mediation of writing processes via speech would predict equal planning scopes in both modalities.

Although the effects observed in both sets of experiments generally point towards the same planning processor underlying both modalities, there was some evidence that less advance planning was required in written production. For example in Experiment 3 of Chapter 2 there was (weak) evidence for an effect for advance lexical planning beyond the first noun in speech but not in writing. Also the results from Experiment 3 in Chapter 3 show that more information was prepared in advance in speech than in writing. As this experiment took place in an interactive setting with an addressee present, the context for the written condition was conceptually very similar to spoken interactions – at least conceptually more similar to speech than the written studies of Experiments 1–3 in Chapter 2. The activation of inner speech would therefore be more likely in Experiment 3 of Chapter 3. Instead there was evidence that writers planned only the first noun in postnominal phrases whilst for speech the entire phrase was preplanned. In particular shorter onset latencies were found for postnominal phrases compared to possessive phrases for unprimed written responses which was not seen in speech. Eye data confirmed planning across the entire phrase for the latter. This finding is evidence that, all else being equal, planning in writing might scope over less linguistic material compared to speech, showing that written output can be initiated after preplanning of smaller units than it is the case for speech. As speech requires more advance planning, writing is unlikely to systematically involve the generation of sentences in inner speech. This, however, is not to say that advance planning in writing does not involve inner speech at all. Instead the data show that the advance planning scope may be adjusted depending on the output modality used. The existence of implicit speech or inner prosody in writing is entirely unstudied and may constitute an area for future research.

In sum, this discussion eliminates two possible concerns for the presented data. First, the absence of an audience in the majority of the presented experiments does not relax fluency pressure on speech. If this were the case, there would be no need to preplan utterances beyond the first noun. This is at odds with the findings of Chapter 2. This shows that the fluency demands on spoken utterances are not conditional on the presence of an audience. However, there is reason to believe that the contrast effect that was investigated in Chapter 3 is dependent on presence of an address. Second, fluency constraints do not impact advance planning in writing and there is nothing in writing that encourages more pre-planning beyond fluency. If either were the case, one would assume that the same information is preplanned in both writing and speech. However, the presented data show that less information is preplanned in writing. Therefore, advance planning in writing is not merely an extension of the process that generates speech.

4.3 Caveats and qualifications

There are five caveats of the presented research that I will qualify in this section.

First, all experiments focused on noun phrases with a particular syntactic form, semantic structure, and lexical content. These phrases were produced in a sentential context – except Experiment 3 in Chapter 3 which focused on noun phrases only (used in elliptic sentences). Sentential frames remained similar within experiments – the verb was kept constant (i.e. *moved*, *is*). Noun phrases comprised simple determiner-noun pairs, coordinated noun phrases with two nominal heads, and nouns with nominal modifiers (possessives and postnominal phrases). It is possible that the observed effects do not generalise to linguistic structures that are more complex, including different event structures (verbs). Evidence from previous research however suggests that this is not the case. Effects do not appear to be verb-specific. Martin et al. (2010) found similar effects for a range of verbs other than *move* (e.g. *bump*, *follow*) in the production of coordinated noun phrases. In relation to the conceptualisation process discussed in Chapter 3, there is evidence that relational and non-relational planning are both in principle available in event descriptions (Konopka & Kuchinsky, 2015; Konopka & Meyer, 2014; Kuchinsky et al., 2011). It remains possible, however, that, even though there were no substantial modality differences in the present experiments, this might not generalise to event descriptions as those are conceptually-relational more complex than the conceptual relations tested in this thesis (i.e. *above*, *and*, *'s*, *of*). Testing event descriptions in writing might be an interesting perspective for future research, as writing provides the opportunity to easily test prediction about the time course of the production process which is not as easily achieved in speech (which is discussed in the next section).

A second issues concerns the fluency argument that was brought up throughout this investigation as one difference between advance planning in speech and writing. The limitation at hand is that any modality differences observed in the present research cannot directly be linked to fluency demands on the output. This is because modality difference and not fluency demands were manipulated in the experimental task. As discussed in Chapter 1, there are many other differences between writing and speech that might lead to different planning strategies. However, it was not the intention of this thesis to use modality comparisons as a proxy of fluency pressure on the scope of advance planning (see F. Ferreira & Swets, 2002, 2005). Instead this investigation addressed whether effects from advance planning reproduce across writing and speech. Any effect that does not reproduce across output modality would not be evidence for a feature of the underlying (modality independent) language production system. Indeed the most obvious difference between writing and speech was not the focus of the present research: while phonological representations are generated in speech, writing requires the preparation of orthographic codes which may or may not involve phonology (Damian et al., 2011; Qu & Damian, 2017). Instead this thesis focused on hypotheses about higher level processes – grammatical encoding and conceptualisation.

The next caveats are related to the empirical basis of our data and general limitations known for cognitive research. The third limitation relates to the problem that the used tasks were relatively complex and involve different processes. In principle, the present data do not allow us to separate visual effects and effects linked to language planning, and thus requirements imposed by the grammatical encoder or conceptualisation. All experimental manipulations involved (minimal) changes of the stimulus screens in order to elicit linguistically different structures. Thus, all differences observed in eye movement data and response times might be either due to the activated linguistic representation or visual features of the stimulus. This problem was discussed in the General Discussion of Chapter 2 with reference of previous studies (Martin et al., 2010; Smith & Wheeldon, 1999; Zhao et al., 2015) that aimed at teasing apart visual and linguistic processing factors. The point that these studies emphasised is that visual factors alone do not lead to changes in advance planning if the linguistic processing component was removed. Visual factors do not account for effects of advance planning – at least not for preplanning of coordinated noun phrases. However, this does not mean that effects emerge during grammatical encoding. Chapter 3 suggests that planning of complex noun phrases (co- and subordinated) originate from a conceptual processing stage. As for the results presented in Experiment 3 of Chapter 3, visual factors cannot explain the reported results. The effects of advance conceptual planning for contrastive nouns requires the detection of an ambiguous referent. The apprehension of the disambiguating (and contrasting) noun is not explainable on the basis of visual factors alone.

The next issue regards the tested population. In cognitive research there is a bias towards using local undergraduate students for experiments. The justification for doing so is linked to the assumption that if findings are based on a cognitive process that is specific to language, one should be able to observe similar effects in other populations. As writing is an acquired cognitive skill, some people are better writers than others and some people have not learned how to write at all. In the context of the present language production studies, participants were required to have a sufficient amount of writing experience. Using university students in the present experiments – and in many of the cited studies – has the advantage that the number of participants with experience in keyboard writing is large (in the context of social-media, mobile phones and far beyond). Thus, the amount of usable data is high and the production process is not superseded by problems at the execution stage. However, as written production tends to include less fluency and more editing of already produced language units than spoken production, it was often the case that more data had to be removed from the analysis. The remaining data revealed a larger variability in writing which is due to the smaller number of data points but also due to the greater variability between fast and very slow responses – finding the correct first key takes more time than initiating speech. This suggests that the number of participants tested for experiments in writing might ideally be larger than for similar studies in speech to reduce the impact of this greater variability. In the end of this section I will show that the collection of larger samples of written data is only a minor effort that is worth taking because data preprocessing in writing, unlike speech, can be widely automatised.

The last caveat to be discussed addresses the problem of ecological validity of the presented data. Language production in this thesis was constrained to the elicitation of short sentences on the basis of simple arrays of images. This comes at the cost of potential generalisability. Generally people might use different planning strategies in writing and speaking in laboratory experiments than in real life or in extemporaneous conversations which is subject to many additional processing factors (Pickering & Garrod, 2004). Hence the presented findings may not generalise to language production outside of the laboratory environment and may be specific to the paradigm that was used to uncover them. However, it was important to constrain language production in the present context for the following reason. The hypotheses that were investigated are specifically about language production experiments. Using controlled experiments allows us to establish evidence for the discussed mechanisms under controlled experimental conditions. As it is generally difficult to control factors that underlie language processing, controlling for possible variance in the data (e.g. used sentence structure, lexical material, syntactic position) is essential to draw cause-effect conclusions which are difficult to realise in spontaneous sentence production and corpus studies. Therefore, using tasks that enable us to isolate particular effects by controlling for other variables makes it possible to establish causality. This is important, as naturalistic behavioural data provide only indirect evidence about internal processes (e.g. speech errors, conversation, text writing).

4.4 Methodological implications

The presented findings have an important methodological implication that will be discussed here. The presented effects reproduced in both modalities. Thus advance planning in writing seems to resemble planning in speech, at least to the extent of the tested effects. This is excellent news as these data demonstrate that keyboard typing can be used as output modality additional to or instead of speech. Using keyboard typing instead of spoken utterances has important advances for psycholinguistic investigations. First, as I proposed in Chapter 2, data from keyboard typing are not contaminated by communicative production demands. Furthermore, using keyboard typing as output medium has desirable methodological advantages in the investigation of language production mechanisms. The analysis of spoken data frequently requires intensive time-consuming manual data processing. Writing, in contrast, enables us to automatise many processing steps such as selecting particular parts-of-speech, counting words or sentence structures, extracting chronometric information such as time to production onset and even detailed time course information of the unfolding production – e.g. keystroke intervals between characters before or within words indicating processing variations while preplanning information or planning information in parallel (Torrance & Ofstad Oxborough, 2011). The latter can be easily aligned with data from eye tracking using existing scripts such as EyeWrite (Wengelin et al., 2009). These information are difficult to obtain for spoken data and would require manual work and subjective decisions (e.g. where does one phoneme/word stop and the next begin, what qualifies as a disfluency). There are ways to extract onset latency information (see Bansal, Griffin, & Spieler, 2001; Roux, Armstrong, & Carreiras, 2016) which are, however, affected by filled pauses in sentence-initial position (e.g. *erm*). Consequently, time constraints frequently prohibit researchers from running high-power studies (i.e. collecting large samples of data). Furthermore, the preprocessing and the analysis of keyboard typed responses allows an entirely transparent and replicateable documentation as many, perhaps all, processing steps can be automatised and recorded in analysis scripts. Using keyboard writing as output modality should, therefore, be appealing for researchers that are interested in language production. This is important as a high number of well-known findings in psychology were recently found to be unreliable; some have termed the resulting debate as a *replication crisis* (Open Science Collaboration, 2015; see also Chambers, 2017). At the root of this crisis is low statistical power and lack of transparency in research. Both factors were addressed in the present thesis.

4.5 Conclusion

The present thesis explored high-level cognitive mechanisms – grammatical encoding and the conceptualisation – underlying the advance scope in sentence planning. Language planning research is speech-focused and has widely neglected other output modalities (Alario et al., 2006). As language production is subject to environmental constraints (F. Ferreira & Swets, 2002; Wagner et al., 2010) that might be specific to spoken language use, theories derived from existing empirical findings might be modality-specific rather than language-general. For the first time mechanisms underlying language planning were explored in keyboard typing and directly compared to spoken language production. On the basis of the evidence that I reported in this thesis I conclude:

(1) Advance planning in coordinated noun phrases (*A and the B*) as frequently reported in the spoken production literature (Martin et al., 2014; Smith & Wheeldon, 1999; Wagner et al., 2010) is not speech-specific but reproduces in writing (see Chapter 2). While coordinated noun phrases require preplanning in both modalities, lexical processing beyond the first noun can be postponed until after production onset.

(2) Advance planning for complex noun phrases is determined on a conceptual processing stage, not by the grammatical encoder (see Chapter 3). While coordinated noun phrase exhibit consistent patterns of advance planning, regardless of output modality, subordinated noun phrases (*The A above the B*) permit incremental planning (Allum & Wheeldon, 2007; Brown-Schmidt & Tanenhaus, 2006) but do require hierarchical planning under certain condition (E.-K. Lee et al., 2013; Nottbusch et

al., 2007). Whether or not preplanning has to extend beyond the first noun is determined by the semantic relation of the phrase elements. This was found for both speech and typing.

(3) The theory developed on the basis of speech-specific evidence, that has been developed over several decades, generalises to other modalities – specifically keyboard typing. The presented results are in line with state-of-the-arts theories of language production (Bock & Ferreira, 2014; V. S. Ferreira & Slevc, 2007; Konopka & Brown-Schmidt, 2014). This is an optimistic outcome as it suggests that existing theories of language production are not merely speech-specific but generalises across language-modality.

4.6 Future directions

In this final section, I will illustrate possible directions for future research suggested by this investigation. There are open questions and potential research alleys for investigating language production mechanisms in the context of keyboard typing that may be addressed by future research.

One open question regards the evidence for advance planning of coordinated noun phrases. Research has consistently found longer onset latencies for coordinated noun phrases, regardless of output modality, but not for subordinated noun phrases. The questions here is what causes these additional processing demands for coordinated noun phrases. One alternative explanation is that advance planning is imposed on coordinated noun phrases on a trial-by-trial basis. In other words, while in some trials planning may scope beyond the first noun, only the first noun is preplanned in the remaining trials. This could be tested by modelling onset latencies as a mixture process – i.e. a combination of two Gaussian distributions (for a similar approach see Vasishth, Chopin, Ryder, & Nicenboim, 2017). If coordinated noun phrases are planned incrementally in some trials, one would expect that one of the distributions is virtually indistinguishable from onset latencies for simple noun phrases representing planning for the first noun only. The other distribution would differ substantially and represent advance planning of both nouns.

Another explanation for advance planning of some complex noun phrases was provided in Chapter 3. On the basis of evidence from noun contrast effects in subordinated noun phrases presented in Chapter 3, I concluded that a conceptual planning stage determines whether or not planning beyond the first noun is required. For coordinated noun phrases this might be that the semantic feature determining the coordinating relationship between the first and the second noun phrase might cause the processing system to plan beyond the first noun. The evidence from Chapter 3 shows that semantic relationships between nouns affects preplanning of subordinated noun phrase, but does not show that the same mechanism induces preplanning of coordinated noun phrases. As noted above, coordinated noun phrases involve additional processing difficulty. For instance, one would need to determine to what extent the fact that the order of coordinated noun phrases is syntactically arbitrary causes additional planning demands (Allum & Wheeldon, 2007). This could be accounted for by using coordinated noun phrases that have conventionally predetermined order such as *salt and pepper*. However, these types of phrases might not be syntactically assembled but might be retrieved as one lexical unit. Alternatively, one might use phrases that have conventionally predetermined order but a different semantic coordination. Instead of the conjunction *and*, the disjunction *or* could be used as in *salt or pepper* as this phrase requires the generation of structure but a conventionally determined order. Such an investigation would help to understand whether the semantic relation between phrase elements determines planning scope in coordinates noun phrases.

Furthermore, in Chapter 3, I argued that decisions made at the conceptual processing stage affect the scope of grammatical encoding. While my data suggest advance conceptual planning under semantic contrast, I provided no direct evidence that this changes the grammatical scope. Planning of the second noun may still be postponed until after production onset. My conclusion assumes an automatic cascade from conceptual processes to either or both the lexical and the syntactic processor. To demonstrate that conceptual planning affects grammatical encoding, one would need to cross the noun contrast manipulation with a structural and a lexical manipulation similar to the paradigm used in Chapter 2. E.-K. Lee et al. (2013) provided evidence that would suggest a cascade from the conceptual to the structural processor but not to the lexical processors. However, in their paradigm the authors do not distinguish the generation of conceptual and assembly of syntactic structures. Therefore, their findings might reflect conceptual planning with or without a cascade into syntax.

I discussed earlier that the experimental paradigms used in the majority of the present experiments and studies reported in the advance planning literature (see e.g. Smith & Wheeldon, 1999; Martin et al., 2010; Zhao & Yang, 2013) systematically confounded visual and linguistic processing. By manipulating the visual array to elicit different linguistic structures, any effects might be attributed to either visual processing or indeed advance linguistic planning. For example, advance planning of coordinated noun phrases could be a by-product of visual group of image pairs (see Martin et al., 2010). To remove this confound, one could use a paradigm that does not require the description of a visual stimulus. Instead, participants could be presented with two sentences, a target sentence and a distractor sentence which appears after the target sentence was removed from the screen. The subject will be asked to type one of these sentences after the second sentence disappeared. The purpose of the second sentence is to prohibit the participant to memorise the first sentence word-by-word. This is because first, the participant does not know whether the first or the second sentence needs to be repeated, and second, the second sentence overwrites visual buffer for the first sentence. Therefore, the participant needs to regenerate the linguistic structure of the sentence from the conceptual representation. There is evidence that memory for linguistic representations decays rapidly while conceptual representations are much more persistent (Christiansen & Chater, 2016). Such a paradigm has the advantage that it removes visual confounds from the production task. Also, this paradigm would allow highly controlled experiments with target structures that contain linguistic properties that cannot easily be elicited using visual stimuli (e.g., anaphors, quantifier noun phrases, wh-fronting, cross-over effects). In combination with keyboard typing, this paradigm would provide fine-grained information about linguistic processes throughout the production process. This could be used to investigate how language production unfolds after production onset and over time by extracting inter-keystroke intervals before, after and within words as dependent variables. Ideally the research would predefine the region of interest in which processing changes are predicted.

Another possible pathway for further research is the connection between implicit prosody and writing. While there is evidence that implicit prosody plays a significant role in reading (e.g. McCurdy et al., 2013; Thomson & Jarmulowicz, 2016; Wade-Woolley & Heggie, 2015), and there is a discussion on the activation of phonological features in written naming (e.g. Bonin & Fayol, 2000; Nottbusch et al., 2005; Rapp et al., 1997; Sahel et al., 2005), there is no direct evidence for the activation of supra-segmental prosodic features in writing. On a sentential level one may investigate whether semantic features – such as contrast and focus – which are typically expressed by prosodic stress (Jackendoff, 1972; Selkirk, 1995) – activate similar grammatical features in writing in which this semantic cue remains opaque for the reader. Also syntactic structures that are ambiguous to the reader necessarily had a concrete meaning for the writer. As there is evidence that implicit prosody helps the reader to disambiguate alternating structures (McCurdy et al., 2013), there must be an equivalent representation active during the writing process. Thus, potentially ambiguous syntactic attachments that are disambiguated by prosodic means could be tested in writing.

As discussed before (see also Chapter 2) inner speech might be activated during writing. To control for this possibility, the articulatory suppression paradigm (see e.g. Saito, 1997, 1998) could be used to inhibit the generation of a phonological representations of the target sentence/text. In this paradigm the participant is instructed to repeatedly produce speech sound/syllables (e.g. *da da da da da*) while performing another, primary task. For example, the participant's primary task could be to plan and write a sentence in response to a visual array of images as in Chapter 2. The simultaneous production of irrelevant sound sequences inhibits the activation of a phonological representation of the target phrase. If advance planning in writing underlies the same mechanism as in speech, the effects reported in Chapter 2 should be reproduced even if the activation of phonological representations is systematically suppressed. Research that uses articulatory suppression would be useful to investigate linguistic planning in writing by controlling the activation of phonological representations and to understand to what extent writing is mediated by a phonological route.

In sum, writing – in particular keyboard typing – is a promising way to study how linguistic processes unfold during planning and output of language that has been undeservedly neglected in language production research. Studying language production in the context of writing promises interesting theoretical insights about language production that would be either difficult to access in spoken data or impossible to understand from spoken data alone. Exploring modalities other than speech is essential to tease apart processes that are modality-specific and processes that are language general. To echo Alario et al. (2006), it should therefore be a priority to extent language production research beyond the spoken output modality.

References

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). Hoboken, NJ: Wiley.

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.

Alario, F.-X., Ferrand, L., Laganaro, M., New, B., Frauenfelder, U. H., & Segui, J. (2004). Predictors of picture naming speed. *Behavior Research Methods, Instruments, & Computers*, *36*, 140–155.

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.

Asker-Árnason, L., Wengelin, Å., Sahlén, B., & Ibertsson, T. (2010). Picture-elicited written narratives, process and product, in 18 children with cochlear implants. *Communication Disorders Quarterly*, *31*, 195–212.

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

Bansal, P., Griffin, Z. M., & Spieler, D. H. (2001). *Epd: Matlab wave file parsing software.* `http://homepage.psy.utexas.edu/homepage/`

`faculty/Griffin/resources.html`.

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

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

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*.

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

Behrns, I., Ahlsén, E., & Wengelin, Å. (2010). Aphasia and text writing. *International Journal of Language & Communication Disorders*, *45*(2), 230–243.

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., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 249–278). New York, NY: Psychology Press.

Bock, J. K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*(4), 653–685.

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., & Chalard, M. (2001). Age of acquisition and word frequency in written picture naming. *The Quarterly Journal of Experimental Psychology: Section A*, *54*(2), 469–489.

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., Meot, A., & Fayol, M. (2006). The scope of advance planning in written picture naming. *Language and cognitive processes*, *21*(1-3), 205–237.

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.

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

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology*, *29*(5), 591–620.

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, R., & McNeill, D. (1966). The "tip of the tongue" phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *5*, 325–337.

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

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

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

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

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*.

Chambers, C. (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press.

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.

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, *39*, 1–72. doi: http://dx.doi.org/10.1017/S0140525X1500031X

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.

Delattre, M., Bonin, P., & Barry, C. (2006). Written spelling to dictation: Sound-to-spelling regularity affects both writing latencies and durations. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(6), 1330–1340.

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., Kaiser, E., & Zhao, P. (2017). How are questions made? A production study of object Wh-questions. In *23rd Architectures* *and Mechanisms for Language Processing (AMLaP)* (p. 242). Lancaster, UK.

Eysenck, M. W., & Keane, M. T. (2015). *Cognitive psychology: A student's handbook* (7th ed.). Taylor & Francis.

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, F., & Swets, B. (2005). The production and comprehension of resumptive pronouns in relative clause "island" contexts. In A. Cutler (Ed.), *Twenty-first century psycholinguistics: Four cornerstones* (pp. 263–278). Mahwah, New Jersey: Lawrence Earlbaum Associates.

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.

Flower, L., & Hayes, J. R. (1980). The dynamics of composing: Making plans and juggling constraints. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 31–50). Hillsdale, N.J.: Erlbaum.

Fuchs, S., Petrone, C., Krivokapić, J., & Hoole, P. (2013). Acoustic and respiratory evidence for utterance planning in German. *Journal of Phonetics*, *41*(1), 29–47. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0095447012000678` doi: http://dx.doi.org/10.1016/j.wocn.2012.08.007

Gabry, J., & Goodrich, B. (2016). rstanarm: Bayesian applied regression modeling via stan [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=rstanarm` (R package version 2.13.1)

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.

Garrett, M. F. (1980). Levels of processing in sentence production. In B. Butterworth (Ed.), *Speech and talk* (Vol. 1, pp. 177–220). London, UK: Academic Press.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic Press.

Garrett, M. F. (1988). Processes in language production. In J. F. J. Newmeyer (Ed.), *The cambridge survey of linguistics: Lan-*

guage: Psychological and biological aspects (Vol. 3, pp. 69–96). Harvard University 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., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* Cambridge: Cambridge University Press.

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. (1983). The acquisition of typewriting skill. *Acta Psychologica*, *54*(1–3), 233–248.

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

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

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). Cambridge, MA: MIT Press.

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

Goldrick, M., Ferreira, V. S., & Miozzo, M. (Eds.). (2014). *The Oxford Handbook of Language Production.* Oxford: Oxford University Press.

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. (2004a). The eyes are right when the mouth is wrong. *Psychological Science*, *15*(12), 814–821.

Griffin, Z. M. (2004b). 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.

Griffin, Z. M., & Spieler, D. H. (2006). Observing the what and when of language production for different age groups by monitoring speakers' eye movements. *Brain and Language*, *99*(3), 272–288.

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.

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.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive neuropsychology*, *5*, 67–104.

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

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*(4), 434–446.

Kampstra, P. (2008). Beanplot: A boxplot alternative for visual comparison of distributions. *Journal of Statistical Software*, *28*(1), 1–9. Retrieved from `http://www.jstatsoft.org/v28/c01/`

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

Kellogg, R. T. (1988). Attentional overload and writing performance: Effects of rough draft and outline strategies. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(2), 355–356.

Kellogg, R. T. (2001). Long-term working memory in text production. *Memory & Cognition*, *29*(1), 43–52.

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

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

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

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

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. (2009). *From seeing to saying: Perceiving, planning, producing* (Unpublished doctoral dissertation). University of Illinois, Urbana-Champaign.

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

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behaviour* (pp. 112–136). New York: Wiley.

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., Le Page, R. B., & Longuet-Higgins, H. C. (1981). The speaker's linearization problem. *Philosophical Transactions of the Royal Society of London*, *296*(1077), 305–315.

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., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and brain sciences*, *22*, 1–75.

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.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive science*, *29*(3), 375–419. Retrieved from `http://www.ncbi.nlm.nih.gov/pubmed/21702779` doi: 10.1207/s15516709cog0000_25

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

McCurdy, K., Kentner, G., & Vasishth, S. (2013). Implicit prosody and contextual bias in silent reading. *Journal of Eye Movement Research*, *6*(2), 1–17.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, *8*(3), 299–325.

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.

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

Momma, S., Bowen, Y., & Ferreira, V. (2017). Non-linear lexical planning in sentence production. In *30th Annual CUNY Conference on Human Sentence Processing.* Boston, MA: MIT.

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.

Nottbusch, G., Weingarten, R., & Will, U. (1998). Schreiben mit der Hand und Schreiben mit dem Computer. *Osnabrücker Beiträge zur Sprachtheorie*, *56*, 11–27.

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.

Olive, T., Alves, R. A., & Castro, S. L. (2009). Cognitive processes in writing during pause and execution periods. *European Journal of Cognitive Psychology*, *21*(5), 758–785.

Olive, T., & Kellogg, R. T. (2002). Concurrent activation of high-and low-level production processes in written composition. *Memory & Cognition*, *30*(4), 594–600.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. doi: 10.1126/science.aac4716

Parkhurst, D., & Niebur, E. (2003). Scene content selected by active vision. *Spatial Vision*, *16*(2), 125–154.

Retrieved from `http://booksandjournals.brillonline.com/content/journals/10.1163/15685680360511645` doi: https://doi.org/10.1163/15685680360511645

Phillips, N. (2016). yarrr: A companion to the e-Book "YaRrr!: The Pirate's Guide to R" [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=yarrr` (R package version 0.1.2)

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

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, *27*(2), 169–189.

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.

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.

Roux, F., Armstrong, B. C., & Carreiras, M. (2016). Chronset: An automated tool for detecting speech onset. *Behavior Research Methods*, 1–18. doi: 10.3758/s13428-016-0830-1

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.

Sahel, S., Nottbusch, G., Grimm, A., & Weingarten, R. (2008). Written production of german compounds: Effects of lexical frequency and semantic transparency. *Written Language & Literacy*, *11*(2), 211–228.

Saito, S. (1997). When articulatory suppression does not suppress the activity of the phonological loop. *British Journal of Psychology*, *88*(4), 565–578.

Saito, S. (1998). Phonological loop and intermittent activity: A whistle task as articulatory suppression. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *52*(1), 18–23.

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.

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

Shen, X. R., Damian, M. F., & Stadthagen-Gonzalez, H. (2013). Abstract graphemic representations support preparation of handwritten responses. *Journal of Memory and Language*, *68*(2), 69–84.

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.

Solomon, E. S., & Pearlmutter, N. J. (2004). Semantic integration and syntactic planning in language production. *Cognitive Psychology*, *49*(1), 1–46.

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. (2015a). *Stan: A C++ library for probability and sampling.* `http://mc-stan.org/`. (R package version 2.8.0.)

Stan Development Team. (2015b). *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. (2013). Effects of conversational pressures on speech planning. *Discourse Processes*, *50*(1), 23–51.

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

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., … Bates, E. (2003). Timed picture naming: Extended norms and validation against previous studies. *Behavior Research Methods, Instruments, & Computers*, *35*(4), 621–633.

Székely, A., D'Amico, S., Devescovi, A., Federmeier, K., Herron, D., Iyer, G., … Bates, E. (2005). Timed action and object naming. *Cortex*, *41*(1), 7–25.

Székely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., … others (2004). A new on-line resource for psycholinguistic studies. *Journal of Memory and Language*, *51*(2), 247–250.

Terzuolo, C. A., & Viviani, P. (1980). Determinants and characteristics of motor patterns used for typing. *Neuroscience*, *5*(6), 1085–1103.

Thomson, J., & Jarmulowicz, L. (2016). *Linguistic rhythm and literacy* (Vol. 17). Amsterdam, Philadelphia: John Benjamins Publishing Company.

Tomlin, R. S. (1995). Focal attention, voice, and word order. In P. Downing & M. Noonan (Eds.), *Word order in discourse* (pp. 517–552). Amsterdam: John Benjamins.

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

Torrance, M., & Ofstad Oxborough, G. (2011, 11–13 July). Micro- and macrostructural effects on the timecourse of spontaneous text production. evidence from keystrokes and eye movement. In *The 21st annual meeting of the society for text and discourse.* Poitiers, France. (Paper presentation)

Van de Velde, M., & Meyer, A. S. (2014). Syntactic flexibility and planning scope: the effect of verb bias on advance planning during sentence recall. *Frontiers in psychology*, *5*.

Van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: describing “easy” and “hard” events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*(1), 124–144.

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

Van Der Wege, M. M. (2009). Lexical entrainment and lexical differentiation in reference phrase choice. *Journal of Memory and Language*, *60*(4), 448–463.

Vasishth, S., Chopin, N., Ryder, R., & Nicenboim, B. (2017). Modelling dependency completion in sentence comprehension as a Bayesian hierarchical mixture process: A case study involving Chinese relative clauses. *ArXiv e-prints*.

Vernon, M., Torrance, M., & Baguley, T. (2017). Re-syllabification effects in keystroke production. In *Proceedings of the 2nd writing word(s) workshop* (p. 24). Bristol, UK.

Wade-Woolley, L., & Heggie, L. (2015). Implicit knowledge of word stress and derivational morphology guides skilled readers' decoding of multisyllabic words. *Scientific Studies of Reading*, *19*(1), 21–30. doi: 10.1080/10888438.2014.947647

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.

Weingarten, R., Nottbusch, G., & Will, U. (2004). Morphemes, syllables, and graphemes in written word production. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (Vol. 157, pp. 529–572). Berlin: Mouton de Gruyter.

Wengelin, Å. (2002). *Text production in adults with reading and writing difficulties* (Unpublished doctoral dissertation). Göteborg University.

Wengelin, Å., Leijten, M., & Van Waes, L. (2010). Studying reading during writing: new perspectives in research. *Reading and Writing*, *23*(7), 735–742.

Wengelin, Å., Torrance, M., Holmqvist, K., Simpson, S., Galbraith, D., Johansson, V., & Johansson, R. (2009). Combined eyetracking and keystroke-logging methods for studying cognitive processes in text production. *Behavior Research Methods*, *41*(2), 337–351.

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.

Will, U., Nottbusch, G., & Weingarten, R. (2006). Linguistic units in word typing: Effects of word presentation modes and typing delay. *Written Language & Literacy*, *9*(1), 153–176.

Wundt, W. (1900). *Die Sprache [Language]* (Vol. 1). Leipzig: Kroner-Engelmann.

Zesiger, P., Orliaguet, J.-P., Boë, L.-J., & Mounoud, P. (1995). The influence of syllabic structure in handwriting and typing production. *Les Cahiers de l'ICP. Rapport de recherche*, *4*, 77–87.

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 (Chapter 2)

A.1 Stimuli: Experiment 1

Table A.1: List of stimulus images for second noun by codability (Experiment 1). File ID indicates the Rossion and Pourtois (2004) image

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

Table A.1: *(continued)*

Item	File ID	Image	H	Codability	Movement N1
10	043	camel	0.0	high	down
10	067	sofa	1.1	low	down
11	044	candle	0.0	high	down
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

Table A.1: *(continued)*

Item	File ID	Image	H	Codability	Movement N1
21	137	lettuce	1.1	low	up
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
32	178	bag	1.1	low	up

Table A.1: *(continued)*

Item	File ID	Image	H	Codability	Movement N1
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
43	155	nose	0.0	high	down
43	235	thumb	1.1	low	down
44	158	orange	0.0	high	down

Table A.1: *(continued)*

Item	File ID	Image	H	Codability	Movement N1
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

A.2 Onset latency: Experiment 1

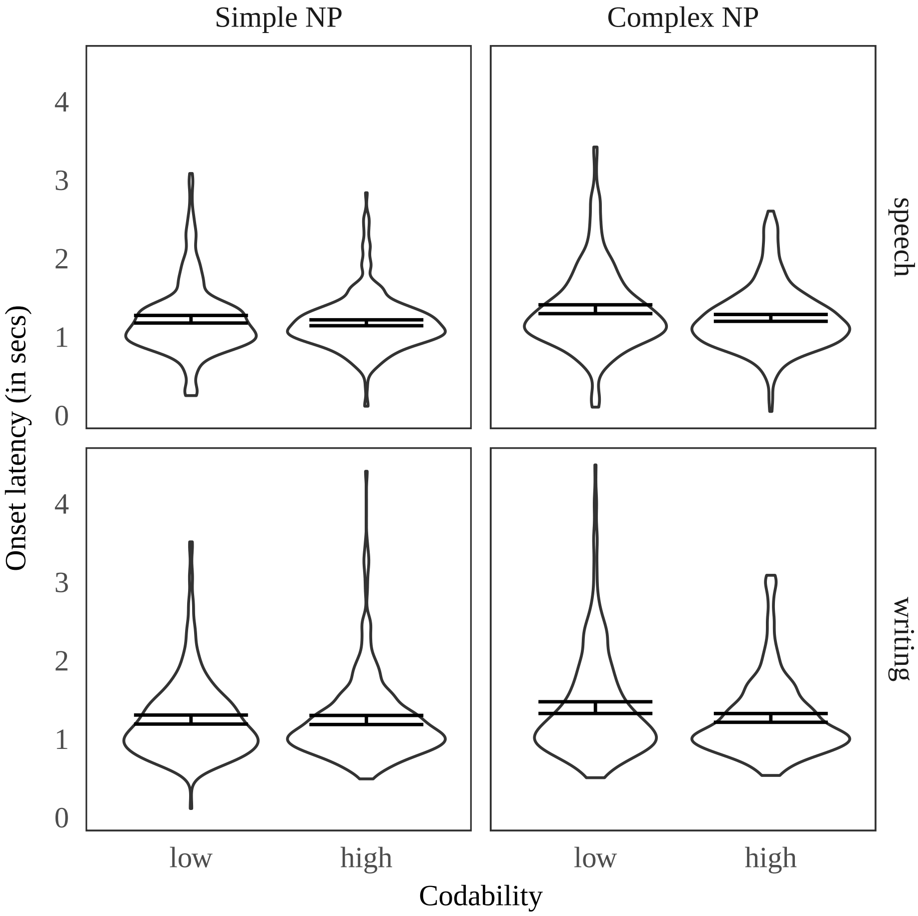


Figure A.1: Bean plots of onset latency (Experiment 1). The beans illustrate the smoothed density of the latency distribution, and the bands show 95% CIs.

A.3 Stimuli: Experiment 2, 3

Table A.2: List of stimulus images for second noun with image name prime and non-word prime (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

Table A.2: *(continued)*

Item	File ID	Image/prime	non-word	Movement N1
21	064	coat	hhfm	up
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
43	117	harp	dqnm	up

Table A.2: *(continued)*

Item	File ID	Image/prime	non-word	Movement N1
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
64	163	peach	duuaq	down
65	165	peanut	hzhwiy	down
66	170	pepper	bttbts	down

Table A.2: *(continued)*

Item	File ID	Image/prime	non-word	Movement N1
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
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

Table A.2: *(continued)*

Item	File ID	Image/prime	non-word	Movement N1
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

A.4 Pilot: priming experiment

In a pilot study we tested whether lexical 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 Table A.2) were tested due to counterbalancing constraints.

Prior to analysis we removed trials with onset latency longer than 10,000 ms (0.53%). Table A.3 shows the descriptive data of the onset latency and the proportion of responses using the most commonly given name by condition.

Table A.3: 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	50ms	1378	66	.84	.03	189
image name	80ms	1229	51	.89	.02	190
non-word	50ms	1624	67	.72	.03	188
non-word	80ms	1688	62	.69	.03	189

Note: Pr(MCN) = proportion of responses using the most commonly given name for a particular image, *M* = sample mean, *SE* = standard error, *N* = number of observations

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 Bayesian linear mixed model are summarised in Table A.4. The model revealed unsubstantial support ($\text{BF}_{10} < 1$) 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 ($\text{BF}_{10} = 2.4$) showing shorter latencies. For non-word primes the model showed negligible evidence ($\text{BF}_{10} < 1$) 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 ($\text{BF}_{10} = 53$). 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 ($\text{BF}_{10} = 2.8$) for a priming effect showing shorter latencies for image names compared to non-words ($\hat{\mu} = 0.16$, 95% CrI[0.06, 0.26], $P(\beta < 0) = .003$). Strong evidence ($\text{BF}_{10} = 352$) for a priming effect was found for 80 ms priming duration ($\hat{\mu} = 0.27$, 95% CrI[0.16, 0.38], $P(\beta < 0) < .001$).

Table A.4: 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)

	Mean	*Lower*	*Upper*	*P($\beta < 0$)*
image name (50 ms)	0.09	-0.01	0.19	.036
image name (80 ms)	0.16	0.05	0.27	.004
non-word (50 ms)	-0.07	-0.12	-0.02	.997
non-word (80 ms)	-0.11	-0.15	-0.07	$> .999$

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that effect β is smaller than 0

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.[1] The results are shown in Table A.5. The proportion of using the most commonly given name increased for image name primes for both 50 ms ($\text{BF}_{10} = 26$) and for

---

[1] The Stan code for the binomial Bayesian linear mixed model is based on Sorensen et al. (2016) and was kindly provided by Bruno Nicenboim.

80 ms ($BF_{10} = 143$) priming duration. The proportion of responses using the most commonly given name remained unchanged for non-word primes at both 50 ms ($BF_{10} < 1$) and 80 ms ($BF_{10} < 1$) priming duration. Comparisons between image name and non-word primes support this effect moderately ($BF_{10} = 5$) for 50 ms priming during ($\hat{\mu} = 1.09$, 95% CrI[0.02, 2.33], P($\beta < 0$) = .022) and substantially ($BF_{10} = 82$) for 80 ms priming duration ($\hat{\mu} = 2.28$, 95% CrI[0.87, 4.09], P($\beta < 0$) < .001).

Table A.5: 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)

	Mean	*Lower*	*Upper*	*P($\beta > 0$)*
image name (50ms)	1.34	0.41	2.49	.004
image name (80ms)	2.35	1.02	4.16	< .001
non-word (50ms)	0.25	-0.37	0.89	.211
non-word (80ms)	0.06	-0.55	0.72	.421

Note: Mean ($\hat{\mu}$) = effect magnitude, *Lower* and *Upper* = 2.5% and 97.5% of 95% CrI, *P($\beta < 0$)* = probability that effect β is smaller than 0

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 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 lexical priming facilitates naming and hence, lexical retrieval.

A.5 Onset latency: Experiment 2

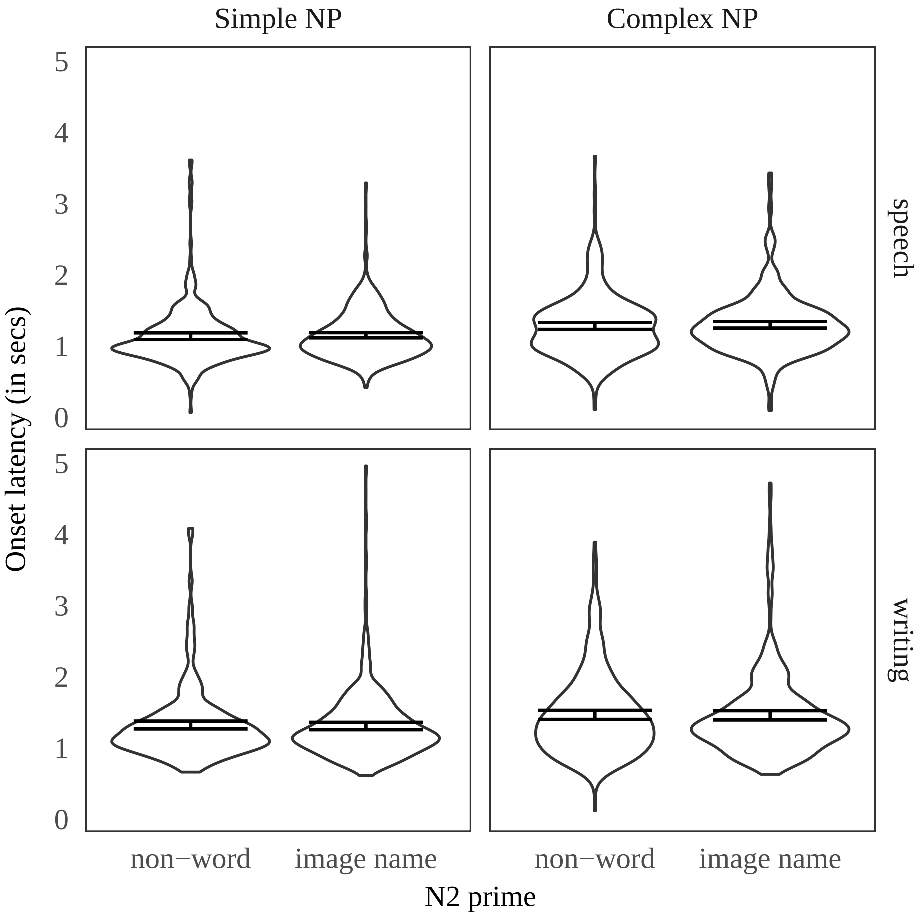


Figure A.2: Bean plots of the production onset latency (Experiment 2). The beans illustrate the smoothed density of the data distribution, and the bands show the 95% CIs.

A.6 Onset latency: Experiment 3

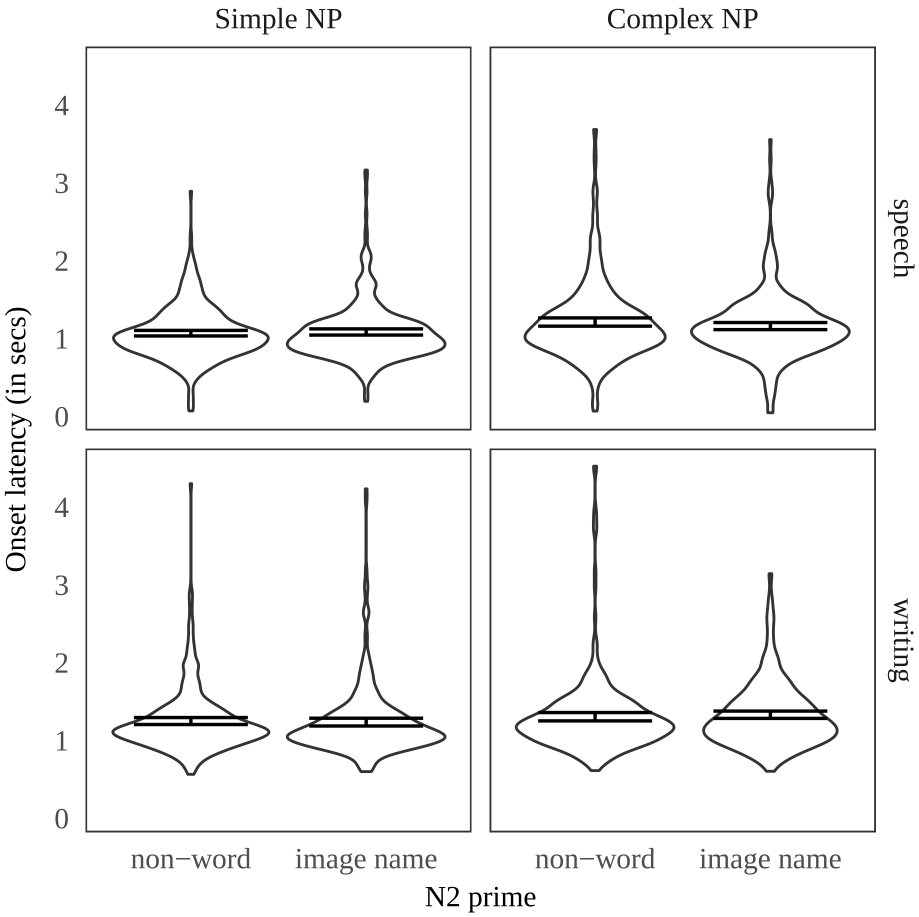


Figure A.3: Bean plots of the production onset latency (Experiment 3). The beans illustrate the smoothed density of the data distribution, and the bands show 95% CIs.

Appendix B (Chapter 3)

B.1 Stimuli: Experiment 1

Table B.1: Stimulus image sets by Noun Contrast (Experiment 1)

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
1	N1	obj131donkey	obj431sun	obj093chicken	obj431sun	red
1	N2	obj131donkey	obj431sun	obj131donkey	obj434swing	red
2	N1	obj334porcupine	obj332popcorn	obj038bee	obj332popcorn	red
2	N2	obj334porcupine	obj332popcorn	obj334porcupine	obj319pillow	red
3	N1	obj432swan	obj390shell	obj233ladybug	obj390shell	red
3	N2	obj432swan	obj390shell	obj432swan	obj399skeleton	red
4	N1	obj249lizard	obj490bricks	obj480unicorn	obj490bricks	red
4	N2	obj249lizard	obj490bricks	obj249lizard	obj492walnut	red

Table B.1: *(continued)*

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
5	N1	obj157fish	obj095church	obj145elephant	obj095church	red
5	N2	obj157fish	obj095church	obj157fish	obj114crackers	red
6	N1	obj402skunk	obj060bread	obj004alligator	obj060bread	red
6	N2	obj402skunk	obj060bread	obj402skunk	obj062bridge	red
7	N1	obj313penguin	obj437table	obj170fox	obj437table	red
7	N2	obj313penguin	obj437table	obj313penguin	obj438tail	red
8	N1	obj045bird	obj340purse	obj239leopard	obj340purse	red
8	N2	obj045bird	obj340purse	obj045bird	obj337present	red
9	N1	obj246lion	obj393shoe	obj416spider	obj393shoe	blue
9	N2	obj246lion	obj393shoe	obj246lion	obj403sled	blue
10	N1	obj516worm	obj327plate	obj266monkey	obj327plate	blue
10	N2	obj516worm	obj327plate	obj516worm	obj314piano	blue
11	N1	obj519zebra	obj367rug	obj297panda	obj367rug	blue

Table B.1: *(continued)*

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
11	N2	obj519zebra	obj367rug	obj519zebra	obj380scissors	blue
12	N1	obj343rabbit	obj181glasses	obj036beaver	obj181glasses	blue
12	N2	obj343rabbit	obj181glasses	obj343rabbit	obj182globe	blue
13	N1	obj409snake	obj345radio	obj290owl	obj345radio	blue
13	N2	obj409snake	obj345radio	obj409snake	obj350razor	blue
14	N1	obj419squirrel	obj196handcuffs	obj034bear	obj196handcuffs	blue
14	N2	obj419squirrel	obj196handcuffs	obj419squirrel	obj193hammer	blue
15	N1	obj286octopus	obj472truck	obj171frog	obj472truck	blue
15	N2	obj286octopus	obj472truck	obj286octopus	obj468trashcan	blue
16	N1	obj388shark	obj080hat	obj207hippo	obj080hat	blue
16	N2	obj388shark	obj080hat	obj388shark	obj083carrot	blue
17	N1	obj111cow	obj460grave	obj086cat	obj460grave	green
17	N2	obj111cow	obj460grave	obj111cow	obj455tire	green

Table B.1: *(continued)*

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
18	N1	obj476turtle	obj063broom	obj302parrot	obj063broom	green
18	N2	obj476turtle	obj063broom	obj476turtle	obj064brush	green
19	N1	obj184goat	obj243lighthouse	obj178giraffe	obj243lighthouse	green
19	N2	obj184goat	obj243lighthouse	obj184goat	obj236leaf	green
20	N1	obj067butterfly	obj108cork	obj501whale	obj108cork	green
20	N2	obj067butterfly	obj108cork	obj067butterfly	obj132door	green
21	N1	obj126dinosaur	obj376saw	obj128dog	obj376saw	green
21	N2	obj126dinosaur	obj376saw	obj126dinosaur	obj397sink	green
22	N1	obj268moose	obj417thread	obj475turkey	obj417thread	green
22	N2	obj268moose	obj417thread	obj268moose	obj426stove	green
23	N1	obj121deer	obj259match	obj273mouse	obj259match	green
23	N2	obj121deer	obj259match	obj121deer	obj260medal	green
24	N1	obj006ant	obj299paper	obj364rooster	obj299paper	green

Table B.1: *(continued)*

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
24	N2	obj006ant	obj299paper	obj006ant	obj298pants	green
25	N1	obj072camel	obj465tractor	obj408snail	obj465tractor	yellow
25	N2	obj072camel	obj465tractor	obj072camel	obj458toilet	yellow
26	N1	obj017baby	obj078canoe	obj284nurse	obj078canoe	yellow
26	N2	obj017baby	obj078canoe	obj017baby	obj082carousel	yellow
27	N1	obj410snowman	obj129doll	obj122dentist	obj129doll	yellow
27	N2	obj410snowman	obj129doll	obj410snowman	obj151feather	yellow
28	N1	obj112cowboy	obj192hamburger	obj357robot	obj192hamburger	yellow
28	N2	obj112cowboy	obj192hamburger	obj112cowboy	obj191brush	yellow
29	N1	obj413soldier	obj235lawnmower	obj061bride	obj235lawnmower	yellow
29	N2	obj413soldier	obj235lawnmower	obj413soldier	obj227kite	yellow
30	N1	obj338priest	obj369saddle	obj342queen	obj369saddle	yellow
30	N2	obj338priest	obj369saddle	obj338priest	obj391boat	yellow

Table B.1: *(continued)*

Item	Noun Contrast	Target 1	Target 2	Comparator 1	Comparator 2	Colour
31	N1	obj513witch	obj149fan	obj489waiter	obj149fan	yellow
31	N2	obj513witch	obj149fan	obj513witch	obj163flower	yellow
32	N1	obj127doctor	obj085castle	obj177ghost	obj085castle	yellow
32	N2	obj127doctor	obj085castle	obj127doctor	obj092chest	yellow

Note: Each Noun Contrast pair was presented in both a possessive and postnominal condition. The coloured target image of the possessive condition is referent 2 and the target image of the postnominal condition is referent 1. *Target 1, 2* = referent of the first and second noun in the target phrase, *Comparator 1, 2* = images presented along the target referents, *Colour* = colour of target image.

B.2 Onset latency: Experiment 1

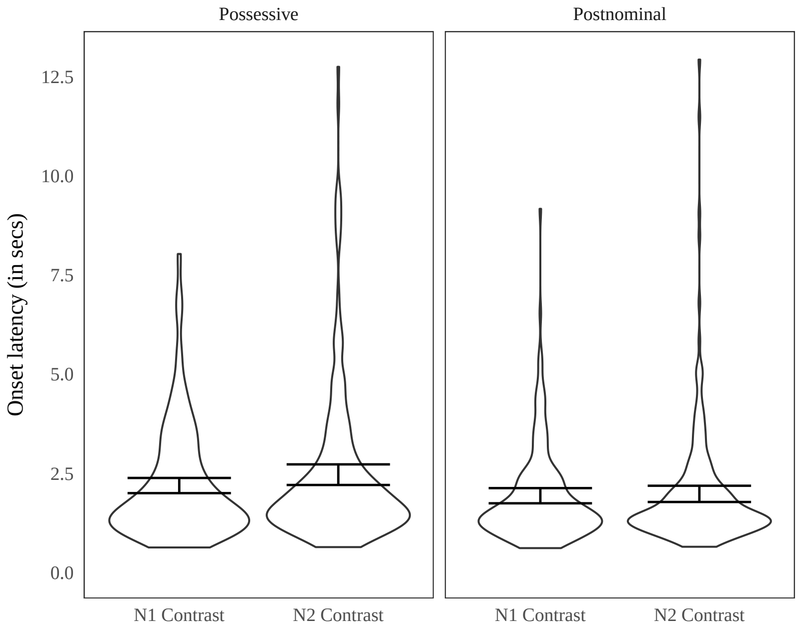


Figure B.1: Bean plots of the onset latency by conditions. The beans illustrate the smoothed density of the distribution of the onset latencies. The bands illustrate the 95% CIs (Experiment 1).

B.3 Stimuli: Experiment 2, 3

Table B.2: Stimulus image sets by Noun Contrast (Experiment 2, 3)

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
1	head	obj006ant	obj213house	obj165fly	obj213house	red
1	modifier	obj006ant	obj213house	obj165fly	obj167football	red
2	head	obj246lion	obj217iron	obj385seal	obj217iron	blue
2	modifier	obj246lion	obj209hoof	obj385seal	obj217iron	blue
3	head	obj286octopus	obj064brush	obj501whale	obj064brush	green
3	modifier	obj286octopus	obj064brush	obj501whale	obj307pear	green
4	head	obj170fox	obj193hammer	obj171frog	obj193hammer	brown
4	modifier	obj170fox	obj396shower	obj171frog	obj193hammer	brown
5	head	obj184goat	obj295palmtree	obj249lizard	obj295palmtree	red
5	modifier	obj184goat	obj295palmtree	obj249lizard	obj219jack	red

Table B.2: *(continued)*

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
6	head	obj034bear	obj395shovel	obj273mouse	obj395shovel	blue
6	modifier	obj034bear	obj244lightning	obj273mouse	obj395shovel	blue
7	head	obj086cat	obj189gun	obj384seahorse	obj189gun	green
7	modifier	obj086cat	obj189gun	obj384seahorse	obj235lawnmower	green
8	head	obj145elephant	obj050bone	obj266monkey	obj050bone	brown
8	modifier	obj145elephant	obj390shell	obj266monkey	obj050bone	brown
9	head	obj133dragon	obj258mask	obj476turtle	obj258mask	red
9	modifier	obj133dragon	obj258mask	obj476turtle	obj070cage	red
10	head	obj130dolphin	obj282net	obj157fish	obj282net	blue
10	modifier	obj130dolphin	obj205highchair	obj157fish	obj282net	blue
11	head	obj251lobster	obj074can	obj313penguin	obj074can	green
11	modifier	obj251lobster	obj074can	obj313penguin	obj437table	green
12	head	obj111cow	obj115crib	obj493walrus	obj115crib	brown

Table B.2: *(continued)*

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
12	modifier	obj111cow	obj263mirror	obj493walrus	obj115crib	brown
13	head	obj032bat	obj332popcorn	obj408snail	obj332popcorn	red
13	modifier	obj032bat	obj332popcorn	obj408snail	obj150faucet	red
14	head	obj067butterfly	obj198harp	obj072camel	obj198harp	blue
14	modifier	obj067butterfly	obj182globe	obj072camel	obj198harp	blue
15	head	obj480unicorn	obj060bread	obj519zebra	obj060bread	green
15	modifier	obj480unicorn	obj060bread	obj519zebra	obj269mop	green
16	head	obj126dinosaur	obj403sled	obj316pig	obj403sled	brown
16	modifier	obj126dinosaur	obj458toilet	obj316pig	obj403sled	brown
17	head	obj344raccoon	obj318piggybank	obj516worm	obj318piggybank	red
17	modifier	obj344raccoon	obj318piggybank	obj516worm	obj073camera	red
18	head	obj121deer	obj183glove	obj178giraffe	obj183glove	blue
18	modifier	obj121deer	obj152fence	obj178giraffe	obj183glove	blue

Table B.2: *(continued)*

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
19	head	obj113crab	obj149fan	obj402skunk	obj149fan	green
19	modifier	obj113crab	obj149fan	obj402skunk	obj255mailbox	green
20	head	obj128dog	obj363roof	obj388shark	obj363roof	brown
20	modifier	obj128dog	obj451thimble	obj388shark	obj363roof	brown
21	head	obj211horse	obj106comb	obj419squirrel	obj106comb	red
21	modifier	obj211horse	obj106comb	obj419squirrel	obj079canopener	red
22	head	obj416spider	obj163flower	obj475turkey	obj163flower	blue
22	modifier	obj416spider	obj124desk	obj475turkey	obj163flower	blue
23	head	obj122dentist	obj027bandaid	obj229knight	obj027bandaid	green
23	modifier	obj122dentist	obj027bandaid	obj229knight	obj090cheese	green
24	head	obj127doctor	obj042bench	obj357robot	obj042bench	brown
24	modifier	obj127doctor	obj315picture	obj357robot	obj042bench	brown
25	head	obj179girl	obj227kite	obj513witch	obj227kite	red

Table B.2: *(continued)*

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
25	modifier	obj179girl	obj227kite	obj513witch	obj094chimney	red
26	head	obj057boy	obj188guitar	obj147eskimo	obj188guitar	blue
26	modifier	obj057boy	obj062bridge	obj147eskimo	obj188guitar	blue
27	head	obj102clown	obj134drawer	obj323pirate	obj134drawer	green
27	modifier	obj102clown	obj134drawer	obj323pirate	obj204helmet	green
28	head	obj342queen	obj056box	obj489waiter	obj056box	brown
28	modifier	obj342queen	obj083carrot	obj489waiter	obj056box	brown
29	head	obj226king	obj254magnet	obj373sailor	obj254magnet	red
29	modifier	obj226king	obj254magnet	obj373sailor	obj066butter	red
30	head	obj061bride	obj356road	obj338priest	obj356road	blue
30	modifier	obj061bride	obj347rain	obj338priest	obj356road	blue
31	head	obj017baby	obj314piano	obj155fireman	obj314piano	green
31	modifier	obj017baby	obj314piano	obj155fireman	obj368ruler	green

Table B.2: *(continued)*

Item	Noun Contrast	Target: modifier	Target: head	Comparator: modifier	Comparator: head	Colour
32	head	obj284nurse	obj063broom	obj410snowman	obj063broom	brown
32	modifier	obj284nurse	obj181glasses	obj410snowman	obj063broom	brown

Note: Each Contrast Noun pair was present in both a possessive and postnominal condition. Noun Contrast on head noun represents N1 contrast in postnominals and N2 contrast in possessives, and *vice versa* in possessives. *Target* = referent of the head and the modifier noun in the target phrase, *Comparator* = images presented along the target referents which constitute the prime trial in Experiment 3. *Colour* = colour of head noun referent (only Experiment 2).

B.4 Onset latency: Experiment 2

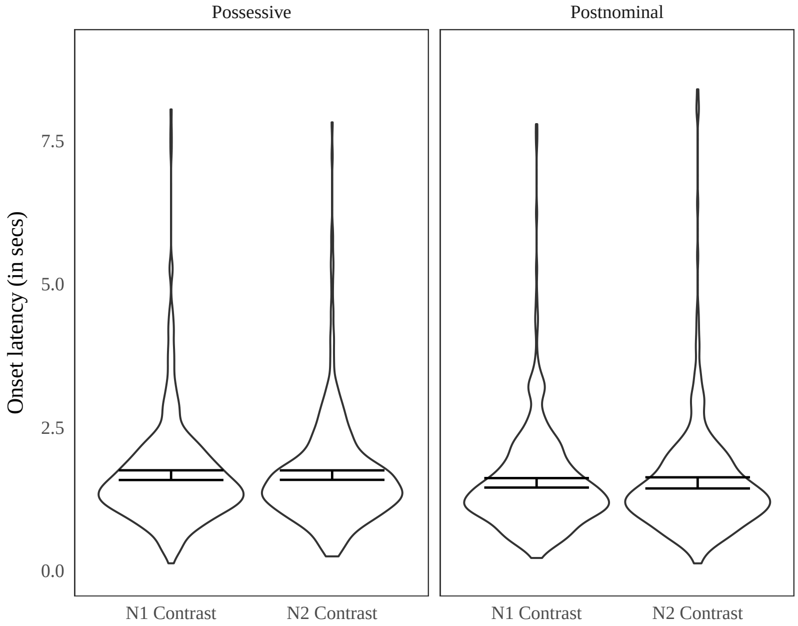


Figure B.2: Bean plots of the onset latency. The beans illustrate the smoothed density of the distribution of the onset latencies. The bands illustrate the 95% CIs (Experiment 2).

B.5 Onset latency: Experiment 3

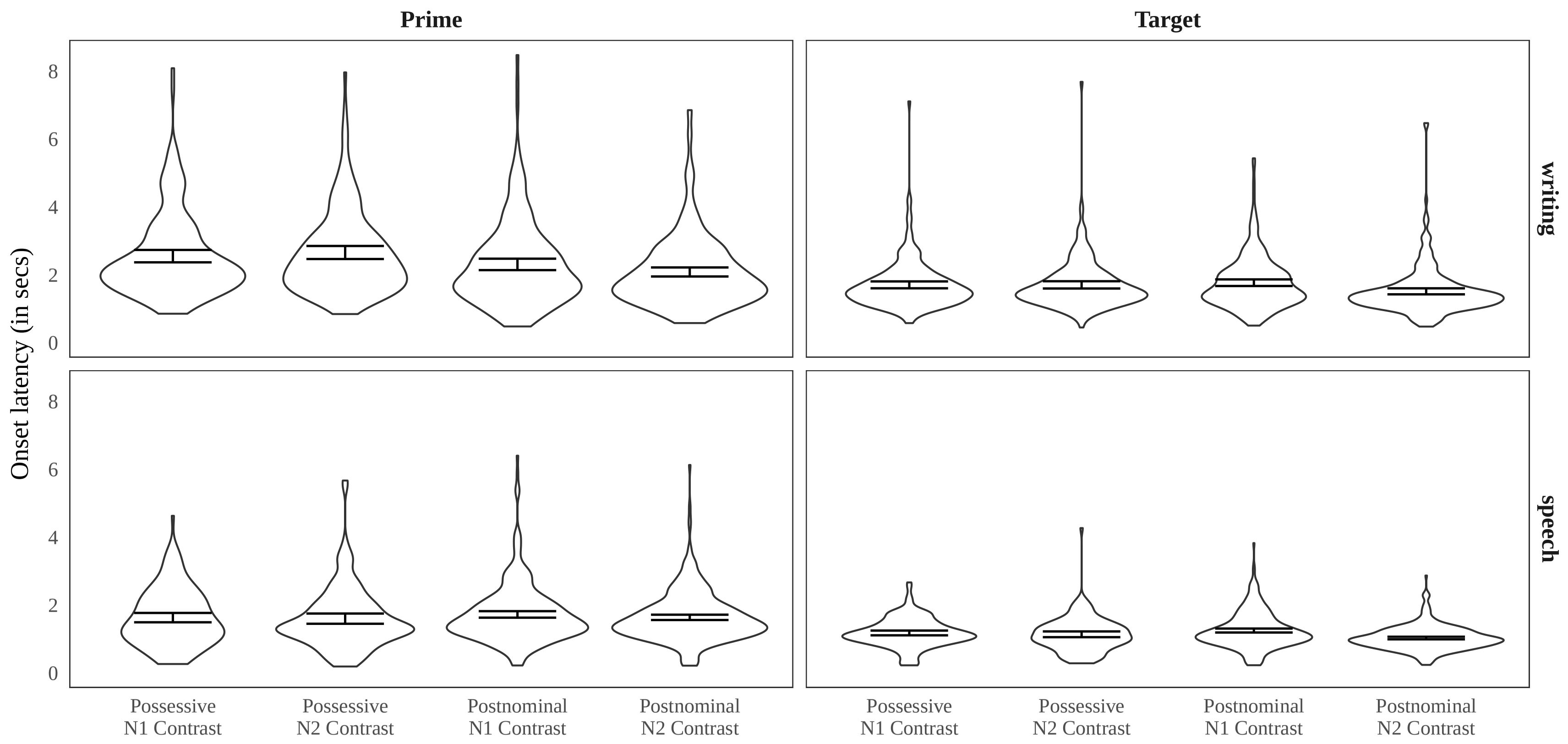


Figure B.3: Bean plots of the onset latency. The data are shown by response modality used. The beans illustrate the smoothed density of the latency distribution. The bands illustrate concentration in 95% CIs (Experiment 3).