Incrementality and flexibility in sentence production

Maartje van de Velde

MPI Series

97

Max Planck Institute for Psycholinguistics

Series

Incrementality and flexibility in sentence production

Maartje van de Velde

Invitation

You are hereby invited to the defense of my PhD thesis entitled

Incrementality and flexibility in sentence production

on Thursday, April 30, 2015 at 10:30 in the aula of the Radboud University Nijmegen, Comeniuslaan 2.

After the defense there will be a reception at the Max Planck Institute, Wundtlaan 1

Maartje van de Velde
maartje.vandevelde@mpi.nl

Paranymphs

Svetlana Gerakaki
svetlana.gerakaki@mpi.nl

Annelies van Wijngaarden
annelies.vanwijngaarden@mpi.nl

Max Planck Institute for Psycholinguistics

Series

INCREMENTALITY AND FLEXIBILITY IN SENTENCE PRODUCTION

ISBN: 978-90-76203-65-2

Printed by Ipskamp Drukkers, Enschede

Incrementality and flexibility in sentence production

Proefschrift
ter verkrijging van de graad van doctor
aan de Radboud Universiteit Nijmegen
op gezag van de rector magnificus prof. dr. Th.L.M. Engelen,
volgens besluit van het college van decanen
in het openbaar te verdedigen op donderdag 30 april 2015
om 10.30 uur precies

door
Maartje van de Velde
geboren op 9 september 1985
te Vlagtwedde

Promotor:

Prof. dr. A.S. Meyer

Copromotor:

Dr. A.E. Konopka

Manuscriptcommissie:

Prof. dr. A.P.A. Roelofs

Prof. dr. P. Hagoort

Dr. L.R. Wheeldon (University of Birmingham, Verenigd Koninkrijk)

TABLE OF CONTENTS

1

INTRODUCTION

1.1 Introduction

Many research articles on word production start with the statement that lexical retrieval, although usually fast and fluent, is an extremely complex process. Now consider the production of an entire sentence: One does not only need to retrieve the correct words for conveying the message in mind, words also need to be combined into organized sequences that obey the grammar rules of the speaker's language. To increase complexity, multiple syntactic structures are often available to frame the same message. Yet, with all these restrictions on the one hand and many choices on the other hand, speaking rate during sentence production is very fast (4.23 syllables per second for Dutch speakers). It is therefore inevitable that part of the speech plan is prepared before speech onset. In this thesis I investigated how much of the speech plan is prepared (i.e., the degree of *incrementality*) under varying circumstances and how speakers deal with the syntactic choices (i.e., syntactic *flexibility*) available to them. Before turning to the key topics of this thesis, I will first provide a general outline of the sentence production process.

1.2 Sentence production

According to most models of language production, sentence planning occurs in three stages (Bock & Levelt, 1994; Chang, 2002; Chang, Dell, & Bock, 2006; Kempen & Hoenkamp, 1987). First a message is constructed that specifies the intended non-linguistic meaning of the utterance. In a second stage, the lemmas (grammatical word units) corresponding to the concepts in the message are retrieved and organized into a syntactic framework via grammatical encoding processes. The third stage is the construction of the sound form of the utterance during phonological encoding.

This thesis focuses on the second stage of sentence production: on *grammatical encoding*. During grammatical encoding, conceptual messages are transformed into linguistically structured word sequences. According to many theories starting with Garrett (1976), grammatical encoding consists of two sub-processes: functional and positional encoding. Functional encoding comprises the retrieval of lemmas (i.e., lexical items with syntactic properties) and the assignment of grammatical roles (e.g., subject, direct object, indirect object) to these lemmas. During positional encoding, lemmas are embedded in syntactic constituents and receive serial order.

It is still debated how functional and positional processes interact. One group of theories suggests that the mapping from a conceptual message to a linguistic structure is lexically mediated (Bock & Levelt, 1994; Kempen & Hoenkamp, 1987). According to these lexically-driven theories, lemma accessibility can directly drive the ordering of constituents in a sentence by prioritizing the encoding of highly accessible units. Other theories posit that there is a direct mapping from the conceptual structure to an abstract hierarchical framework (Chang, 2002; Chang, et al., 2006). Within this framework, slots associated to the thematic roles (e.g., agent, patient, recipient) in the message are assigned a functional-grammatical role. Only during a subsequent positional stage are the slots filled with the associated lexical items. Therefore, according to this group of structurally-driven theories, the generation of a syntactic structure is independent of lexical factors.

1.3 Incrementality

From these two accounts of sentence production, different assumptions can be derived about the amount of pre-planning needed to initiate an utterance (i.e., the degree of *incrementality)* and about the structure of the advance planning scope. According to the lexically-driven models, the formulation and pronunciation of a sentence can be initiated as soon as the first lexical item is encoded, as lemmas in a sentence structure are activated in a word-by-word fashion solely driven by the accessibility of the individual increments. This planning strategy has been referred to in the literature as *linearly* or *lexically* incremental planning (Bock, Irwin, Davidson, & Levelt, 2003).

Conversely, accounts of *hierarchically* incremental planning assume that the number of lemmas retrieved prior to speech is constrained by the grammatical structure of the sentence (Lee, Brown-Schmidt, & Watson, 2013; Wheeldon, Ohlson, Ashby, & Gator, 2013). More specifically, it is not the linear order of the surface structure, but the underlying hierarchical relationships that determine how planning proceeds. According to these theories, planning of a sentence-initial increment also incorporates establishing an overarching structural framework that supports the production of the first increment as well as later ones (Bock, Irwin, & Davidson, 2004; Bock, et al., 2003).

As for the structure of the sentential planning scope, *linear* incrementality predicts that words within the scope of planning are linearly organized. Contrarily, *hierarchical* incrementality predicts that the words within the planning scope are organized in a hierarchical fashion (Lee, et al., 2013). Some grammar formalisms assume that the verb plays a central role in this hierarchy (Ferreira, 2000; Joshi, 1987; Kroch & Joshi, 1985). According to this *verb-centered* hierarchical incrementality, when planning the production of a clause, speakers first lexically encode the main verb, thereby automatically activating its associated subcategorization frames. The verb therefore plays an important role in organizing functional role assignment and subsequent structure choices. Verb-centered hierarchical incrementality and linear incrementality yield different predictions regarding the planning of clauses. For example, consider the Dutch example sentence *De jongen, die de apen de bananen gaf, rende weg* 'The boy who gave the monkeys the bananas ran away'. While the subject *die* 'who' and the verb *gaf* 'gave' in the embedded dative clause share a dependency relation and are thus hierarchically proximate, in the linear surface form of the sentence they are separated by the indirect and direct object of the embedded dative clause. Consequently, *linear incrementality* would predict that after the retrieval of *die* 'who', the indirect object (*de apen* 'the monkeys') would be prepared, whereas *hierarchical incrementality* predicts preparation of the head verb of the clause (*gaf* 'gave') prior to the planning of the indirect object.

So far, studies examining advance planning during sentence production have yielded mixed results. A large body of research has focused on the advance planning of sentences with varying phrase structures (e.g., *the dog and the hat move above the fork*) and has found evidence for the phrase as the default planning unit, for instance by showing that speakers take longer to initiate sentences starting with complex (e.g., *the dog and the hat move*) than simple noun phrases (e.g., *the dog moves*) (Martin, Crowther, Knight, Tamborello, & Yang, 2010; Smith & Wheeldon, 1999; Wheeldon, et al., 2013). However, in other studies using the same sentence types, lexical planning scopes ranging from single lexical items to entire clauses have been found (Griffin, 2001; Meyer, 1996). The production of more complex sentences (such as descriptions of transitive events) has also proven to be flexible, with speakers sometimes prioritizing the encoding of single linguistic elements, and under other circumstances encoding an abstract plan of the utterance before speaking (Gleitman, January, Nappa, & Trueswell, 2007; Kuchinsky & Bock, 2010).

These findings suggest that speakers may choose between hierarchically incremental and linearly incremental planning strategies, based on linguistic and extra-linguistic factors. Research trying to identify these factors, has found that the scope of advance sentence planning decreases under a) cognitive load, b) time pressure, and c) in fast speakers, i.e., speakers with short onset latencies (Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010). There are also linguistic factors that may change the scope of planning. One of these factors is the ease with which the sentence structure can be assembled. In several studies, it has been found that planning scope is expanded when a structure is repeated throughout an experimental procedure as opposed to when different utterances have to be produced (Konopka, 2012; Oppermann, Jescheniak, & Schriefers, 2010; Wagner, et al., 2010). Conversely, planning scope can be narrowed if lower level linguistic and perceptual factors receive emphasis during the sentence planning process. For example, Gleitman et al. (2007) found that speakers prioritized the encoding of a single event character when this referent was perceptually cued.

Importantly, higher level linguistic factors (such as ease of conceptual encoding) have been found to constrain the influence of lower level perceptual factors on advance planning strategies. Kuchinsky and Bock (2010) found that perceptual cuing only influenced the selection of a sentential starting point (and hence the emphasis on encoding only one character) in depicted events that were difficult to encode conceptually. It is therefore important that studies of advance planning take into account multiple factors known to influence the message-to-structure mapping. In this thesis, such a comprehensive approach is taken, by identifying the relative contribution of a range of lower and higher level linguistic factors in influencing sentence planning (Chapter 2). In addition, the possible influence of a new linguistic factor–the availability of multiple syntactic frames–on advance sentence planning is examined (Chapter 4). Finally, an important, new contribution to research on sentence planning to date is made by examining the structure of the advance planning scope in spontaneous speech, using a corpus based approach (Chapter 5). In the thesis outline (see below) each of the experimental chapters is introduced in more detail.

1.4 Syntactic flexibility

Even with grammar putting up certain restrictions on the linguistic form of utterances, speakers are still faced with syntactic choices. For instance, when describing an event featuring an agent (here: Walter), a theme (an apple), and a recipient (Anna), a speaker can choose to produce a prepositional object dative (a PO; *Walter gives the apple to Anna*) or a double object dative (a DO; *Walter gives Anna the apple*). Both structures are grammatically correct and the message that is conveyed is unaffected by the chosen grammatical form (but see construction-based approaches for an alternative view, e.g., Stefanowitsch & Gries, 2003). So what makes a speaker choose one alternative over the other?

The accounts outlined in the section 'Sentence production' lead to different views on how structure choices come about. Lexically-driven models assume that lexical and conceptual accessibility can drive structure choices by prioritizing encoding of highly accessible units during sentence linearization. For example, if at the outset of sentence formulation the theme (*apple*) is more accessible to the speaker than the recipient (*Anna*), it will tend to be placed early in the clause, giving rise to a PO structure (e.g., *Walter gives the apple to Anna*). Studies have shown that lexical and conceptual accessibility (e.g., promoted by salience, recency, animacy or concreteness) play an important role in structure choices (e.g., Altmann & Kemper, 2006; Bock & Warren, 1985; Bresnan, Cueni, Nikitina, & Baayen, 2007; Ferreira, 1994; Kempen & Harbusch, 2004).

Structurally-driven theories, on the other hand, place more emphasis on higher level conceptual and structural factors driving structure choices (Chang, et al., 2006). Priming studies, starting with Bock (1986), provide evidence for these theories by showing that the activation of a syntactic framework can bias the choice for a structural alternative. For instance, when a speaker has just produced a DO dative, he is more likely to re-use this sentence structure on a subsequent occasion when describing a dative event. But activation levels of structural alternatives also depend on existing frequencies reflecting speakers' life-long exposure to written and spoken language. Frequently used structures, such as actives, have higher base level activation than infrequently used (i.e., dispreferred) structures, such as passives (Bock, 1982). Sometimes, the frequency of a structure is tied to the verb it occurs with (Colleman, 2009). In some models, this frequency information is represented via weighted links between verb lemmas and structural representations (Chang, 2002; Chang, Janciauskas, & Fitz, 2012). For instance, a verb that is typically used with the DO dative (e.g., *opleveren* 'yield') will have a connection to the DO structure with a higher weight than to the PO structure (i.e., it has a DO preference). In sum, syntactic priming, existing frequency differences, and verb preferences may all lead to differences in activation levels between structures, thereby promoting the usage of one structural alternative over the other.

However, there are also situations in which syntactic alternatives are equipotent and differences in activation levels do not provide a cue for choosing from alternatives. This *syntactic flexibility* may either facilitate or delay grammatical encoding processes. Ferreira (1996) investigated the influence of syntactic flexibility on sentence production by comparing the speed of speakers' production of constrained sentences (featuring non-alternating verbs, such as *donate*, or the constraining preposition *to* or the pronoun *it*) and unconstrained (i.e., syntactically flexible) sentences (featuring alternating verbs and no constraining constituents). He found shorter sentence onset latencies and fewer errors when participants produced sentences in the unconstrained condition than when they did so in the constrained conditions and concluded that syntactic flexibility facilitates grammatical encoding processes. Ferreira explained this facilitatory effect with an incremental account, in which syntactic flexibility speeds up sentence production by offering speakers more possibilities for filling slots with lexical items. Results were taken as evidence against the existence of competition between structural alternatives that are roughly equipotent.

This thesis has also focused on the question of whether competition can arise between syntactic alternatives during grammatical encoding (Chapter 3). However, instead of contrasting an incremental vs. a competition view, it follows the theoretical distinction of competitive versus non-competitive models that has been proposed in the comprehension literature. In this literature, there is extensive debate on how parsing decisions are made under syntactic ambiguity (i.e., the equivalent of *syntactic flexibility* in production research) and studies have found mixed evidence for competitive and non-competitive theories (see Clifton & Staub, 2008 for a review). In recent language *production* research, results have shown increasing evidence in line with competitive theories (Hwang & Kaiser, 2013; Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013).

1.5 Thesis outline

In four chapters I investigated two key questions related to sentence production:

(1) How do factors acting on different levels of the message-to-language mapping influence the degree of incrementality during advance sentence planning?
(2) How is syntactic flexibility implemented in the grammatical encoding system and how do speakers deal with syntactic choice available to them?

Chapter 2 starts off with a comprehensive assessment of perceptual and linguistic factors influencing sentence planning and formulation in two eye-tracking experiments. In these experiments, participants' eye movements were tracked as they described events that were "easy" or "hard" to encode conceptually (i.e., high or low event codability) using preferred and dispreferred structures (actives and passives). In Experiment 1, ease of (passive) structure assembly and perceptual salience of individual referents were additionally manipulated through cumulative priming of the passive and a perceptual cuing manipulation, respectively. In Experiment 2, only the ease of conceptual encoding of events and agents (event codability and agent codability) was manipulated. Together these experiments aimed at assessing the influence of lower-level perceptual and linguistic factors (i.e., ease of accessing and linguistically encoding individual event referents) vs. higher-level linguistic factors (ease of structure assembly and conceptual encoding) on sentence planning strategies. In line with earlier findings

(Konopka, 2012; Kuchinsky & Bock, 2010), speakers were expected to shift from a *linearly* towards a more *hierarchically* incremental planning strategy as the ease of message encoding and structure assembly increased. Only when events were difficult to encode, were speakers expected to resort to linearly incremental planning, prioritizing the encoding of a highly salient (perceptually or linguistically) event character (Gleitman, et al., 2007).

Chapter 3 addresses the *second* key question of this thesis. By comparing the speed of speakers producing sentences with varying syntactic flexibility, it assesses whether the election of a syntactic frame during grammatical encoding is competitive or non-competitive. According to competition-based models, syntactic flexibility causes structural alternatives to compete until one candidate is selected, thereby delaying sentence production. Conversely, non-competitive models predict that syntactic flexibility speeds up sentence production by providing more possibilities for filling sentence slots with lexical items. In Chapter 3, syntactic flexibility was manipulated by eliciting ditransitive sentences featuring verbs with a preference for one syntactic alternative (i.e., low syntactic flexibility) and sentences featuring verbs without a preference (i.e., high syntactic flexibility), see 1a and 1b.

1a) *De ober serveert de klant de maaltijd.* (weak DO bias)
'The waiter serves the customer the meal.'

1b) *De ober schotelt de klant de maaltijd voor.* (strong DO bias)
'The waiter dishes the customer the meal out'
['The waiter dishes out the meal to the customer']

To elicit these specific target sentences with fixed wording and structure, I used a rapid-serial visual presentation (RSVP) paradigm. In this paradigm, participants are presented with a sentence in a word-by-word fashion at a very high speed (100 ms per word). Subsequently they perform a short distractor task, and then see a sentence preamble (e.g., the subject noun phrase) which they have to complete to form the presented sentence. It is assumed that the fast presentation of the sentence leads to the formation of a conceptual representation of the sentence, which has to be 'rebuilt' during later recall (Potter & Lombardi, 1990).

Faster onset latencies for sentences with low syntactic flexibility would imply that having a syntactic choice delays sentence production, in line with competition-based models. If syntactic flexibility indeed leads to competition between alternatives, the question arises how speakers resolve this competition. One candidate mechanism for promoting fast competition resolution is *executive control* (EC). Executive control is a broad term for the control of cognitive processes, including working memory, reasoning, task flexibility and problem solving. Within the language domain, and more specifically sentence production, one EC subcomponent may be particularly important: selective inhibition. To examine the possible role of selective inhibition in facilitating competition resolving, speakers' performance on the flanker task (Eriksen & Eriksen, 1974) was compared with their speed of producing syntactically flexible sentences.

Chapter 4 combines the themes of incrementality and flexibility by examining the influence of *syntactic flexibility* on grammatical advance planning scope during sentence recall. Syntactic flexibility, i.e., the availability of two (or more) alternative structures to frame one message, may promote elaborate syntactic preplanning due to the necessity to make a syntactic choice. In a study using the RSVP paradigm, speakers produced dative sentences with varying degrees of syntactic flexibility as in Chapter 3 (i.e., sentences with strong versus weak bias verbs, here: *voorschotelen* vs. *serveren* 'dish out' vs. 'serve'). In addition, the frequency of the first (Experiment 1) and second (Experiment 2) post-verbal noun (N1 and N2) in these sentences was manipulated. In Experiment 1, half of the to-be-produced sentences featured a low-frequency N1 and the other half featured a high frequency N1 (see sentences 2a and 2b) while the frequency of the N2 was kept constant (low frequency).

2a) *De ober serveert /schotelt* ***de monarch*** *het feestmaal voor*. (low frequency N1)
'The waiter serves/dishes out **the monarch** the banquet.'

2b) *De ober serveert/schotelt* ***de koning*** *het feestmaal voor*. (high frequency N1)
'The waiter serves/dishes out **the king** the banquet.'

In Experiment 2, I manipulated the frequency of the second noun (N2; high vs. low frequency) while the first noun always had high frequency (see 3a and 3b).

3a) *De ober serveert/schotelt de klant* ***het feestmaal*** *voor*. (low frequency N2)
'The waiter serves/dishes out the customer **the banquet**.'

3b) *De ober serveert/schotelt de klant* ***de maaltijd*** *voor*. (high frequency N2)
'The waiter serves/ dishes out the customer **the meal**.'

By examining how both frequency manipulations affected speech onsets, inferences can be made about the lexical planning scope for sentences with a high and low degree of syntactic flexibility. If speakers lexically encode their sentence up to and including the first noun at the outset of sentence formulation, then faster onsets are expected for sentences starting with high frequency than medium frequency nouns. If their lexical planning scope reaches, by default, up to the second noun, then N2 frequency should predict onset latencies as well. More interestingly, if planning scope changes with the syntactic flexibility of the verb, then noun frequency and syntactic flexibility (i.e., weak vs. strong verb bias) should interact in predicting onset latencies.

Finally, Chapter 5 focuses on the structure of the advance planning scope during spontaneous language production. More specifically, a corpus study was conducted to examine whether words within the scope of advance planning are linearly or hierarchically organized. To this end, effects of verb biases on structure choices were examined in dative clauses with verb-second (VO) order and verb-final (OV) order. In dative OV clauses, the head verb is produced after the objects and hence after the choice for a syntactic structure has been made. Theories of linear and hierarchical incrementality yield opposite predictions regarding the effect of verb bias on structure choices in OV clauses: while linear incrementality predicts no effect at all (only in VO sentences can verb biases influence structure choice), hierarchical incrementality predicts that verb biases can exert their influence regardless of the position of the verb in the sentence. Chapter 6 summarizes and discusses the findings.

References

Altmann, L. J. P., & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes, 21*(1-3), 322-354.
Bock, J. K., & Levelt, W. J. (1994). Grammatical Encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945–984). San Diego: Academic Press.
Bock, K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review, 89*(1), 1.
Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355-387.
Bock, K., Irwin, D. E., & Davidson, D. J. (2004). Putting first things first. *The interface of language, vision, and action: Eye movements and the visual world*, 249-278.
Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. [Article]. *Journal of Memory and Language, 48*(4), 653-685.
Bock, K. J., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*(1), 47-67.
Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. *Cognitive foundations of interpretation*, 69-94.
Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language, 66*(4), 833-849.
Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. [Article]. *Cognitive Science, 26*(5), 609-651.
Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review, 113*(2), 234-272.
Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass, 6*(5), 259-278.
Clifton, C., & Staub, A. (2008). Parallelism and competition in syntactic ambiguity resolution. *Language and Linguistics Compass, 2*(2), 234-250.
Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences, 31*(5), 593-611.
Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics, 16*(1), 143-149.
Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715-736.
Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. *Aspects of language production*, 291-330.
Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language, 46*(1), 57-84.
Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*(5), 724-755.
Garrett, M. (1976). The Analysis of Sentence Production1. *Psychology of learning & motivation: V. 9, 9*, 133.
Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. [Article]. *Journal of Memory and Language, 57*(4), 544-569.
Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1-B14.

Hwang, H., & Kaiser, E. (2013). Having a syntactic choice is not always better: the effects of syntactic flexibility on Korean production. *Language, Cognition and Neuroscience*(ahead-of-print), 1-17.

Joshi, A. K. (1987). An introduction to tree adjoining grammars. *Mathematics of language, 1*, 87-115.

Kempen, G., & Harbusch, K. (2004). A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (pp. 173-181). Berlin: Mouton De Gruyter.

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*(1), 143-162.

Kroch, A. S., & Joshi, A. K. (1985). The linguistic relevance of tree adjoining grammar. *Technical Reports (CIS)*, 671.

Kuchinsky, S. E., & Bock, J. K. (2010). *From seeing to saying: Perceiving, planning, producing.* Paper presented at the 23rd meeting of the CUNY human sentence processing conference.

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

Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., 2nd, & Yang, C. L. (2010). Planning in sentence production: evidence for the phrase as a default planning scope. [Research Support, N.I.H., Extramural]. *Cognition, 116*(2), 177-192.

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.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: the case of English and Russian. [Research Support, Non-U.S. Gov't]. *Quarterly Journal of Experimental Psychology, 66*(8), 1601-1619.

Oppermann, F., Jescheniak, J. D., & Schriefers, H. (2010). Phonological advance planning in sentence production. *Journal of Memory and Language, 63*(4), 526-540.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. [Article]. *Journal of Memory and Language, 46*(3), 586-605.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*(6), 633-654.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73*(3), 205-246.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics, 8*(2), 209-243.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the Flexibility of Grammatical Advance Planning During Sentence Production: Effects of Cognitive Load on Multiple Lexical Access. *Journal of Experimental Psychology-Learning Memory and Cognition, 36*(2), 423-440.

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

2

MESSAGE FORMULATION AND STRUCTURAL ASSEMBLY

Describing "easy" and "hard" events with preferred and dispreferred syntactic structures

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.

Abstract

When formulating simple sentences to describe pictured events, speakers look at the referents they are describing in the order of mention. Accounts of incrementality in sentence production rely heavily on analyses of this gaze-speech link. To identify systematic sources of variability in message and sentence formulation, two experiments evaluated differences in formulation for sentences describing ‘‘easy’’ and ‘‘hard’’ events (more codable and less codable events) with preferred and dispreferred structures (actives and passives). Experiment 1 employed a subliminal cuing manipulation and a cumulative priming manipulation to increase production of passive sentences. Experiment 2 examined the influence of event codability on formulation without a cuing manipulation. In both experiments, speakers showed an early preference for looking at the agent of the event when constructing active sentences. This preference was attenuated by event codability, suggesting that speakers were less likely to prioritize encoding of a single character at the outset of formulation in ‘‘easy’’ events than in ‘‘harder’’ events. Accessibility of the agent influenced formulation primarily when an event was ‘‘harder’’ to describe. Formulation of passive sentences in Experiment 1 also began with early fixations to the agent but changed with exposure to passive syntax: speakers were more likely to consider the patient as a suitable sentential starting point after cumulative priming. The results show that the message-to-language mapping in production can vary with the ease of encoding an event structure and of generating a suitable linguistic structure.

2.1 Introduction

Sentence production involves a sequence of complex operations. As the process of formulating a pre-verbal message and a sentence unfolds over time, speakers make numerous choices about the content and form of their utterances (Levelt, 1989). Such choices can be influenced by long-term biases, such as a general preference for a frequent sentence structure over a less frequent structure (Bock, 1982), as well as recent linguistic (Bock, 1986) and non-linguistic experience (Gleitman, January, Nappa, & Trueswell, 2007). Here we examine how the process of sentence formulation depends on the ease of both formulating the message itself and of expressing this message with a preferred and dispreferred structure (active syntax vs. passive syntax).

Much of what we know about message and sentence formulation comes from eye-tracking studies showing a tight link between gaze and speech. When asked to describe pictured events (e.g., a cat catching a mouse; Fig. 1), speakers normally direct their gaze to the characters in the event in the order of mention (Gleitman et al., 2007; Griffin, 2004; Griffin & Bock, 2000). Gaze shifts from the first character to the second character are typically initiated when speakers finish retrieving the name of the first character. Despite the apparent systematicity of this process, however, it is still debated how speakers begin to direct their gaze to the two characters in such an orderly fashion. This question concerns the way speakers prioritize encoding the different types of information that become available to them during early viewing of an event.

On one proposal, speakers may immediately direct their gaze to the character that attracts their attention at picture onset, irrespective of its status in the event, and begin encoding it linguistically. For example, when perceptual cues are presented in the location of the agent or the patient in an upcoming picture, speakers tend to shift their gaze to this location within 200 ms of picture onset and then begin their sentences with this character rather than the uncued character (Gleitman et al., 2007; Kuchinsky & Bock, 2010; Myachykov, Tomlin, & Posner, 2005; Tomlin, 1997).

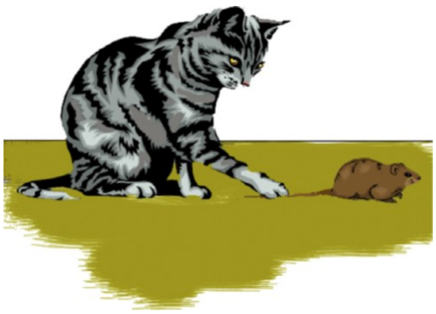

Figure 1. Example of a target item. The modal active description of this event is ‘‘The cat is catching the mouse.’’

Thus, studies employing visual cuing manipulations have suggested that speakers can begin message and sentence formulation by encoding as little as one character (e.g., ‘‘*cat*’’). Planning of the *second* character (‘‘*mouse*’’), as well as its relationship to the first character (‘‘*catching*’’), must then happen when speakers shift their gaze to the second character. This type of planning is referred to as *linear incrementality* (Bock, Irwin, & Davidson, 2004): in each increment, speakers plan only enough information to describe the referent they are fixating and the sentence is built up from the addition of these increments, one by one, to the developing message.

A competing proposal is that speakers direct their attention preferentially to one character (the subject character) only after apprehending the gist of the event. Based on a study carried out without manipulations directing speakers’ attention to a specific character at picture onset (such as subliminal cuing), Griffin and Bock (2000) proposed that apprehension occurs within 400 ms of picture onset: in this time window, speakers may not yet fixate either character preferentially and thus fixations do not predict sentence form. On this account, sentence planning is *hierarchically incremental*: instead of ‘‘zooming in’’ on a single character, speakers first construct a broad but rudimentary conceptual framework for the event that includes information about the relationship between event characters (‘‘*catching*’’). They are then more likely to select a starting point on conceptual grounds, rather than via a bottom-up process like attention capture, and shifts of gaze occurring after 400 ms are more likely to be goal driven (Bock et al., 2004).

The two accounts differ critically in the emphasis they place on lower-level (perceptual) and higher-level (conceptual) influences on early formulation. Most likely, however, speakers can make use of *both* types of information when they begin formulating a message and a sentence. Thus an important question for production models is whether and why speakers might prioritize encoding of either perceptually salient information or conceptual information about the event under different circumstances. Indeed, studies of planning scope in simple utterances (such as numerals or noun phrases) have shown that speakers may prepare larger or smaller increments of a message and sentence before speech onset depending on conversational pressures and resource constraints (Ferreira & Swets, 2002; Konopka, 2012; Wagner, Jescheniak, & Schriefers, 2010). Formulation of more complex utterances (such as descriptions of transitive events) may also be flexible, falling anywhere on a continuum between preparation of small increments that include information about one character (*linear incrementality*) and preparation of increments supported by a larger conceptual framework (*hierarchical incrementality*). The prediction that follows from the accounts outlined above is that the balance between these planning strategies should depend on two key factors: on the one hand, shifts in speakers' focus of attention (Gleitman et al., 2007), and on the other hand, the ease of mapping message-level information onto language (Griffin & Bock, 2000). The *first* of these factors concerns primarily the *selection of starting points*, while the *second* extends the predictions of linear and hierarchical incrementality to the *entire timecourse* of formulation.

Investigating the role of attention in the *selection of starting points*, Kuchinsky and Bock (2010) highlighted important limitations of attention-driven accounts of formulation. As in Gleitman et al. (2007), participants in their experiments described pictures of two-character events in which one of the characters had been subliminally cued. The effectiveness of perceptual cues in biasing speakers to begin sentences with the cued character varied with the ease of encoding event gist – operationalized in terms of the *codability* of the action that the two characters were engaged in. Codability reflects consensus across speakers about the conceptual structure of an event: speakers tend to converge on a small set of suitable verbs for higher-codability events, but use a wider range of verbs for lower-codability events. Kuchinsky and Bock (2010) showed that cued characters were placed in subject position less often when the event was easy to describe (high-codability events) than when it was harder to describe (low-codability events). This suggests that perceptual cues were only weak predictors of starting points when speakers could quickly decide which character to start the sentence with on conceptual grounds (i.e., based on their construal of the event gist). The ease of naming the characters themselves (*character* codability) influenced structure choice primarily in hard-to-describe events.

Whereas Kuchinsky and Bock's work concerned the way speakers select starting points for their utterances, here we examine in detail the entire timecourse of sentence formulation, from the initial selection of a starting point until speech onset, for different types of events. In other words, once speakers have decided where to start, how do they proceed to map the pre-verbal message onto language? Our focus is primarily on conceptual properties of events and on the ease of generating the linguistic material needed to describe the event. Broadly speaking, encoding event gist requires encoding the relational, who-did-what-to-whom structure of an event. At the sentence level, this information must be expressed with a suitable syntactic structure. Thus we hypothesize that the timing of the message-to-language mapping throughout the formulation process should be sensitive to the ease of encoding structural information at two levels. Firstly, the message-to-language mapping should be easy to execute when speakers can quickly decide what they want to communicate about the event (i.e., to encode the event gist or *event structure*) and more difficult otherwise. Secondly, extending this prediction to processes responsible for building *linguistic structure* (Konopka, 2012), we hypothesize that the message-to-language mapping may also be sensitive to the ease of generating a *linguistic* structure. Specific predictions are outlined below.

2.2 Overview of experiments

We describe two eye-tracking experiments that compared formulation of active and passive descriptions of transitive events. Native speakers of Dutch saw a long series of event pictures and described each event with one sentence. They received no further instructions about the content and form of their descriptions to elicit responses in an unconstrained manner. The characters participating in these events varied in ease of naming, and the actions performed by these characters were also either easier or harder to describe. Further, we manipulated the perceptual salience of the two characters and the ease of describing the events with active and passive syntax (see below). Thus we examine whether shifts in the timecourse of formulation towards either linear or hierarchical incrementality can be predicted by variables acting at different levels of the message-to-language mapping process.

For comparison with earlier work, we first consider the effects of early attention shifts (Gleitman et al., 2007; Kuchinsky & Bock, 2010) on the selection of starting points in this item pool. Then we turn to the predictions of linear and hierarchical incrementality for the timecourse of formulation.

Starting points

On attention-based accounts and linearly incremental accounts of formulation (e.g., Gleitman et al., 2007; Tomlin, 1997), the selection of starting points should be predicted by the perceptual salience of individual characters and by early attention shifts. To test this hypothesis, Experiment 1 used a subliminal cuing manipulation to direct speakers' attention to one of the two characters at picture onset (see Gleitman et al., 2007; Kuchinsky & Bock, 2010). Both experiments also considered the relationship between first fixations (irrespective of attention capture) and sentence form.

We also consider whether character-specific properties other than perceptual salience can influence the assignment of a character to prominent structural positions (like the subject slot in a sentence). We focus on character *codability* and character *animacy* (see Kuchinsky & Bock, 2010). Speakers have a strong preference for placing conceptually and linguistically accessible referents in subject position (e.g., Bock, 1982; Bock & Warren, 1985; Christianson & Ferreira, 2005; Ferreira, 1994; McDonald, Bock, & Kelly, 1993), so they should produce more active sentences to describe events with higher-codability agents and with human agents, and, conversely, more passive sentences to describe events with higher-codability patients and with human patients. Both variables reflect the ease of identifying and naming individual characters, and are largely independent of event content. So, like perceptual cuing, they can influence the selection of starting points by processes responsible for encoding of discrete pieces of information. We consider the relationship between first fixations and character properties in both experiments. Effects of character properties on starting point selection would be consistent with a weaker version of linear incrementality, where early formulation is sensitive to processes intrinsic to production proper rather than to lower-level, non-linguistic factors like perceptual salience (replicating Kuchinsky & Bock, 2010; animacy effects on sentence form are well documented, so we focus primarily on character codability).

Timecourse of formulation

Event properties: More importantly, we examine in detail how the timecourse of formulation varies with the ease of encoding and expressing the gist of an event. As shown by Kuchinsky and Bock (2010), event codability can affect the selection of starting points, and we expected the timing of the message-to-language mapping to vary with event codability *throughout* the formulation process as well. To ensure high variability between items on this dimension, Experiment 1 used pictures of events from a wide event codability spectrum and Experiment 2 used items from the higher and lower ends of this spectrum.

If event codability influences processing from the earliest stages of formulation, fixation patterns for higher- and lower-codability events should differ immediately after picture onset. Specifically, in higher-codability events, fixations to the two characters should diverge *slowly* in the first 200–400 ms of each trial: this would be consistent with the assumption of *hierarchical incrementality,* that speakers do not prioritize encoding of one character over the other character at the outset of formulation. In contrast, fixations should diverge more *rapidly* in low-codability events. Here, speakers should find it more difficult to generate a conceptual framework for the event, and – in the absence of such a framework – they should fall back on a different encoding strategy: consistent with *linear incrementality*, they should fixate and prioritize encoding of the character that is most suitable for a sentential starting point based on lower-level factors like perceptual salience or on linguistic factors like ease of naming.

Although linear and hierarchical incrementality make strong predictions about the distribution of early fixations, any effects of event codability can be expected to carry over into later time windows as well (i.e., fixations occurring after 200 or 400 ms). Thus we analyze the distribution of fixations beyond 400 ms to examine the consequences of early differences in fixation patterns for later formulation. In principle, when speakers deploy their gaze to the character that they will mention first (whether it be before or after 400 ms; Gleitman et al., 2007, vs. Griffin & Bock, 2000), they tend to fixate it until they have retrieved its name (Griffin, 2004; Meyer, Sleiderink, & Levelt, 1998) and then shift their gaze to the second character. However, the duration of fixations to the first character may differ across items, depending on what information speakers relied on to select this character as the starting point of their sentence in the first place. On the one hand, if speakers choose a starting point within 200 ms of picture onset on the basis of perceptual salience or ease of character naming (Gleitman et al., 2007; *linear incrementality*), the distribution of fixations after the 0–200 ms time window should primarily show sensitivity to properties of the first character (with shorter gaze durations when this character is easy to name). On the other hand, if speakers select a starting point based on an early construal of the event gist in the first 400 ms of picture viewing (*hierarchical incrementality*), gaze durations on the first character and gaze shifts from one character to another should also be supported by a larger conceptual plan. As a result, gaze shifts to the second character may occur earlier in high-codability than low-codability events, as speakers should find it easier to add the second character to the sentence.

Sentence structure: Formulation may also be sensitive to the ease of generating the linguistic structure that supports expression of the pre-verbal message. The conceptual structure of a transitive event is compatible with both active and passive syntax, but linguistic structures differ in frequency and ease of assembly, reflecting speakers' lifelong exposure to written and spoken language (in Dutch, as in English, actives are preferred to passives; Hartsuiker & Kolk, 1998). On the basis of this frequency difference alone, one might expect the distribution of fixations to agents and patients to be more consistent in active sentences (easy sentences) than passive sentences (harder sentences), showing a higher degree of confidence that speakers can build sentences with an agent rather than a patient in subject position. In addition, since speakers use passive syntax primarily when the patient in an event is animate or is easier to name than the agent, formulation of passives may show more sensitivity to character-specific variables than formulation of actives. To test these hypotheses, Experiment 1 examined formulation of sentences with the preferred active and dispreferred passive structure across different types of events. Additionally, Experiment 1 included a cumulative priming manipulation to facilitate assembly of passive sentences and verify experimentally whether recent experience with a structure changes the timecourse of formulation.

If the message-to-language mapping is indeed sensitive to the ease of structural encoding, generation of an easy structure may leave speakers with enough processing resources to extend their early scope of planning to more than one character (see Konopka, 2012). When assembling an active sentence, therefore, speakers may show little to no preference for either character in the event before 400 ms, consistent with *hierarchical incrementality*. After 400 ms, they should fixate the agent consistently until approximately speech onset and then turn their attention to the patient. As outlined earlier, the timing of gaze shifts throughout the formulation process may be modulated by both event codability and agent codability. In contrast, fixations during formulation of a passive sentence should reflect a strong influence of properties of the patient. Importantly, formulation of passives should proceed in a more hierarchically incremental fashion, showing more sensitivity to event codability only when passives become easier to generate (i.e. after passive priming).

An alternative prediction is that fast assembly of a structure ‘‘primes’’ encoding of the first-mentioned character without influencing processing of the second character. For active structures in particular, there is a natural parallel between a preference to quickly identify the most agent-like character in an event (e.g., Konopka & Meyer, 2011) and the use of a structure that allows early mention of the agent. So if the preferred word order for descriptions of transitive events is agent-initial, high familiarity with active syntax may create a bias to prioritize encoding of the agent at the expense of the patient during early viewing. Similarly, in passive sentences, exposure to passive syntax may create a bias for prioritizing the patient over the agent during early viewing. For both actives and passives, this strategy resembles *linear incrementality* in its emphasis on single-character encoding.

In sum, by assessing the effects of the lower-level perceptual variables and higher-level linguistic and conceptual variables outlined above, we provide a comprehensive analysis of the way speakers coordinate encoding of different types of information in preparation for speaking. In each experiment below, we describe effects of variables expected to produce shifts towards *linear incrementality* first and variables expected to produce shifts towards *hierarchical incrementality* second. We discuss the relative magnitude of these shifts in the General discussion.

2.3 Experiment 1

Eye-tracked Dutch native speakers described pictures of two-character, transitive events embedded in a long list of unrelated filler pictures. We selected a wide range of target events to reflect the variability in messages encountered in every-day speech: the events differed in the ease of gist encoding (event codability) and the characters participating in these events differed in ease of naming (character codability). Codability scores were calculated from the distribution of verbs and character names used to describe each picture.

To assess the effect of lower-level perceptual factors on formulation, the experiment employed a subliminal attentional cuing manipulation analogous to Gleitman et al. (2007) and Kuchinsky and Bock (2010): target pictures were preceded by a 60-ms black dot shown in the location of the agent or the patient in the upcoming picture or in a neutral location. Effects of higher-level structural factors on formulation were assessed by comparing the distribution of agent-directed and patient-directed fixations over time across items.

The task also included a cumulative priming manipulation midway through the experiment to facilitate production of passive descriptions (Kaschak, 2007; Kaschak, Kutta, & Schatschneider, 2011; see Pickering & Ferreira, 2008, for a review of structural priming). The experimental session was divided into three blocks: in Blocks 1 and 3, speakers described pictures of events, while Block 2 consisted of a reading task where speakers read a list of sentences from the computer screen and performed a simple memory test. Speakers were exposed to an unequal distribution of sentences with the two target structures in this task (3 active and 27 passive sentences, i.e., a 1:9 ratio of active to passive sentences).

Analyses were carried out in several steps. We first verified whether the rates of production of active and passive descriptions changed with factors previously shown to influence sentence form: perceptual cuing and the distribution of first fixations (Gleitman et al., 2007; Kuchinsky & Bock, 2010), character codability (Kuchinsky & Bock, 2010), character animacy (e.g., Ferreira, 1994), and cumulative structural priming (Pickering & Ferreira, 2008). Based on earlier work, speakers were expected to begin their sentences with characters that first attracted their attention and characters that were easy to name. They were also expected to generally prefer active syntax over passive syntax, but to produce more passive sentences after performing the reading task in Block 2.

Next, we examined how these factors bear on the formulation process. Timecourse analyses were performed over a series of early and late time windows. Analyses of Block 1 served to assess baseline differences in the production of sentences with preferred and dispreferred syntax, and Block 3 tested whether the timecourse of formulation changed with exposure to the dispreferred passive construction. If formulation is sensitive to properties of the events, fixations to agents and patients in higher- and lower-codability events should differ across the entire experiment. Similarly, if formulation is modulated by the ease of generating a linguistic structure, differences in the gaze-speech link for active and passive sentences should be present across the entire experiment: comparing production of actives and passives should clarify to what extent speakers are hierarchically or linearly incremental when using the preferred active structure and thus how its formulation differs from the passive. Any changes in formulation of passive sentences from Block 1 to Block 3 should then demonstrate whether recent experience with a dispreferred linguistic structure changes the message-to-language mapping.

2.3.1 Method

Participants

48 adult native speakers of Dutch (ages 18–40 years) with normal or corrected-to-normal vision participated for payment.

Materials and design

60 Target pictures of transitive events were constructed from the Microsoft Clipart collection and materials from earlier experiments (e.g., Bock, 1986; see Fig. 1 for an example and Appendix A for a full list). The pictures showed two target characters (an agent and a patient) engaged in simple actions (e.g., kicking, pushing), as well as a third, unrelated object (e.g., tree, cloud). The target characters were chosen to include characters encountered in every day speech. Hence, some characters occurred more often than others (e.g., *boy, girl, car* vs. *waiter, pirate, windmill*). There were 31 pictures with human agents, 13 with animate agents, and 16 with inanimate agents; 30 of these pictures included human patients, 8 included animate patients, and 22 included inanimate patients.

Two mirror-reversed versions of each picture were created to counterbalance the placement of agents and patients on different trials. Thus, agents appeared on the left and right side of the screen equally often within one experimental list. Since agent orientation did not influence visual scan paths (left-oriented and right-oriented agents both attracted 54% of first fixations) and had a weak influence on structure choice (speakers produced 80% actives to describe pictures with left-oriented agents and 77% actives to describe pictures with right-oriented agents), all analyses collapsed across this factor.

The pictures were divided into two blocks (Blocks 1 and 3). Each block included 30 unique target pictures, and the order of the blocks was counterbalanced across participants. The target pictures were interspersed among 130 filler pictures (65 in each block) showing a range of events with one or two characters that could be described with a variety of structures. Two fillers intervened between any two target trials. Pictures within lists were arranged to minimize semantic overlap and repetition of content words on adjacent trials. Pictures were displayed at sizes ranging from 450 x 450 to 600 x 600 pixels, depending on the size of the characters. All target pictures were preceded by a subliminal cue (a black dot, 32 pixels in diameter, subtending a 1^{o} visual angle), presented for 60 ms in the location of the agent, patient, or neutral object before picture onset. Experimental lists were constructed so that each picture appeared in each of the three cuing conditions on different lists, resulting in a 2 (Block) x 3 (Cue condition) within participant and within-item design. The animacy of agents and patients was also balanced across Blocks and Cue conditions.

Block 2 of the experiment was a reading task. The materials used in this task consisted of 3 active, 27 passive, and 60 intransitive sentences that were unrelated to the events shown in target pictures.

Procedure

Participants were seated at an SR Research EyeLink 1000 Tower Mount eye-tracker (500 Hz sampling rate), 70 cm away from a 22-in. monitor. They were told that they would describe sequences of pictures (Blocks 1 and 3) and perform a short reading task (Block 2). For the picture-description task, they were instructed to produce grammatically correct sentences, mentioning all event characters, as quickly and fluently as possible. The experiment began with seven practice trials. Each trial started with a fixation cross at the top of the screen, followed by a 60 ms blank screen (filler trials) or a 60 ms attentional cue (target trials). Participants then saw a picture and produced their response. The experimenter pressed a button to continue to the next trial.

Block 2 was presented as a reading and memory task. Participants were asked to read sentences out loud and to remember them for a later memory test. Sentences were presented individually: participants read each sentence and the experimenter pressed a button to continue to the next trial. One third of all trials (1 active, 9 passive, and 20 filler sentences) were followed by a question mark, prompting participants to repeat the sentence they had just read from memory. Due to the large number of fillers, the unequal distribution of active and passive sentences went unnoticed, as participants declared in a post-experimental debriefing.

Sentence scoring

For all target trials, fixations were scored as falling within the interest area of the agent, patient, neutral object, or outside these areas. Sentences were scored as actives, passives (e.g., *The cat caught the mouse; The mouse was caught by the cat*), and truncated passives (*The mouse was caught*; these sentences were not included in the analyses). Sentences with intransitive syntax or other constructions were also excluded, as were sentences with indefinite pronouns (e.g., *someone*) or restarts (e.g., *a woman shakes uhm- congratulates uhm a man – the man congratulates a woman*). Finally, we eliminated responses with onset latencies longer than 5000 ms and 3 standard deviations above the grand mean. The final dataset consisted of 1664 sentences (79% actives, 21% passives). Exclusion rates were similar across blocks (42%) and item categories (ranging from 40 to 45% for lower- and higher-codability items and items with lower- and higher-codability agents).

Codability scoring

The ease of encoding event gist and identifying characters was operationalized in terms of event and character codability (Kuchinsky & Bock, 2010). Codability was estimated with Shannon's entropy on the basis of the frequency and number of different verbs and nouns used to describe the action performed by the two characters and the characters themselves.

Events described consistently with a small number of verbs received low entropy scores (low verb heterogeneity, and thus high event codability) and event eliciting descriptions with a wider range of verbs received higher entropy scores (high verb heterogeneity, and thus low event codability). For example, the picture of a woman massaging a man was consistently described with one verb and received a high codability score; the picture of a bodyguard pulling a speaker aside was described with a variety of verbs and received a lower codability score. Likewise, characters described consistently with a small number of nouns received low entropy scores (low noun heterogeneity, and thus high character codability) and characters described with a wider range of nouns received higher entropy scores (high noun heterogeneity, and thus low character codability). Table 1 lists mean codability scores after a median split.

Event codability was not correlated with Agent or Patient codability (r = .19 and .09), showing that the ease of describing the event and the ease of identifying its protagonists were independent sources of encoding difficulty in this item pool. However, Agent and Patient codability were positively correlated (r = .47). All codability scores were included in the models as continuous predictors (for clarity, figures show codability results for events, agents, and patients after a median split).

Table 1. Mean codability scores (with standard deviations) for high- and low-codability events with high- and low-codability agents and patients in both experiments.

Event codability (median split)	Agent codability (median split)	Event codability score	Agent codability score	Patient codability score
Experiment 1				
High	High	.79 (.43)	.36 (.41)	.47 (.44)
High	Low	.83 (.53)	1.56 (.41)	1.39 (.89)
Low	High	1.89 (.30)	.19 (.27)	.73 (.55)
Low	Low	2.19 (.51)	1.93 (.55)	1.37 (1.00)
Experiment 2				
High	High	.86 (.37)	.44 (.40)	.85 (.75)
High	Low	.63 (.44)	1.63 (.38)	1.04 (.78)
Low	High	1.98 (.58)	.29 (.33)	.78 (.49)
Low	Low	2.05 (.55)	1.65 (.53)	.61 (.58)

Analyses

The data were analyzed with mixed logit models (for analyses of structure choice) and linear mixed effects models (for analyses of speech onsets), treating participants and items as random effects (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). Timecourse analyses were carried out with by-participant and by-item quasi-logistic regressions (Barr, 2008). Speech onsets are discussed after the timecourse results, performed for sentences without initial disfluencies (1576 trials). All factors were centered and random slopes for these factors were included when they improved model fit to capture additional variability at the subject and item level (models with maximal random structures showed similar results and are therefore not listed; cf. Barr, Levy, Scheepers, & Tily, 2013). Interactions were included when they improved model fit. Effects were considered to be reliable at $\alpha < .05$.

2.3.2 Results

We first tested whether first fixations predicted sentence form. Next, we compared the influence of Event codability, Character codability, Character animacy, and the priming manipulation on sentence form and on the timecourse of formulation in active and passive sentences.

First fixations and structure choice

The majority of first fixations fell on agents (55%), followed by fixations to patients (35%) and neutral areas (10%). Speakers looked at the cued character on 58% of all trials, and the distribution of first fixations varied by cuing condition: speakers looked at cued agents and cued patients on 87% and 62% of all scored trials respectively (Fig. 2).

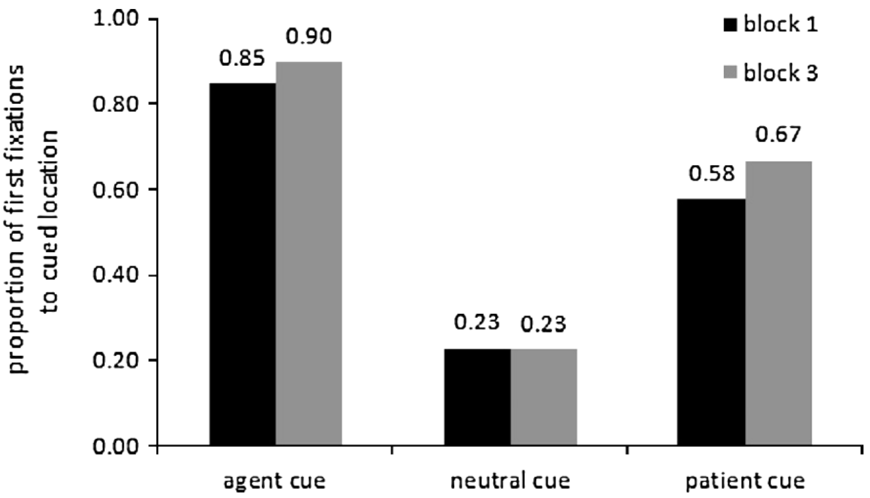

Figure 2. Proportion of first fixations directed to the cued location (agent, patient, and neutral object) for all scored trials in Experiment 1.

Despite the effectiveness of the subliminal cues in directing attention to cued characters, first fixations did not reliably predict structure choice. Descriptively, speakers produced actives at comparable rates after agent and neutral cues, and tended to produce fewer actives after patient cues; however, on trials with successful attention capture (Fig. 3a), the effect of cue condition varied considerably across items and did not reach significance (β = -.58, z = 1.46, with random by-item slopes for Cue condition; the majority of the items did not show an effect in the expected direction).

The difference in the proportion of active sentences produced after fixations to the agent and patient was only .03 for high-codability events and .07 for low-codability events, (consistent with Kuchinsky & Bock, 2010), but the effect of Event codability failed to reach significance. The effectiveness of the cues also did not differ with respect to Agent or Patient codability (all *z*s < 1.5 for the main effects and interactions with the location of first fixations) and did not change from Block 1 to Block 3.

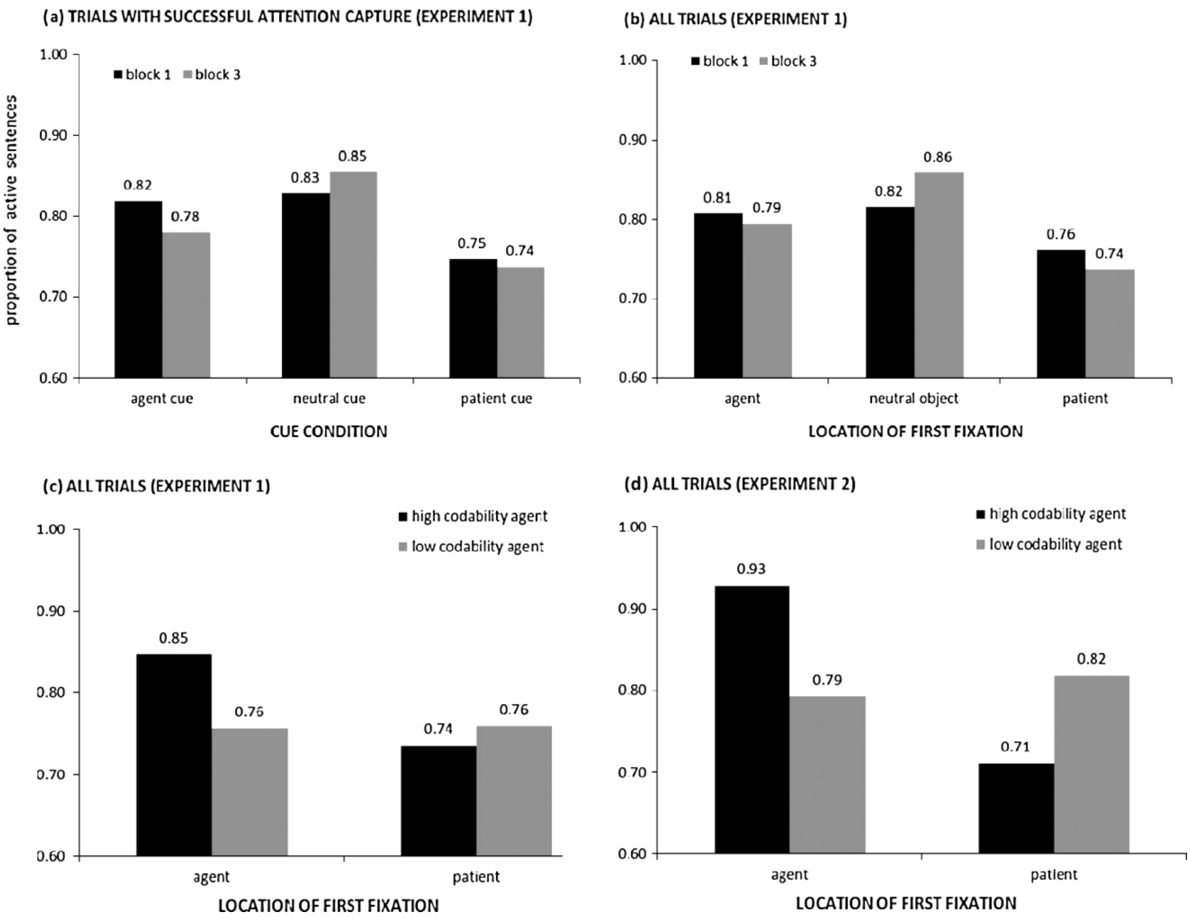

Figure 3. Proportion of active sentences produced in (a)–(b) Experiment 1 when first fixations were directed to agents and patients. Figure (a) shows results for trials with successful attention capture; figure (b) shows results for all trials, irrespective of attention capture. Fixations to the neutral area were not included in the analyses due to small *n*'s. Figures (c) and (d) show the proportions of active sentences produced in Experiments 1 and 2 when first fixations were directed to agents and patients in items with higher- and lower-codability agents.

A complementary analysis was carried out to compare production of actives across trials where speakers fixated agents first and patients first, irrespective of cue condition and attention capture. The results from this larger dataset showed a similar pattern (Fig. 3b), but the effect of

First fixations on sentence form was again not reliable (β = -.19, z = 1.03) and did not differ across blocks (β = .18, z = 1.03). There were also no effects of Event, Agent or Patient codability (all zs < .9): numerically, however, speakers produced more active sentences if they first looked at an easy-to-name agent than a hard-to-name agent (Fig. 3c; this result is reported for comparison with Experiment 2). Thus contrary to the predictions of the strong version of linear incrementality (Gleitman et al., 2007), these results show that, across a wide range of events, starting points are not reliably predicted by the distribution of first fixations.

Structure choice

Sentence form was not predicted by Event codability (z = -.10), so the ease of encoding an event did not directly influence the selection of active or passive syntax. Instead, structure choice was more sensitive to the ease of generating structures and of character naming (Fig. 4a): exposure to passive syntax in the reading task (Block 2) increased production of passive sentences in Block 3 when the agent was easy to name (β = .54, z = 2.96, for the Block by Agent codability interaction). Speakers produce fewer actives when they found the agent difficult to name (low-codability agents), and this tendency did not change from Block 1 to Block 3. Production of passives was not reliably predicted by Patient codability (z = -.58).

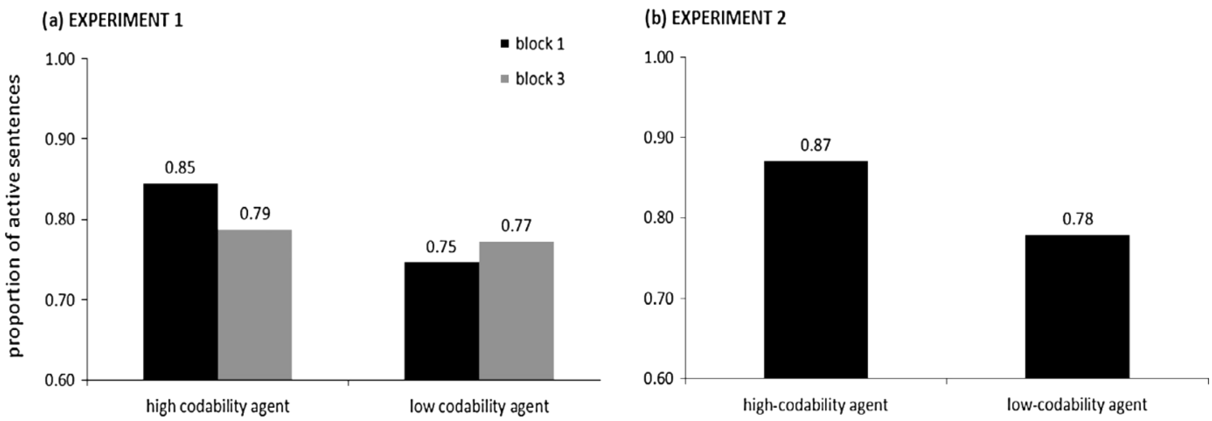

Figure 4. Proportion of active sentences produced with respect to Agent codability in (a) Experiment 1 (across blocks) and (b) Experiment 2.

A separate analysis compared structure choice in sentences with human agents and patients to sentences with non-human agents and patients (animate and inanimate characters were combined into one category, as items with these characters did not differ systematically). As expected, speakers produced more active sentences to describe events with human agents than non-human agents (β = 2.31, z = 4.08, for the main effect of Agent animacy; model fit was improved by including by-participant random slopes for this factor). Similarly, they produced more passive sentences to describe events with human than non-human patients (β = -2.73, z = -4.98, for the main effect of Patient animacy). Importantly, in a model including both character animacy and character codability as predictors of structure choice, Agent and Patient animacy did not change the effects of Agent and Patient codability described above. Therefore, we limit further discussion of character-specific variables in sentence formulation to character codability.

Timecourse of sentence formulation

Timecourse analyses were carried out to test whether the formulation process was sensitive to conceptual properties of the events (Event codability) and to the same variables that influenced structure choice (Agent codability and the ease of structural assembly). To restate the predictions, if the mapping of a pre-verbal message onto language depends on speakers' understanding of the event itself, fixations to the two characters should diverge more slowly in higher-codability events, where speakers should be more likely to encode information about both characters (consistent with hierarchical incrementality) than in lower-codability events, where speakers might quickly prioritize encoding of a single character (consistent with linear incrementality). Event codability should thus also determine the extent to which character-specific properties (i.e., character codability) influence early planning: character codability should play a weak role in high-codability events and a stronger role in low-codability events. If the message-to-language mapping is also sensitive to the ease of structural assembly, differences in formulation should be observed between active and passive sentences as well as between Block 1 and Block 3 of the experiment.

Active sentences

Figure 5 shows the timecourse of formulation for active sentences and speech onsets across item categories and conditions. Analyses were carried out in three steps. Fixations to each character were aggregated into 10-ms bins. The *first* analysis examined differences in gaze patterns within 200 ms of picture onset across item categories and blocks, testing the prediction of linear incrementality that speakers will quickly direct their gaze to the character selected to be the starting point (Gleitman et al., 2007). Fixations to the agent were added up into four 50-ms bins for this analysis. The *second* analysis compared fixations in a wider time window—between 0 and 600 ms after picture onset—after aggregating fixations into three consecutive 200-ms windows. This analysis assessed whether effects seen in the 0–200 ms time window were specific to first fixations or whether subsequent gaze shifts were also sensitive to event properties. The *third* analysis compared fixations to the agent across the two blocks between 600 and 1800 ms (i.e., over six consecutive 200-ms windows until speech onset).

The regression models included Event codability, Agent codability, Block, and Time bin as fixed factors. By-participant analyses were performed after a medial split of items into high and low Event codability and Agent codability; by-item analyses included the two codability variables as continuous predictors. Interactions between predictors were included when they improved model fit and only the highest-order interactions of theoretical interest are discussed below. The results of by-participant and by-item analyses are largely consistent for effects of theoretical interest. For clarity, Agent codability was only included in analyses for active sentences and Patient codability in analyses of passive sentences.

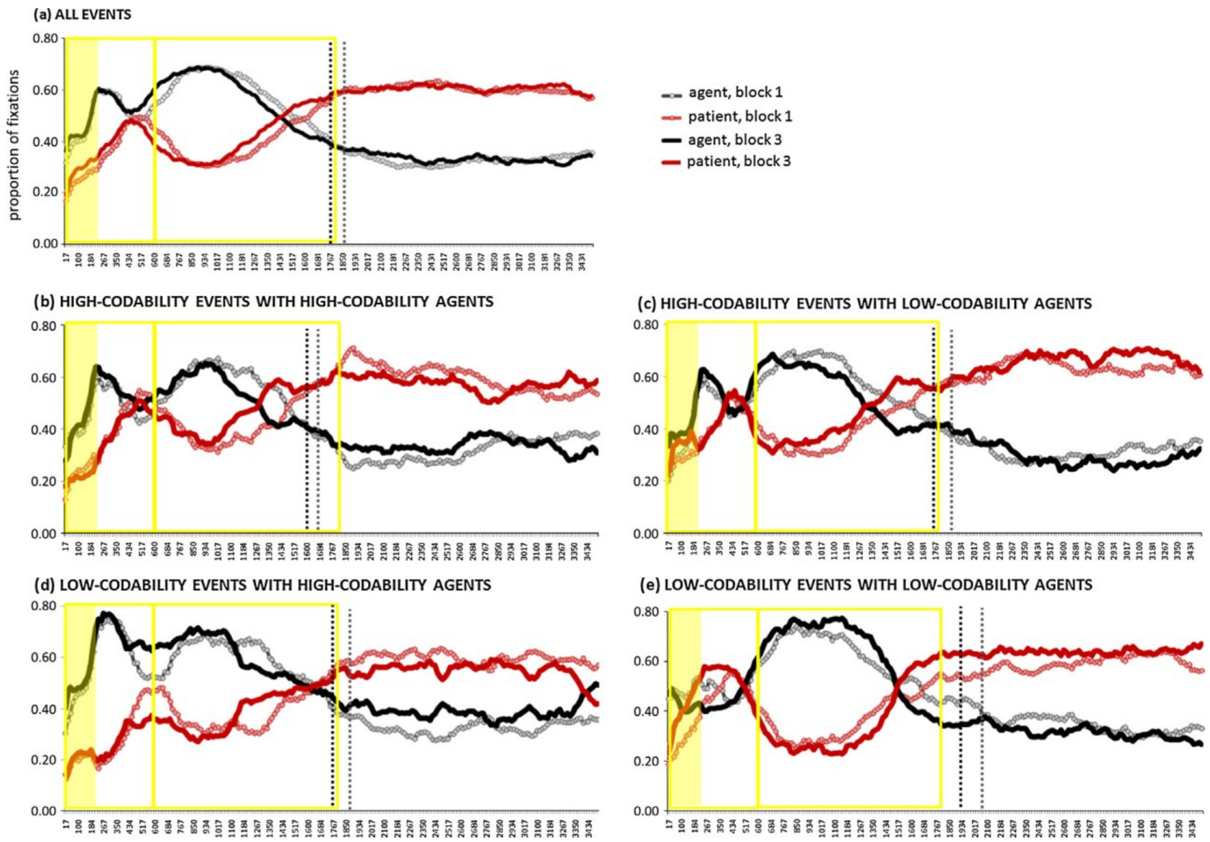

Figure 5. Timecourse of formulation for active sentences in Blocks 1 and 3 of Experiment 1 for all events (a) and separately for higher- and lower-codability events with higher- and lower-codability agents (b)–(e). Yellow boxes show the time windows included in the analyses (0–200 ms, 0–600 ms, 600–1800 ms). The vertical dashed lines indicate speech onsets in Block 1 and 3.

1. Early fixations to the agent (0–200 ms)

Formulation of active sentences (Fig. 5a) generally began with fixations to the agent, but this agent preference was modulated by both Event and Agent codability (Table 2).

Table 2. Results of by-participant and by-item quasi-logistic regressions comparing agent-directed fixations within 200 ms of picture onset in active sentences in Experiment 1. In all tables, (*s*) indicates the inclusion of random slopes.

	By-participants			By-items		
Fixed effects	Estimate	SE	t-value	Estimate	SE	t-value
Intercept	-.68	.06	-11.45*	-.68	.05	-13.30*
Time bin	2.90 *(s)*	.29	10.03*	3.74 *(s)*	.20	19.01*
Block	.17	.09	1.90†	.33	.05	6.37*
Event codability	.69 *(s)*	.10	6.68*	.25	.07	3.49*
Agent codability	.04	.09	.44	-.17	.06	-2.73*
Block * Time bin	-1.49	.48	-3.13*	-2.26	.25	-9.00*
Event cod. * Time bin	-3.28	.46	-7.10*	-1.63	.28	-5.84*
Event cod. * Block	-.44	.17	-2.58*	.23	.07	3.10*
Event cod. * Block * Time bin	-1.87	.92	-2.03*	-.35	.34	-1.01
Agent cod. * Time bin	-2.16	.48	-4.55*	-.23	.24	-.95
Agent cod. * Block	.13	.18	.72	.29	.06	4.60*
Agent cod. * Block * Time bin	-2.76	.95	-2.90*	-2.70	.31	-8.58*
Event cod. * Agent cod.	.52	.17	3.04*	.02	.08	.20
Event cod. * Agent cod. * Time bin	-4.83	.92	-5.23*	-1.13	.33	-3.40*
Event cod. * Agent cod. * Block	-.46	.34	-1.33	-.47	.09	-5.31*
Event cod. * Agent cod. * Block * Time bin	6.07	1.85	3.29*	4.03	.43	9.29*

* $p<.05$ in all tables
† $p<.10$ in all tables

The difference in fixations directed to the agent and patient was relatively small in events that were easy to describe (Fig. 5b and c) and was only weakly modulated by Agent codability. In contrast, the distribution of fixations to the two characters depended strongly on Agent codability in events that were harder to interpret (Fig. 5d and e): speakers were more likely to fixate high-codability than low-codability agents, so the difference in fixations directed to the two characters was *larger* when the agent was easy to name and *smaller* when the agent was difficult to name. This suggests that speakers attempted to begin their sentences with the agent in subject position by default, but considered the patient as an alternative starting point when the agent was difficult to name. The difference in fixations directed to the two characters was also reduced in Block 3 of the experiment as speakers directed more fixations to the patient in this time window in Block 3 than Block 1. These effects jointly resulted in a four-way interaction between Event codability, Agent codability, Time bin and Block. Together, they show that speakers were more likely to engage in *hierarchically* incremental planning when event gist was easy to encode and in *linearly* incremental planning when event gist was hard to encode.

2. Fixations to the agent between 0 and 600 ms

Speakers did not look consistently at the agent after 200 ms, but shifted their gaze briefly towards the patient before fixating the agent again after 600 ms. This pattern suggests that the information encoded within 200 ms of picture onset was likely not sufficient to support continued deployment of attention to the first-mentioned character; instead, it implies that speakers sought information about the second character before continuing to encode a sentence with the agent in subject position.

Comparing fixations to the agent across item categories between 0 and 600 ms showed an interaction between Event codability, Agent codability, and Time bin (Table 3a). When describing higher-codability events, speakers distributed their attention between the agent and patient equally (the proportion of fixations to the agent and patient was approximately .50 between 400 and 600 ms). This is consistent with hierarchical incrementality as it shows that speakers take longer to direct their gaze preferentially to the first-mentioned character in events that are easier to encode and describe. The results for low-codability events also confirm the observation from the first analysis (0–200 ms) that early eye-movements show sensitivity to character-specific properties: speakers maintained fixations on the agent when this character was easy to name and turned their attention to the patient otherwise.

3. Fixations to the agent between 600 and 1800 ms

Linguistic encoding is assumed to begin either after 200 or 400 ms (Gleitman et al., 2007; Griffin & Bock, 2000), and the duration of gazes to a character during linguistic encoding can index the difficulty of lexical retrieval (name-related gazes; Griffin, 2004). However, we also expected the deployment of fixations to the subject character and the timing of later gaze shifts to the second character to be sensitive to properties of the events. Specifically, if formulation begins with the generation of a coarse structural plan for higher codability events, this plan should facilitate adding the second character to the developing sentence and speakers should direct their gaze to the second character more quickly in higher-codability compared to lower-codability events. Conversely, if formulation begins with encoding of one character in lower-codability events, speakers should shift their gaze to the second character only after they have finished encoding the subject character.

Table 3. Results of by-participant and by-item quasi-logistic regressions comparing agent-directed fixations in active sentences in the a) 0-600 ms and b) 600-1800 ms time windows in Experiment 1.

	By-participants			By-items		
Fixed effects	Estimate	SE	t-value	Estimate	SE	t-value
a) 0-600 ms						
Intercept	-.11	.01	-7.41*	-.20	.02	-10.94*
Time bin	.71	.05	14.58*	.88 *(s)*	.08	11.69*
Block	.02	.03	.86	.01	.02	.44
Event Codability	.16	.03	5.73*	.12	.03	4.60*
Agent Codability	-.33	.03	-12.38*	-.20	.02	-9.06*
Block * Time bin	.29	.10	2.95*	.22	.07	3.03* *(a)*
Event cod. * Time bin	-.06	.10	-.60	-.15	.10	-1.43
Agent cod. * Time bin	.40	.10	4.13*	.45	.09	4.91* *(a)*
Event Codability * Agent Codability	-.09	.05	-1.72†	-.08	.02	-3.33* *(a)*
Event Codability * Block	-.06	.03	-1.80†	.03	.02	2.00* *(a)*
Agent Codability * Block	-.11	.03	-3.44*	-.07	.02	-4.71* *(a)*
Event Cod. * Agent Cod. * Time bin	-1.09	.20	-5.59*	-.08	.02	-3.49* *(a)*
b) 600-1800 ms						
Intercept	.55	.02	29.24*	.74	.02	41.10*
Time bin	-.67 (*s*)	.03	-21.08*	-.90 (*s*)	.03	-35.74*
Block	.07	.02	3.07*	.05	.01	3.55*
Event codability	.18	.01	13.06*	-.01	.03	-.45
Agent codability	.17	.02	7.43*	.27	.02	12.26*
Block * Time bin	-.24	.04	-6.15*	-.23	.02	-9.26*
Event cod. * Time bin	--	--	--	.13	.04	3.49*
Event cod. * Block	.19	.02	6.90*	.34	.02	18.74*
Event cod. * Time bin * Block	--	--	--	-.49	.03	-14.95*
Agent cod. * Time bin	-.17	.04	-4.37*	-.08	.03	-2.74*
Event cod. * Agent cod.	--	--	--	.26	.03	8.30*
Event cod. * Agent cod. * Time bin	--	--	--	-.59	.04	-13.51*

(a) this interaction does not improve model fit and is treated as suggestive.

Supporting these predictions, agent fixations between 600 ms and speech onset (approximately 1800 ms) were sensitive both to Agent and Event codability (Table 3b). First, as expected, Agent codability reduced the length of name-related gazes: speakers looked at the agent for less time when describing events with easy-to-name than hard-to-name agents (producing an interaction of Agent codability with Time bin). This effect was independent of Block, indicating that repeatedly describing events with easy-to-name agents did not influence the timing of gaze shifts from one character to another. Second, more importantly, speakers also looked at the agents for less time in high-codability events, and this effect was larger in Block 3 than Block 1 (producing an interaction of Event codability with Block). These results suggest that speakers were able to shift their gaze to the second character and integrate information about this character into the developing sentence *earlier* when they had been able to encode event gist during early formulation, and that they found it progressively easier to do so over the course of the experiment.

Speech onsets

Finally, speech onsets were also sensitive to Agent and Event codability. Speakers initiated sentences with easy to- name agents faster than hard-to-name agents (β = 169, t = 3.85, for the main effect of Agent codability). Consistent with the observation that properties of the agent influenced formulation primarily when the event itself was difficult to encode, the effect of Agent codability on speech onsets was stronger in lower-codability than higher-codability events (β = 118, t = 1.96, for the interaction of Agent and Event codability). Onsets were also faster in Block 3 than Block 1 (β= 107, t = 3.32; model fit was improved by including by-participant slopes for the Block factor).

Passive sentences

Passives are relatively difficult to produce and constituted only 21% of the data (Fig. 6a). For clarity, Figure 6b and c and Figure 6d and e show the timecourse of formulation for higher- and lower-codability events and events with higher- and lower-codability patients separately.

We tested whether formulation of sentences with the dispreferred passive structure changed from Block 1 to Block 3 after exposure to passive syntax in Block 2 (cumulative structural priming). Hierarchical incrementality predicts that exposure to passive syntax should increase the likelihood of speakers beginning formulation by encoding information about two characters (i.e., distributing their attention between agents and patients); linear incrementality predicts that speakers should be primed to consider the patient a suitable sentential starting point. Both accounts therefore predict a reduction in the early preference (0–200 ms) for the agent and an increase in the proportion of fixations directed to the patient in Block 3 relative to Block 1. However, linear incrementality, which assumes that fixation times primarily reflect the ease of encoding individual characters, does not predict additional changes in the gaze pattern beyond 200 ms from Block 1 to Block 3. In contrast, hierarchical incrementality predicts that familiarity with passive syntax should facilitate encoding throughout the formulation process: speakers should find it easier to *continue* building the sentence from their chosen starting point – and hence should look more consistently at the patient – after 200 ms in Block 3 than Block 1.

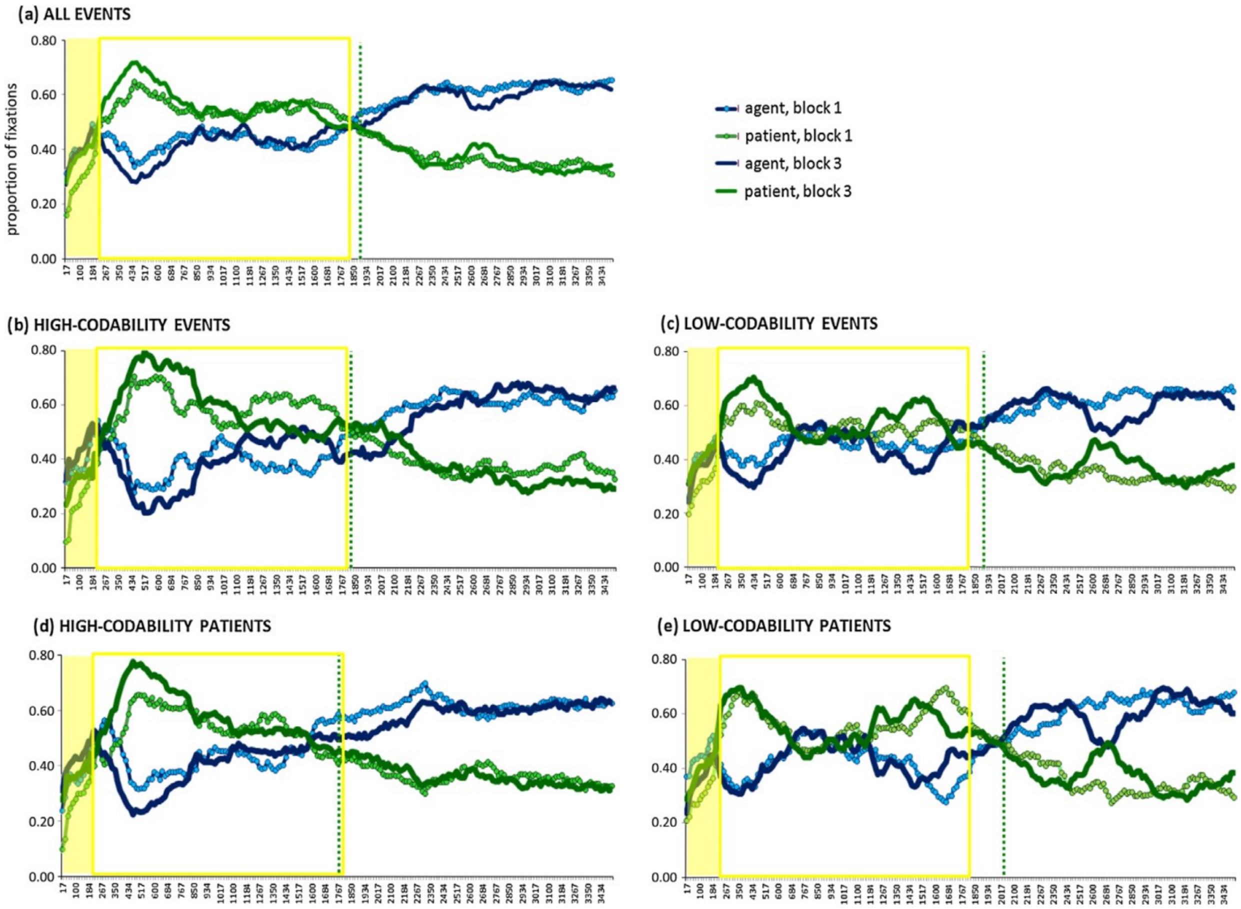

Figure 6. Timecourse of formulation for passive sentences in Blocks 1 and 3 of Experiment 1 for all events (6a) and separately for higher- and lower-codability events (b and c) and higher- and lower-codability patients (d and e). The vertical dashed lines indicate speech onsets collapsed across Blocks.

Analyses were carried out in two steps. The *first* analysis tested whether the pattern of early fixations to the patient (0–200 ms) differed in Blocks 1 and 3. Due to sparse data, the by-participant analysis included only Event codability and the by-item analyses included both Event and Patient codability as predictors (the by-participant analysis requires subdividing the data into high- and low-codability item categories based on a median split, which reduces statistical power). The *second* analysis compared fixations to the patient across the two blocks between 200 and 1800 ms, after aggregating the data into eight consecutive 200-ms windows (this is a wider time window than in the analysis of active sentences as the data do not warrant an additional comparison of fixations between 0 and 600 ms). The by-participants and by-items analyses included both Event and Patient codability as predictors.

1. Early fixations to the patient (0–200 ms)

Production of passives in Block 1 began with fixations to the agent (Fig. 6a; Table 4a). Early fixations did not show sensitivity to Event and Patient codability (Fig. 6b–e). Importantly, the early preference for fixating the agent was reduced after exposure to passive syntax in Block 2: speakers directed more early fixations to patients in Block 3 than Block 1 (resulting in a main effect of Block; the effect was reliable only in the by-item analysis).

2. Fixations to the patient between 200 and 1800 ms

Unlike formulation of active sentences, speakers continued fixating the first-mentioned character (the patient) between 200 ms and 600 ms. The first indication that speakers began shifting their attention and gaze back to the agent occurred relatively late (approximately 800 ms after picture onset).

The distribution of fixations to the patient between 200 and 1800 ms was predicted by both Event codability (Fig. 6b–c) and Patient codability (Fig. 6d–e). Speakers spent more time fixating patients in higher-codability than lower-codability events and in events with higher-codability than lower-codability patients (see Table 4b for main effects). The effect of Event codability suggests that, in higher-codability events, speakers may have encoded enough information about the event as a whole after picture onset (0–200 ms) to warrant selection of the patient as the sentential starting point on conceptual grounds. This likely increased the likelihood of speakers fixating the patient more consistently between 200 ms and speech onset. In contrast, when describing lower-codability events, speakers were more likely to distribute their attention between the patient and the agent, suggesting that they were less certain that they could continue their sentence from their chosen starting point and that they continued planning what they wanted to communicate well after 200 ms. The effect of Patient codability shows that gaze shifts from the patient to the agent occurred earlier when patients were easy to name.

Results concerning the timing of gaze shifts from the patient to the agent with exposure to passive syntax are less clear. Patient gazes were generally longer in Block 3 than Block 1 (producing a main effect of Block), so facilitating production of passive sentences reduced the degree of gaze fluctuations. Interactions of Block with Event codability were only present in the by-item analyses and are treated as suggestive.

Table 4. Results of by-participant and by-item quasi-logistic regressions comparing patient-directed fixations in passive sentences a) within 200 ms of picture onset and b) between 200 and 1800 ms after picture onset in Experiment 1. In model (a), Patient codability was only included in the by-item analysis. In model (b), Event and Patient codability were included in both the by-participant and by-item analyses.

	By-participants			By-items		
Fixed effects	Estimate	SE	t-value	Estimate	SE	t-value
a) 0-200 ms						
Intercept	-1.18	.12	-9.59*	-1.33	.11	-12.48*
Time bin	3.69 *(s)*	.68	5.44*	3.97 *(s)*	.50	7.97*
Block	.37	.23	1.58	.49	.18	2.66*
Event codability	.31	.24	1.33	-.02	.13	-.14
Patient codability	--	--	--	-.01	.13	-.01
b) 200-1800 ms						
Intercept	.19	.03	6.07*	.26	.03	8.89*
Time bin	-.09	.04	-2.60*	-.03 *(s)*	.03	-.99
Block	.08	.04	2.25*	.20	.04	4.60*
Event codability	-.12	.04	-3.10*	-.17	.04	-4.84*
Patient codability	-.12	.06	-1.94†	-.11	.04	-3.17*
Block * Time bin	--	--	--	-.23	.06	-3.71*
Event cod. * Time bin	--	--	--	.19	.04	4.89*
Event cod. * Block	--	--	--	-.34	.05	-6.84*
Patient cod. * Time bin	.30	.07	4.21*	.20	.04	5.08* *(a)*
Event cod. * Block *Time bin	--	--	--	.58	.07	7.73*

(a) this interaction does not improve model fit and is treated as suggestive.

Speech onsets

Analyses of speech onsets show a joint influence of Patient codability, Event codability, and Block (β = 264, t = 3.22, for the three-way interaction between these factors). Speakers initiated sentences with easy-to-name patients faster than sentences with hard-to-name patients. Consistent with the prediction that properties of the patient should influence formulation primarily when the event itself is difficult to encode, the effect of Patient codability on speech onsets was stronger in lower-codability than higher-codability events in Block 1. However, this effect was absent in Block 3, where onsets were faster in all events with high-codability patients than low-codability patients. In other words, speech onsets show that exposure to passive syntax led speakers to prioritize encoding of patients before speech onset; thus, effects of Event codability on formulation were observable only in eye movements and not in speech.

2.3.3 Discussion

Experiment 1 considered the effects of multiple variables on structure choice and sentence formulation to identify the conditions under which speakers might make use of a more linearly incremental or hierarchically incremental planning strategy. As in earlier work, we tested the effects of early attention shifts as a first indication of what type of information speakers use to select a starting point for their sentences. We then compared the entire timecourse of formulation for different types of items and sentences.

Effects of early gaze shifts on sentence form test the strong version of linear incrementality, and Experiment 1 showed little support for this account. The subliminal cuing manipulation we used was analogous to that of earlier studies (Gleitman et al., 2007; Kuchinsky & Bock, 2010), but cuing did not systematically influence structure choice. Supporting a weaker version of linear incrementality, speakers were more likely to assign a first-fixated agent to subject position if this character was easy to name than when it was harder to name. The comparison of effects of gaze shifts on selection of starting points against the influence of a higher-level variable like character codability suggests a hierarchy of factors shaping formulation from its earliest stages, with higher-level conceptual and linguistic factors playing a stronger role in this process than lower-level perceptual salience.

Looking beyond analyses of first fixations, speakers showed a strong bias for selecting the agent as the default starting point. Agents were assigned to object position primarily when they were difficult to name, consistent with earlier demonstrations of conceptual and lexical accessibility effects on sentence form (e.g., Bock & Warren, 1985). Since selection of a character as a starting point on the basis of its role in the event (i.e., agenthood) implies some degree of encoding of the who-did-what-to-whom event structure, these findings support a weak version of both linear and hierarchical incrementality: the accessibility of a single character can bias structure choice and facilitate production, but agents appear to play a more prominent role than patients.

Timecourse analyses largely confirmed conclusions about the relative mediation of sentence formulation by higher-level and lower-level factors. Formulation of active sentences showed evidence of hierarchically incremental planning within 600 ms of picture onset when events were easy to describe (the difference in early agent-directed and patient-directed fixations was relatively small), and evidence of linearly incremental planning when events were harder to describe (the distribution of early fixations to the two event characters depended strongly on properties of the subject character). Effects of event and agent codability carried over to later windows as well (between 600 ms and speech onset), where differences between higher- and lower-codability events were observable as differences in the timing of gaze shifts to the second character.

We also observed a change in formulation of active sentences across blocks. Speakers shifted their gaze to the second character earlier in Block 3 than Block 1, and this effect was modulated by Event codability: shifts of gaze occurred earlier in higher-codability events than lower-codability events in Block 3 than Block 1. Speakers naturally gained experience producing active sentences in this paradigm as the experiment progressed, so these effects show that repeated use of a sentence structure can create a shift towards hierarchically incremental planning.

Formulation of sentences with the dispreferred passive structure was substantially different. Here, only later fixations (200–1800 ms) showed sensitivity to differences in Event and Patient codability across items: speakers looked more consistently at the subject character (the patient) in events that were easy to interpret and where the patient was easy to name. The timing of gaze shifts from patients to agents was only predicted by Patient codability, consistent with linear incrementality. Importantly, formulation of passives changed with exposure to passive syntax. Speakers were more likely to fixate patients at the outset of formulation in Block 3 relative to Block 1, and changes in the proportion of fixations to patients to agents across blocks were weakly modulated by Event codability. These changes suggest a weak shift towards hierarchically incremental planning.

2.4 Experiment 2

In the next experiment, we limit our investigation of flexibility in sentence formulation to the factors that exerted the strongest influence on planning of *active* sentences in Experiment 1: codability of events and codability of agents. Since Experiment 1 used a cuing manipulation to direct speakers' gaze to selected characters at picture onset, it is important to replicate the timecourse results with a paradigm that does not influence the distribution of early fixations with exogenous cues. In addition, while codability scores were calculated post hoc based on the distribution of responses in the first experiment, in Experiment 2 we used only items with the highest and lowest codability scores selected from earlier experiments. Properties of the patient were weak predictors of formulation in Experiment 1, so patient codability was held constant across the four item categories.

Thus taking only higher-level variables into account, we focus on the contribution of one factor affecting conceptual- structural encoding (the relationships between event characters) and one affecting encoding of discrete pieces of information (individual characters). As before, we predict that the extent to which formulation is linearly or hierarchical incremental will depend on the ease of encoding an event structure, and that properties of individual characters (specifically, properties of agents) should play a role in formulation only when events are difficult to encode.

2.4.1 Method

Participants

A new group of 40 adult native speakers of Dutch participated for payment.

Materials, design, and procedure

Forty-eight pictures displaying transitive events were selected on the basis of codability scores obtained in Experiment 1 and three similar experiments conducted in the lab. We selected items with consistently high and low Event codability and Agent codability scores across experiments (some with new characters). The final item set thus consisted of four categories of items: 24 items showing high-codability events (12 with high- and 12 with low-codability agents) and 24 items showing low-codability events (12 with high- and 12 with low-codability agents).

To evaluate speakers' perceptions of Event and Agent codability in the current experiment, codability scores were also calculated post hoc from speakers' responses in the current experiment: as expected, high-event codability items had higher codability scores than low-event codability items, $t(1,45) = 4.76$, and items with high-codability agents received higher agent codability scores than items with low-agent codability, $t(1,45) = 4.97$. However, since assignment of items to high- or low-codability categories depends on the distribution of responses within an experiment, the codability ratings calculated post hoc did not correspond exactly to the expected codability ratings (some of the items in the expected high or low codability categories now turned out to be relatively low or high in codability). The analyses of Experiment 2 were thus based on the codability values calculated from the responses given by the participants of this experiment since these values best reflected the processing difficulty of the items for this group of participants (see Table 1 for means and Appendix B for a full list).

Event codability and Agent codability were again not correlated in this item set ($r = -.04$). Patient codability did not differ systematically between the four item categories and did not correlate with Agent codability ($r = .09$; there was a weak correlation with Event codability, $r = -.25$, $p < .10$).

The animacy of the agent and patient characters was balanced across the four item categories. There were 30 pictures with human agents, 10 with animate agents, and 8 with inanimate agents; 22 of these pictures included human patients, 7 included animate patients, and 19 included inanimate patients.

Target pictures were interspersed among 96 filler pictures, organized such that there was no semantic overlap or repetition of content words in adjacent pictures. This list of items was then divided into two blocks, and the order of the blocks was counterbalanced across participants. Placement of the agent on the left and right hand-side of the screen was again controlled by counterbalancing mirror-reversed versions of each picture across experimental lists, yielding four stimuli lists in total. Each participant saw all 48 target pictures and 96 filler pictures, and the procedure used to elicit descriptions was the same as in Experiment 1.

Sentence scoring and analyses

The scoring criteria were identical to those in Experiment 1. Excluding incorrect responses (i.e., non-transitive constructions, sentences with indefinite pronouns and restarts), responses with long speech onsets and trials where first fixations fell into the agent or patient interest area left 931 sentences (83% actives and 17% passives) for analysis. Truncated passives were not included in the analyses.

Analyses of structure choice were conducted as in Experiment 1, with Event codability and Agent codability being included in the models as continuous predictors. Analyses of first fixations were again carried out to test the strong version of linear incrementality. We report effects of Agent and Patient animacy on structure choice briefly as they were largely consistent with the results of Experiment 1. Timecourse analyses are reported only for active sentences due to the low rate of production of passives.

2.4.2 Results

First fixations and structure choice

The majority of first fixations on scored trials were directed to the agent (.70). Unlike Experiment 1, first fixations were reliable predictors of sentence structure: speakers produced more actives when they looked first at the agent (.87) than when they looked first at the patient (.76; $\beta = .99$, $z = 2.97$).

The location of First fixations interacted with Agent codability to determine sentence form (Fig. 3c; $\beta = 1.04$, $z = 2.03$, with random by-participant slopes for Agent codability). Speakers produced more actives to describe events with easy-to-name agents when they first fixated the agent than when they first fixated the patient. Production of actives in events with harder-to-name agents did not vary as a function of the location of First fixations. This finding is similar to the results of Experiment 1 in showing that the selection of starting points reflects a combined influence of lower-level and higher-level factors: speakers are more likely to assign a first-fixated character to subject position if it is easy to encode linguistically. Event codability did not predict production of actives ($z < -.25$), and there was no interaction between Event and Agent codability.

A separate analysis considered the effects of Agent and Patient animacy on structure choice. The results were again similar to those of Experiment 1: speakers produced more active sentences to describe events with human agents than non-human agents (.88 vs. .76; $\beta = 1.44$, $z = 2.48$, for the main effect of Agent animacy), and more passive sentences to describe events with human patients than non-human patients (.77 vs. .89; $\beta = -1.64$, $z = -2.88$, for the main effect of Patient animacy). The two factors did not interact ($z = .37$). A full model including both codability predictors (Event and Agent codability) and animacy (Agent and Patient animacy) again showed the expected main effects of Agent codability, Agent animacy, and Patient animacy (all zs > 1.74 in an additive model), but no interactions between these variables. As expected, properties of the patient exerted only a weak influence on structure choice, so the effect of Patient codability did not reach significance.

Timecourse of formulation for active sentences

Analyses were carried out for three time windows selected based on visual inspection of the distribution of agent and patient fixations before speech onset (Fig. 7; the windows we chose are somewhat different from those in Experiment 1). Formulation began with a brief period where speakers first fixated the agent and then the patient. Thus the *first* analysis compared the proportions of agent-directed fixations within 400 ms of picture onset across the four item categories (after aggregating data into eight 50-ms bins for a sensitive comparison across codability categories). Speakers then briefly looked back at the patient, so the *second* analysis examined differences in gaze patterns across item categories between 0 and 700 ms after picture onset (after aggregating data into seven 100-ms bins). The *third* analysis examined the difference in gazes to agents between 700 ms and speech onset (1900 ms; after aggregating data into six 200-ms bins). All models included Event codability, Agent codability and Time bin as fixed factors.

1. Early fixations to the agent (0–400 ms)

Early fixations showed an immediate preference for the agent of the event (Fig. 7a). This preference was again modulated by Event codability (Fig. 7b, c vs. Fig. 7d and e). The difference in the proportion of agent-directed fixations before 400 ms was smaller in higher- than lower codability events (see Table 5a for the interaction of Event codability with Time bin), confirming that early formulation varied with the ease of encoding event gist. Speakers were also more likely to fixate high-codability than low codability agents. The effect of Agent codability was again negligible in higher-codability events (Fig. 7b and c), but the interaction between Event and Agent codability did not reach significance.

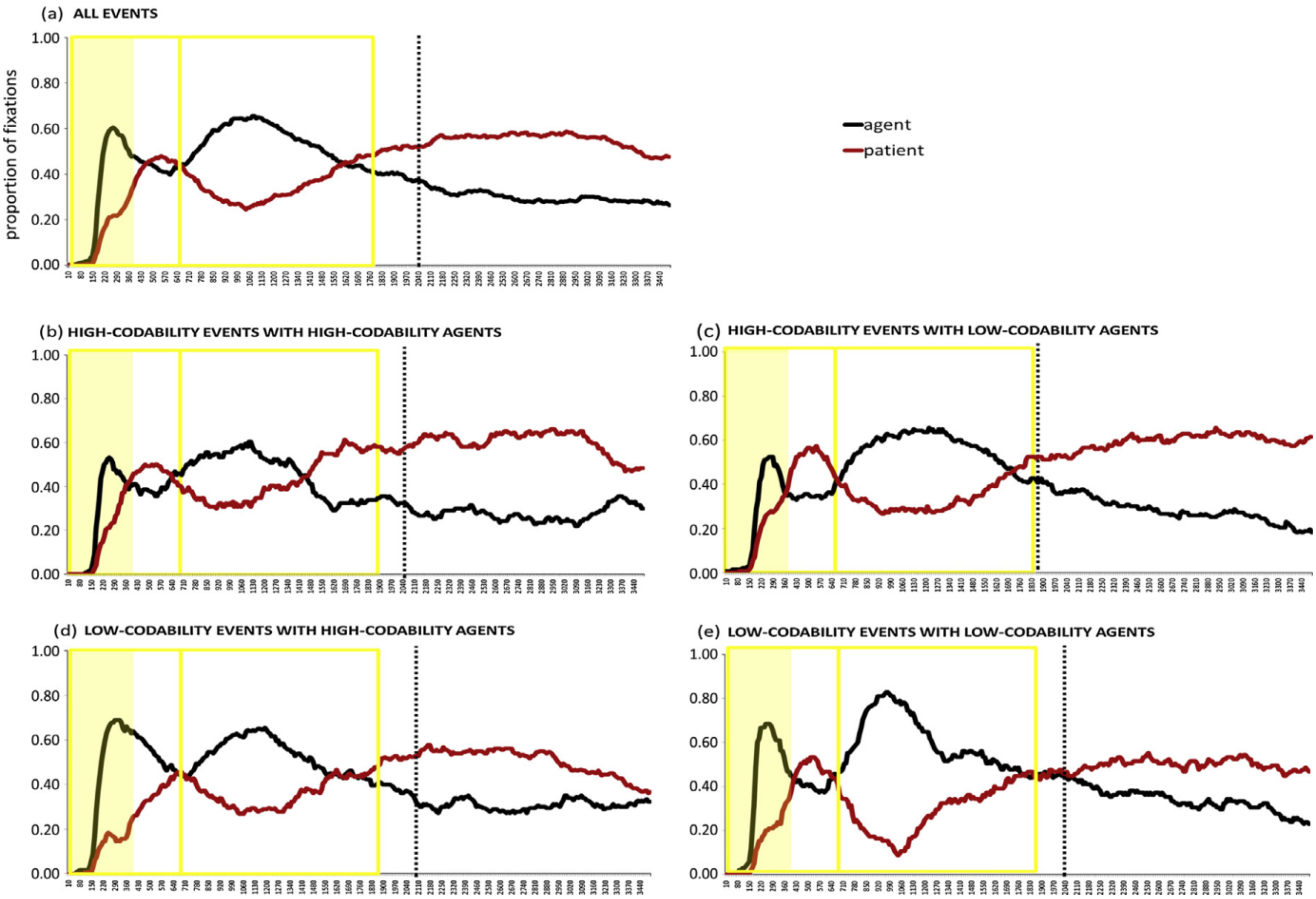

Figure 7. Timecourse of formulation for active sentences for all events in Experiment 2 (7a) and separately for higher- and lower-codability events with higher- and lower-codability agents (b)–(e). Yellow boxes show the time windows included in the analyses (0–400 ms, 0–700 ms, 700–1900 ms). The vertical dashed lines indicate speech onsets.

2. Fixations to the agent between 0 and 700 ms

Speakers looked briefly at the patient after 400 ms, suggesting that they preferred to encode some information about this character before continuing to formulate a sentence with the agent in subject position. The distribution of fixations varied with both Event and Agent codability (Table 5b): speakers were generally less likely to fixate agents in high-codability than low-codability events, but more likely to fixate agents when agents were easy to name than when they were harder to name. Despite the lack of interactions, the results again illustrate that the ease of encoding event gist and the ease of naming characters have opposite effects on the formulation process.

Table 5. Results of by-participant and by-item quasi-logistic regressions comparing agent-directed fixations in active sentences across three time windows in Experiment 2.

	By-participants			By-items		
Fixed effects	Estimate	SE	t-value	Estimate	SE	t-value
a) 0-400 ms						
Intercept	-2.28	.06	-37.78*	-2.12	.06	-34.19*
Time bin	8.45 *(s)*	.22	38.21*	7.77 *(s)*	.30	25.50*
Event codability	.05	.12	.47	-.32	.13	-2.51*
Agent codability	.06	.06	.94	.55	.13	4.34*
Event cod. * Time bin	1.78	.44	4.03*	3.25	.62	5.27*
Agent cod. * Time bin	--	--	--	-2.85	.62	-4.64*
b) 0-700 ms						
Intercept	-.77	.03	-24.52*	-.69	.03	-21.00*
Time bin	1.40 *(s)*	.09	16.18*	1.11 *(s)*	.10	11.39*
Event codability	.33	.03	10.19*	.24	.04	8.45*
Patient codability	-.09	.03	-2.79*	-.12	.04	-3.12*
b) 700-1900 ms						
Intercept	.35	.02	20.03*	.51	.02	29.45*
Time bin	-.53 *(s)*	.03	-17.17*	-.74 *(s)*	.03	-26.34*
Event codability	.13	.02	7.29*	.17	.02	9.38*
Patient codability	.25	.02	13.51*	.33	.02	18.03*

3. Fixations to the agent between 700 and 1900 ms

As in Experiment 1, differences in Event and Agent codability across items influenced the distribution of fixations to the two characters between 700 ms and speech onset as well (Table 5c). Speakers spent less time fixating agents in high-codability than low-codability events, and less time fixating high-codability than low-codability agents. Despite the lack of interactions, the results again show that speakers began adding the second character to the sentence earlier in events that were easy to describe and when they needed less time to encode the first character.

Speech onsets

Speakers initiated descriptions of events with high codability agents more quickly than descriptions of events with low-codability agents (β = 134, $z > 2.08$; model fit was improved by adding random by-participant slopes for this factor). Event codability had no effect on speech onsets (z = .77), and the two factors did not interact.

2.4.3 Discussion

Experiment 2 replicates the main findings of Experiment 1 using a sample of high- and low-codability items. We again compared the effects of lower-level predictors (testing the strong version of linear incrementality) and higher-level predictors (testing the strong version of hierarchical incrementality) on structure choice and on the timecourse of formulation.

As in Experiment 1, analyses of first fixations were carried out to compare conceptual and linguistic accessibility effects against the effects of shifts of visual attention during early inspection of the pictured events. First fixations predicted structure choice in the current experiment, but the likelihood of a first-fixated character being assigned to subject position depended on the codability of the agent: these results are again consistent with a weak version of linear incrementality as they shows effects of both lower-level perceptual factors and higher-level conceptual and linguistic factors on selection of starting points.

More importantly, timecourse analyses again showed a joint influence of Event and Agent codability on early and late fixations to agents and patients in active sentences. Formulation of high-codability events began with speakers distributing their attention between agents and patients, while formulation of low-codability events began with priority encoding of the agent. Event and Agent codability continued to influence the pattern of fixations until speech onset: speakers shifted their gaze away from the agent and towards the patient more quickly in high-codability events and events with high-codability agents.

2.5 General discussion

Two experiments examined the timecourse of sentence formulation across a range of two-character transitive events that were easier or harder to interpret and describe linguistically. In a flexible production system, the form and content of such sentences, as well as the process of assembling sentences incrementally online, can be sensitive to a range of conceptual and linguistic factors. Here we evaluated the extent to which the mapping of messages onto language depends on the perceptual salience and ease of naming of individual characters on the one hand,

and on the ease of performing the conceptual and linguistic operations that encode the who-did-what-to-whom content of the event on the other hand. At issue is the difference between linear incrementality, or theories of sentence formulation that postulate a simple mapping between the uptake of discrete pieces of visual information and linguistic encoding (Gleitman et al., 2007), and hierarchical incrementality, or theories that postulate more wholistic encoding of event gist prior to linguistic encoding (Griffin & Bock, 2000).

To test the strong version of linear incrementality, Experiment 1 employed a manipulation of perceptual salience that was meant to draw speakers' attention to the agent or patient character in target events at picture onset, and we examined whether shifts of attention biased assignment of cued characters to subject position (Gleitman et al., 2007; Kuchinsky & Bock, 2010). A weaker version of linear incrementality was tested by varying the ease of naming the individual characters (i.e., character codability) in pictured events in both experiments. To test the strong version of hierarchical incrementality, the timecourse of sentence formulation was evaluated with respect to speakers' ability to generate a rudimentary conceptual framework or structural plan for the sentence shortly after picture onset. On the assumption that generating this plan requires encoding information about the relationship between characters in an event, we compared the timecourse of formulation across pictures where the event gist was easier or harder to encode and across sentences where a linguistic structure was easier or harder to assemble.

The results showed three findings of interest. *First*, subliminal cues in Experiment 1 did not reliably predict production of active and passive descriptions. An analogous pattern was obtained in analyses evaluating the relationship between the location of first fixations (irrespective of attention capture) and structure choice. In Experiment 2, the location of first fixations predicted structure choice only when the first-fixated character was easy to encode linguistically. Although there were no interactions with event codability as in Kuchinsky and Bock (2010), the results show that selection of a starting point was only weakly influenced by lower-level factors like attention capture or the serial order of visual inspection of characters in a display. This provides evidence against the strong version of linear incrementality (Gleitman et al., 2007): early gaze shifts influenced assignment of a character to subject position if this character was a good candidate for a sentential starting point on the basis of higher-level properties, but did not influence structure choice directly.

Second, structure choice was more sensitive to Agent codability than Patient codability, likely due to the overall preference for active syntax over passive syntax. From the perspective of formulation, the influence of character-specific properties on structure choice supports a version of linear incrementality where the availability of discrete pieces of linguistic information (character names) biases assessments of their suitability for sentential starting points. At the same time, the strong bias for choosing the agent as the default starting point of a sentence, rather than choosing the most accessible character as the default starting point, implies mediation of this process by a conceptual framework. The simplest account of these results is therefore one where speakers quickly extract enough relational information about the event to identify the most agent-like character and immediately verify its suitability for the subject role by attempting to retrieve its name.

Third, the key finding of these experiments concerns the joint influence of factors modulating encoding of structural information (event structure and linguistic structure) on the timecourse of formulation for active and passive sentences. Overall, speakers directed their gaze to the agent very quickly after picture onset: i.e., consistent with the strong, agent-first bias observed in the analyses of structure choice, speakers gave priority to agents during the early stages of sentence formulation. Critically, however, this pattern was found to vary across items with differences in event codability and ease of structural assembly, and effects were observable *throughout* the formulation process.

In *active* sentences, rapid selection of a starting point was modulated by event codability: speakers were more likely to encode information about both the agent and patient in high-codability events (consistent with hierarchical incrementality) than in low codability events, where they attempted to encode information about the most accessible character first (consistent with linear incrementality). Comparable effects were observed immediately after picture onset (within 400 ms of picture onset) and in a wider time window (600–700 ms post-picture onset). This is consistent with Kuchinsky and Bock's (2010) observation that event codability modulates the influence of early attention shifts on selection of starting points, and extends it to describe the influence of codability on the timecourse of sentence formulation itself: speakers begin sentence formulation by generating a coarse conceptual framework when they can quickly decide what they want to say about the event and rely on properties of the two characters to begin formulation otherwise.

Importantly, the benefits of generating such a framework shortly after picture onset extended to later time windows (between 600–700 ms and speech onset), i.e., to the part of the formulation process where gaze durations and gaze shifts are assumed to primarily index linguistic encoding of the fixated characters (Griffin, 2004; Griffin & Bock, 2000). In both experiments, the timing of gaze shifts from agents to patients before speech onset in active sentences depended only in part on the ease of encoding the agent: speakers shifted their gaze towards the patient more quickly in higher-codability than lower-codability events, suggesting that they began incorporating information about the patient into the sentence earlier when formulation was supported by a conceptual framework. In sum, speakers used different planning strategies to formulate their event descriptions: easy encoding of event gist facilitated generation of a conceptual plan that guided further formulation, while difficulty in encoding the causal structure of an event promoted selection of a more linearly incremental planning strategy.

Experiment 1 showed weaker evidence of event codability influencing formulation of sentences with the dispreferred passive structure. The pattern of fixations to the two characters was generally less consistent than in active sentences, and we can attribute high variability in fixations partly to sparse data and partly to the overall difficulty of producing passives: repeated gaze shifts from one character to another indicate that speakers likely continued planning what they would say even as late as during linguistic encoding of the patient. On this interpretation, high event codability simply aided the deployment of gaze to the two characters in the order of mention, and high patient codability supported sustained fixations on the patient character.

Altogether, these results show that the ease of executing an intended message-to-language mapping had a strong impact on the timecourse of sentence formulation, but with different implications for the formulation of sentences with active and passive syntax. Formulation of active sentences appeared to be more hierarchically incremental than formulation of passive sentences. This difference between structures may not be surprising since passives were produced primarily when the preferred mapping of agents to subject position was difficult to execute. A more direct test of how the ease of structural assembly influences sentence formulation was thus provided by the inclusion of a cumulative priming manipulation in Experiment 1 to facilitate generation of passives.

As expected, passive descriptions were produced more often after the reading task than before it, and this cumulative structural priming effect was accompanied by a shift away from the agent-first bias in early fixations (0–200 ms): speakers were more likely to direct attention to the patient at picture onset in Block 3 than Block 1, thereby reducing the difference in fixations to the two characters in this time window. Speakers also directed more fixations to the patient after 200 ms in Block 3 than Block 1, again indicating that they were more confident in continuing to build their sentence from their chosen starting point when they could assemble a passive structure more easily.

The shift in the distribution of early fixations to the agent and patient in passive sentences may be interpreted as supporting hierarchical incrementality as it implies that speakers encoded information about both characters during early formulation. However, this effect may also have occurred, for example, because of increased activation of the patient's thematic role (the *theme*) and thus a higher probability of speakers considering the patient as a possible sentential starting point. Supporting this conclusion is the fact that a comparable increase in early patient-directed fixations was observed in active sentences with lower- codability agents, where speakers briefly directed more fixations to the patient in Block 3 than Block 1 of Experiment 1 in an effort to find an alternative starting point. Such shifts in the timecourse of sentence formulation in Block 3 suggest that facilitating the assembly of an infrequent structure via cumulative structural priming may have influenced the process of selecting starting points in a non-hierarchical fashion.

Nevertheless, if sentence formulation is sensitive to ease of structural assembly itself, then the process of building sentences with infrequent structures should follow a learning trajectory from linear to hierarchical incrementality. In this case, increasing the viability of patients as starting points with exposure to passive syntax may be a first step in creating a shift away from reliance on a linearly incremental planning strategy: repeated assignment of a patient to subject position can provide speakers with the necessary experience of mapping a theme (i.e., a character whose role is defined by virtue of its relationship to the agent in the event) onto a relationally-defined slot in a syntactic structure to eventually begin mapping larger conceptual units (an event representation) onto larger structures. In other words, repeated use of passive syntax may first facilitate the assignment of the patient to subject position and, given time, may begin to facilitate the formulation of a framework that supports encoding of material needed to *complete* a sentence with the patient in subject position.

There is support for this prediction in three places. First, Experiment 1 showed changes in formulation of active sentences from Block 1 to Block 3 in late time windows (speakers shifted their gaze to the second character earlier in Block 3). The presence of this effect for sentences with the *preferred* structure suggests that the effect of experience with production of particular sentence types is to promote hierarchical rather than linear planning. Second, a similar (albeit weaker) effect of event codability on gaze shifts from patients to agents before speech onset was observed in passive sentences in Block 3 compared to Block 1. Third, learning a novel message-to-language mapping has been shown to result in hierarchically incremental planning in simpler messages. For example, Kuchinsky, Bock, and Irwin (2011) tracked the formulation of time expressions when speakers told time from a series of clock displays using either familiar expressions (e.g., *two forty*) or ''inverted'' expressions (*eight ten* to describe the same clock display). Their results showed that, when producing the novel expressions, speakers quickly directed their gaze to the first-mentioned hand of the clock (the longer minute hand) rather than to the first-mentioned hand of the clock in conventional time-telling (the shorter hour hand), and that they spent more time planning the first term (the first hand of the clock) than the second term (the second hand) of the expression before speech onset. This implies that they had quickly learned the novel message-to-language mapping and prepared a structural framework for their responses early during formulation. Thus, it is possible that developing sufficiently abstract mapping procedures to support hierarchical planning in more complex sentences (such as descriptions of two-character events) also requires higher levels of exposure or consistent use of a syntactic structure over longer time frames (see Chang, Dell, & Bock, 2006, for simulations of the acquisition of syntactic structures).

2.6 Implications and conclusions

Differences in gaze-speech coordination observed across the different types of events used in these studies suggest that the timecourse of message and sentence formulation depends on how easily speakers decide what they want to communicate about the event and how easily they can convey this information linguistically. An important parameter in this message-to-language mapping is therefore the degree to which formulation is supported by a conceptual or linguistic structure that facilitates early planning and implementation of this plan. Incremental encoding of messages and sentences in small, independent, single-character units may be inevitable if speakers find it hard to encode a structure or a framework expressing the relational content of an event early during sentence formulation.

At the same time, the results highlight that there may be no ‘‘default’’ form of incrementality in a flexible production system where speakers can make use of multiple sources of information during sentence formulation. Although the strong version of linear incrementality was not supported by these results because shifts of visual attention were poor predictors of structure choice, timecourse analyses showed relatively rapid selection of a starting point on the basis of character-specific information (also see Kuchinsky & Bock, 2010). Similarly, the strong version of hierarchical incrementality cannot account for effects of character codability, but predicts selection of agents as default starting points on the basis of fast encoding of event gist and predicts differences in early and late gaze-speech coordination across higher- and lowercodability events. The process of sentence formulation may therefore be described best as involving a balance of linearly and hierarchically incremental planning, with shifts from linear to hierarchical incrementality, or vice versa, reflecting changes in the efficiency of various encoding operations. This flexibility may facilitate production by allowing speakers to rapidly integrate whatever conceptual and linguistic information is at their disposal as they formulate their messages and sentences.

Acknowledgments

We would like to thank Annelies van Wijngaarden and student assistants from the Psychology of Language Department, especially Moniek Schaars and Samantha van den Hoogen, for help with data collection and processing.

Appendix A

Target pictures showing two-participant events in Experiment 1 grouped by Event and Agent codability category (following a median split of items on each dimension).

Items with high event codability, high agent codability

Ambulance hitting a car
Lightning striking a church
Ducklings following a mother duck
Fireman saving a child
Thief stealing a painting
Ball hitting a boy
Frog catching a fly
Stork bringing a baby
Horse kicking a cow
Boxer hitting a cheerleader
Girl pushing a boy on a sled
Girl tickling a boy
Snake frightening a girl
Mother feeding a baby
River breaking a dam
Avalance approaching the skiers

Items with low event codability, high agent codability

Santa dragging a Christmas tree
Wind blowing away papers
Bomb hitting a ship
Elephant lifting a clown
Shark eating a man
Alarm waking up a boy
Cat catching a mouse
Helicopter pulling a diver out of the water
Coach hugging a tennis player
Dog chasing a mailman
Pirate digging up a chest of gold
Robot crushing a computer
Train crushing a bus
Bird pulling a worm out of the ground

Items with high event codability, low agent codability

Dog licking a puppy's face
Cop stopping a car
Dog leading a blind person
Bishop crowning a king
Mayor revealing a statue
Man shooting a woman
Painter painting a wall
Photographer filming a model
Technician fixing a tv
Football player scoring a goal
Journalist interviewing an actor
Masseuse massaging a man
Professor congratulating a student
Firetruck hitting a woman

Items with low event codability, low agent codability

Truck towing a car
Waiter kicking out a man
Astronaut catching an alien
Cooks preparing a cake
Snow covering a house
Bodyguard protecting the politician
Hunter roasting a pig
Cop arresting a man
Bee stinging a man
Woman dressing a boy
Sprinkler splashing an old woman
Maid throwing out the garbage
Conductor directing an orchestra
Security searching bags
Army attacking a castle
Bulldozer tearing down a building

Appendix B

Target pictures showing two-participant events in Experiment 2 grouped by Event and Agent codability category (one item yielded no scorable responses and is not included).

Items with high event codability, high agent codability

Girl pushing grandpa on a sled
Woman dressing a boy
Horse kicking a cow
Superhero rescuing a diver
Girl taking a cookie
Doctor examining a patient
Cop arresting a girl
Conductor conducting an orchestra
Ambulance ramming a car
Cowboy catching a sheriff
Ballerina tickling a boy

Items with low event codability, high agent codability

Cook preparing a cake
Frog eating a fly
Crab cutting a boy
Mayor unveiling a statue
Dog chasing a mailman
Bird pulling a worm out
Stork bringing a baby
Cat catching a mouse
Man shooting a woman
Pirate burying a treasure
Santa dragging a tree
Shark attacking a man

Items with high event codability, low agent codability

Bee stinging a bear
Professor congratulating a student
Technician fixing a car
Bulldozer destroying a building
Bishop crowning a king
Woman interviewing an astronaut
Painter painting a wall
Diver pushing a paparazzi
Dog licking a puppy
Woman massaging a man
Assistant painting a model
Wind blowing papers away

Items with low event codability, low agent codability

Security searching a bag
Scout roasting a pig
Maid throwing an umbrella out
Truck towing a car
Windmill hitting a farmer
Bodyguard protecting the queen
Army attacking a castle
Thief stealing a painting
Lightning striking a church
Bomb hitting a ship
Punk crushing a computer
Firetruck hitting a trafficlight

References

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

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

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

Bock, J. K. (1986). Syntactic persistence in language production. *Cognitive Psychology*, *18*, 355-387.

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

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*, 47–67.

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

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition, 98*, 105–135.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715-736.

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,* 57-84.

Gleitman, L., 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*, 544-569.

Griffin, Z. M. (2004). Why look? Reasons for speech-related eye movements. In F. Ferreira & M. Henderson (Eds.), *The integration of language, vision, and action: Eye movements and the visual world* (pp. 191-211). New York: Psychology Press.

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

Hartsuiker, R. J., & Kolk, H. J. (1998). Syntactic persistence in Dutch. *Language and Speech, 41*, 143-184.

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

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition, 35*, 925–937.

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., & Meyer, A. S. (2011, September). *Effects of lexical and structural priming on sentence formulation*. Paper presented at the 17th Annual Conference on Architectures and Mechanisms for Language Processing [AMLaP], Paris, France.

Kuchinsky, S. E., & Bock, K. (2010, March). *From seeing to saying: Perceiving, planning, producing.* Paper presented at the 23rd meeting of the CUNY Human Sentence Processing Conference, New York, NY.

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*, 748-756.

Levelt, W. J. M. (1989). Speaking*: From intention to articulation.* Cambridge, MA: MIT Press.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology, 25*, 188-230.

Myachykov, A., Tomlin, R. S., & Posner, M. I. (2005). Attention and empirical studies of grammar, *Linguistic Review, 22*, 347-364.

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

Tomlin, R. (1997). Mapping conceptual representations into linguistic representations: the role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162–189). Cambridge: Cambridge University Press.

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*, 423–440.

3

THE PARADOX OF SYNTACTIC CHOICE

Why syntactic flexibility leads to production difficulty

Abstract

One of the factors driving syntactic choices is verb bias. While some verbs are typically used with one structure (e.g., *voorschotelen* 'dish out′ with the double object dative in Dutch), other verbs have a weaker bias towards one syntactic frame (e.g., *serveren* 'serve'), allowing for syntactic flexibility during production. In this study we investigate whether syntactic flexibility leads to competition between structural alternatives, delaying sentence production. If competition between structural frames exists, sentence production may benefit from a mechanism that helps resolve competition between two syntactic frames by suppressing one frame to enable fast selection of the other frame. We hypothesized that inhibitory control can mediate this selection process, facilitating structure selection in the weak verb bias condition. Testing these predictions, we compared the speed of producing Dutch datives featuring verbs with a strong versus weak bias towards the double object or prepositional object dative structure. In line with the competition view, participants were quicker to initiate utterances featuring strong bias verbs than utterances featuring weak bias verbs. Additionally, participants with good selective inhibition (as assessed with the flanker task) showed a smaller verb bias effect than participants with poor selective inhibition. In contrast to earlier studies of syntactic flexibility with dative verbs, the results provide an indication that equipotent syntactic alternatives can compete during sentence production and suggest that inhibitory control can play a role in resolving this competition.

3.1 Introduction

When speaking, we are constantly faced with choices, ranging from which words to use to describe a scene to choosing which language to use. Likewise, assembling sentence structures often involves multiple syntactic options. Structure assembly requires, according to some models of language production, the incremental mapping of a conceptual representation onto an abstract structural plan encoding the hierarchical and positional relationships among words (Chang, 2002; Chang, Dell, & Bock, 2006; De Smedt, 1990; De Smedt, 1994; Kempen & Hoenkamp, 1987). This mapping has to conform to restrictions imposed by the grammar of the language, but leaves certain degrees of freedom for syntactic choices. For instance, when describing the event of *serving* in English we can choose between the prepositional object (PO) dative *The waiter serves the meal to the customer* and the double object (DO) dative *The waiter serves the customer the meal*. The aim of this study is to investigate whether these syntactic options can compete for selection and if so, how competition is resolved by the speaker.

Syntactic choices can be driven by several factors. Among the most studied are effects of lexical and conceptual accessibility: research has shown that linguistic entities with high accessibility (e.g., due to concreteness, animacy or recency factors) tend to be incorporated into the speech plan first, sometimes driving the choice for a syntactic alternative (Altmann & Kemper, 2006; Bock, 1982; Bock & Warren, 1985; Bresnan, Cueni, Nikitina, & Baayen, 2007; Christianson & Ferreira, 2005; Ferreira, 1994; Gleitman, January, Nappa, & Trueswell, 2007; Gropen, Pinker, Hollander, Goldberg, & Wilson, 1989; Kelly, Bock, & Keil, 1986; Kempen & Harbusch, 2004; Prat-Sala, 2000; Tomlin, 1997). Priming studies, on the other hand, suggest that the activation of a syntactic framework can bias the choice for a structural alternative (Bock, 1986; Bock & Griffin, 2000; Bock, Loebell, & Morey, 1992; Hartsuiker, Bernolet, Schoonbaert, Speybroeck, & Vanderelst, 2008). Activation levels of structural alternatives do not only depend on recent use of structures, but also on existing frequencies reflecting speakers' life-long exposure to written and spoken language. Frequently used structures, such as actives, have higher base level activation than infrequently used (i.e., dispreferred) structures (Bock, 1982). In line with implicit learning accounts of priming, actives are less susceptible to priming effects than their dispreferred structural alternatives, passives (Chang, Janciauskas, & Fitz, 2012; Hartsuiker & Westenberg, 2000; Segaert, Menenti, Weber, & Hagoort, 2011).

In some models, the frequency of a structure is tied to the verb it occurs with. This frequency information is represented via weighted links between verb lemmas and structural representations (Chang, 2002; Chang, et al., 2012). For instance, a verb that is typically used with the DO dative (e.g., *opleveren* 'yield') will have a connection to the DO structure with a higher weight than to the PO structure. In contrast, verbs that are used equally often in a DO and a PO frame (e.g., *toegooien* 'throw at'), will have equipotent connections to both alternatives. In Dutch (as in English) there is considerable variation in the strength of associations between dative verbs and their preferred structural configurations (Colleman, 2009). The strength of these verb biases determines the likelihood that structural alternatives will be selected when the verb is used. Hence verbs without a preference for one sentence structure will support the selection of both alternatives to a similar degree, which leads to syntactic flexibility in the grammatical encoding system. Melinger and Dobel (2005) found that reading a ditransitive verb in isolation primed the use of its associated structural frames during subsequent picture naming, suggesting that structural frames are automatically activated upon the retrieval of a verb.

In the current study we aimed to investigate whether syntactic flexibility, caused by the absence of verb bias, leads to competition between syntactic frames. Ferreira (1996) addressed the same question by investigating the production of structures with alternating verbs (i.e., verbs with two subcategorization frames: DO and PO dative) versus non-alternating verbs (i.e., verbs with only one subcategorization frame: PO). The task used in this study required participants to create a sentence from a preamble (e.g., *I gave*) and two or more sentence fragments (e.g., *children/toys*). As well as varying the flexibility of the verb (alternator vs. non-alternator) in the preamble, syntactic flexibility of alternator verbs was manipulated by prompting participants to use a preposition (Experiment 1, e.g., *to*) or a pronoun (Experiment 2, e.g., *it*) in half of their utterances. This constrained the order of the sentence to a PO dative. Thus, only in the alternator/unconstrained condition (i.e., *I gave... children/toys*), were participants free to make a syntactic choice. Ferreira (1996) found shorter sentence onset latencies and fewer errors when participants produced sentences in this condition than in the non-alternator and constrained conditions. This result was taken as evidence against a competition view and in favor of an *incremental* view of grammatical encoding. The latter posits that sentences are planned in a piecemeal fashion without competition between structural alternatives. Instead, syntactic flexibility should facilitate sentence production by offering more possibilities for filling slots. According to this view, the accessibility of the individual words drives structure selection and predicts the fluency of production at any given point in the sentence.

Here, we do not contrast competition and incremental views, as we do not consider these to be mutually exclusive. Sentence production can proceed in a piecemeal, incremental manner but competition could still arise at different points in this process, e.g., between the direct and the indirect object in dative structures. Conversely, if sentence planning involves elaborate pre-planning of a structural frame before sentence onset, structures would not necessarily have to compete for selection but could be engaged in an independent 'race' with the structure reaching threshold activation first being selected (cf. Van Gompel, Pickering, Pearson, & Liversedge, 2005). Here we follow the theoretical distinction of competitive versus non-competitive models that has been proposed in the comprehension literature.

Whereas few studies have investigated syntactic choices in language production, comprehension studies investigating parsing decisions under syntactic ambiguity are numerous. Theories of syntactic ambiguity resolution roughly fall into two categories. The first class of theories (*constraint-based*), assumes that alternative interpretations are activated in parallel and compete with each other (although there are noncompetitive parallel models as well, see Mohamed & Clifton (2011) for an overview). According to constraint-based theories, there is no restriction to the information sources that can provide support for one analysis over the other. Competition between alternative interpretations is assumed to be particularly high and to slow down parsing when a) constraints support both alternatives equally well, b) constraints favor one analysis when the ambiguity arises, but disambiguating information favors the alternative (e.g., Green & Mitchell, 2006; Kempen & Vosse, 1989; McRae, Spivey-Knowlton, & Tanenhaus, 1998; Vosse & Kempen, 2000). Some theories have proposed that an executive control mechanism mediates the selection of a parsing alternative by inhibiting the to-be-suppressed alternative (e.g., Novick, Trueswell, & Thompson-Schill, 2010)

The second class of theories (*two-stage theories)* postulates that sentence processing is serial and non-competitive (e.g., Van Gompel, et al., 2005; Van Gompel, Pickering, & Traxler, 2001). The *unrestricted race model* is an example of a two-stage theory. This model assumes, unlike other theories in this class, that there is no restriction on the information that can provide support for one analysis over the other. It proposes that alternative analyses are engaged in a race until the winner of the race is adopted as the initial analysis (*first stage*). If this analysis is inconsistent with later information, reanalysis has to take place (*second stage*). In contrast to the predictions of a competition model, the existence of multiple alternative analyses does not slow down parsing, but can speed it up. The "winner of the race" is the fastest analysis to exceed some threshold. Hence, its completion time should, in most cases, be shorter than the average completion time of any particular alternative. Therefore, the existence of multiple alternative analyses increases the chances of a fast winner. Additionally, the construction of different analyses is assumed to draw upon independent processing resources and therefore does not place an additional burden on the parsing system.

Studies have found mixed evidence for competitive and non-competitive theories (see Clifton & Staub (2008) for a review). A series of eye-tracking experiments showed that ambiguous sentences (where constraints equally supported two alternative analyses) were read faster than sentences that were disambiguated towards either the dominant or the non-dominant interpretation, which was taken as evidence against the existence of a competition mechanism (Van Gompel, et al., 2005; Van Gompel, et al., 2001). However, model simulations suggest that the eye-tracking data from these experiments could also be interpreted in a model that incorporates competition as a mechanism for ambiguity resolution (Green & Mitchell, 2006; Vosse & Kempen, 2009).

In language production, the study described above (Ferreira, 1996) provided evidence against a competition-based model by showing that participants were faster to produce sentences when multiple syntactic alternatives were available. However, the study has some limitations, which might complicate the interpretation of its results. Firstly, the non-alternating verbs used in the study were less frequent and longer than the alternating verbs (see Myachykov, Scheepers, Garrod, Thompson, & Federova (2013) for an analysis), which could have caused longer onset latencies for sentences with non-alternating verbs. Secondly, in the non-flexible conditions participants were always forced to produce the PO dative. In English, PO datives occur less frequently than DO datives (e.g., Kaschak, 2007). Hence, constraining participants to the use of one structural option (by requiring the use of the preposition *to* or the pronoun *it* in the constraining conditions - see above) could have led to more errors and slower production because participants were constrained to using the less preferred structure, the PO dative.

On an alternative account, competition between syntactic frames is suggested by a recent study comparing sentence production in English and in Russian (Myachykov, et al., 2013). This study examined whether the number of available syntactic options used in a picture naming task could predict sentence-initial processing load in a second picture naming experiment. In Experiment 1, speakers of English and Russian were encouraged to produce as many structural alternatives for a set of depicted events as possible within 15 seconds per picture. Since Russian grammar offers more options for building transitive constructions than English, Russian participants were hypothesized to use a greater variety of syntactic options than English participants. In Experiment 2, a new group of English and Russian participants described the same set of pictures while their eye-gaze was measured. The goal of the study was to link syntactic flexibility in the first experiment to sentence initial processing load in the second experiment, indexed by sentence onset latency and eye-voice span (i.e., the temporal lag between having the last fixation to a referent and producing that referent's name).

As expected, Russian speakers made more active use of the range of syntactic alternatives available to them than English speakers. In Experiment 2 they were also slower to initiate their sentences and showed longer eye-voice spans for sentence-initial subject nouns than for object nouns, whereas English speakers showed longer eye-voice spans for object nouns than subject nouns. More importantly, in both groups syntactic flexibility in Experiment 1 predicted sentence initial latency effects in Experiment 2: more syntactic alternatives led to longer onset latencies, after accounting for the effect of language (i.e., English vs. Russian). Hence, in line with a competition-based view, the availability of more structural alternatives led to an increase of associated sentence-initial processing load.

The present study contributes to the discussion on syntactic competition by comparing the production of Dutch dative sentences with varying degrees of syntactic flexibility. In an experiment using rapid serial visual presentation (RSVP; Lombardi & Potter, 1992; Potter & Lombardi, 1990, 1998) participants produced dative sentences featuring verbs with a bias towards the PO or DO dative. Syntactic flexibility was manipulated by using verbs varying in the strength of their bias for one structural alternative. Following a competition-based view, we hypothesize that speakers should be quicker to initiate a sentence when syntactic flexibility is low, i.e., when the bias towards one of the frames is strong. When a verb does not exhibit a significant bias towards either syntactic alternative, structures may compete until one candidate is selected, delaying production. We also propose that an inhibitory control mechanism might help speakers to overcome competition between syntactic frames more rapidly.

The specific inhibitory control mechanism that may play a role in the selection of a sentence structure is selective inhibition, or interference control. Selective inhibition refers to the ability to suppress interference specifically from responses that are strong competitors to the target response (Nigg, 2000; Ridderinkhof, 2002; Spaulding, 2010), e.g., ignoring a color word in the Stroop paradigm. This form of inhibition can be assessed with the Eriksen flanker task (Eriksen & Eriksen, 1974) where participants respond to a target stimulus as quickly as possible while ignoring distracting flankers. In language production, selective inhibition has been shown to play a role in the suppression of competing alternative representations, such as a semantic competitor in a picture naming task (Shao, Meyer, & Roelofs, 2013). Similarly, selective inhibition may be involved in the suppression of competing syntactic frames. Therefore, we hypothesize that selective inhibition becomes more important as competition between structural frames increases. Hence, when producing a sentence with a verb that only has a weak preference for one structural alternative, selective inhibition could facilitate frame selection by suppressing activation from the alternative structure. Conversely, in sentences featuring a verb that is strongly biased towards one structure, selective inhibition might play a less prominent role: here, the difference in activation levels between the structures preferred by the verb and the alternative structure naturally favors selection of the preferred structural frame. Statistically speaking, we thus expect to find an interaction between verb bias and selective inhibition (as indexed with flanker task performance) on sentence onset latencies.

To evaluate whether structure selection is specifically facilitated by this selective inhibitory control mechanism, and not by a more general top-down control system, we included a measure for nonselective inhibition in our study. Nonselective inhibition is associated with the top-down suppression of any (non-specific) prepotent response, e.g., stopping a 'hit-response' when the ball goes out of bounds in tennis (Miyake et al., 2000; Nigg, 2000). Although some researchers have argued that selective and non-selective inhibition represent ‘two sides of the same coin’ (Mostofsky & Simmonds, 2008), they have different developmental trajectories (Bjorklund & Harnishfeger, 1990; Jones, Rothbart, & Posner, 2003), different neural correlates (Forstmann et al., 2008), and show separable effects in relation to picture naming (Shao, Meyer, & Roelofs, 2013). Nonselective inhibition is normally assessed with the stop-signal task (Verbruggen & Logan, 2008) where participants report the identity of a stimulus presented to them on a screen, but have to withhold their response when they hear an auditory signal. In language production, nonselective inhibition was found to be related to naming speed in two picture naming studies, such that good nonselective inhibitors showed shorter onset latencies than poor nonselective inhibitors (Shao, et al., 2013; Shao, Roelofs, & Meyer, 2012). More importantly, in one of the studies a picture-word interference paradigm was used and nonselective inhibition contributed to faster naming speed for pictures with both semantically related and unrelated distractors (Shao, et al., 2013). Hence, regardless of the strength of the competitor, nonselective inhibition speeded up production. Similarly, nonselective inhibition could be related to production speed in sentences, regardless of the strength of competition from the structural alternative. In other words, good nonselective inhibition may lead to faster sentence onset latencies regardless of the bias of the verb. Most importantly, we hypothesize that there will be a discrepancy between the effects of selective and nonselective inhibition on the production of datives with varying syntactic flexibility. Whereas we expect to find an interaction between verb bias and selective inhibition in a model predicting onset latencies, either a main effect or no effect on speech initiation is expected for nonselective inhibition.

In sum, we hypothesize that a) sentences featuring strong bias verbs are initiated faster than sentences with weak bias verbs; b) selective inhibition mediates the selection of a structure when two equipotent alternatives are available, facilitating structure selection in sentences with weak bias verbs only. Hence we expect to find an interaction between verb bias and selective inhibition in a model predicting onset latencies. Finally, c) nonselective inhibition does not enter into interaction with verb bias, but may increase overall production speed (i.e., resulting in a main effect of nonselective inhibition on onset latencies)

3.2 Method

Participants

36 adult native speakers (ages 18 -30 years) of Dutch gave informed consent and participated in the experiment for payment. All participants had normal or corrected-to-normal vision. Consent for conducting the study had been obtained from the Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen.

Materials

Dative verbs with weak and strong biases towards the prepositional dative and double object dative structures were selected based on a corpus analysis of the Dutch dative alternation (Colleman, 2009). In this corpus analysis, collostructional strength was identified for 252 alternating dative verbs. Collostructional strength is the degree of association between one lexical item (in this case a verb) and two or more functionally similar abstract constructions. This degree of association is based on the frequencies of the verb occurring in each of these constructions and on the overall frequencies of the construction in the corpus, and is computed using the Fisher exact test. An index of distinctive collostructional strength was calculated as -log(Fisher exact, 10). The higher the index, the stronger the verb's preference for the construction. For example, *give* has a strong preference for the DO structure and *show* has a weak preference, with collostructional strengths of 5.56 and 0.27 respectively. Although *give* is an example of a high frequency verb, its degree of strength is not correlated with lexical frequency. The study revealed a wide range of collostructional strength for verbs preferring DO and PO constructions, with PO dative preferring verbs displaying a wider range (0.17 – 69.07) than DO dative preferring verbs (0.16 – 40.8).

A total of 28 verb pairs with low (*M* =1.04, *SD* = 0.95) and high (*M* = 13.85, *SD* = 13.80) collostructional strength were selected from Colleman (2009), resulting in a weak bias and strong bias condition (e.g., *voorlezen* 'read out' and *leren* 'teach'*)*. For each verb pair, one sentence was constructed which could accept both verbs (e.g., *De docent leest het verhaal voor aan de studenten* 'The teacher reads out the story to the students' vs. *De docent leert het verhaal aan de studenten* 'The teacher teaches the story to the students') in both a DO and PO sentence frame.

Sentences consisted of one main clause and were written in the present tense. They included only medium to high frequency nouns (loglinear form frequency range: 0.48 - 2.94*)*, which increases the likelihood of participants correctly recalling the nouns and rules out the possibility that structure selection would be driven by differences in accessibility of the direct and indirect object. Verb bias conditions and dative structures were matched for verb log lemma frequency, syllable count and separability (CELEX Lexical Database, Baayen, Piepenbrock, & Gulikers, 1995). The latter refers to the possibility of separating the verb core and its particle, as for example in *terug-betalen* 'pay back'. Dutch and German are notable for having many separable verbs. Importantly, Dutch has two placement options for the particle in PO datives: before or after the canonical position of the indirect object. In our experimental stimuli the particle always preceded the indirect object, e.g., *het kind betaalt het geld terug aan de moeder* 'the child pays the money back to the mother'.

Verb lemma frequency was uncorrelated to verb collostructional strength (*r* = .14). Table 1 shows the descriptive statistics for the verbs organized by structural preference

Table 1. Descriptive statistics for DO dative and PO preferring verbs used in the experiment.

		Collostructional strength	Lemma frequency	Plausibility norming
Verb preference				
DO dative	average	5.07	1.40	5.54
	Stdev	6.93	.77	1.03
	Range	.20 - 32.83	.00 - 3.11	2.25 - 6.88
PO dative	average	9.62	1.24	5.34
	Stdev	14.18	.59	.90
	Range	.33 – 66.26	.00 – 2.26	3.00 – 7.00

The experiment used a 2 (structure) x 2 (verb bias) within-participant and within-item factorial design. Thus each item (sentence frame) was presented with a verb biased towards the preferred structure half of the time and with a verb biased towards the alternative structure the other half, leading to two congruent and two incongruent verb-structure pairings for each verb pair. Table 2 shows an example of one sentence frame in each of the four conditions.

Table 2. Example of a full item set for the verb pair *voorschotelen/serveren* 'dish out/serve'.

Sentence	Verb bias	Verb-Structure Congruence
de ober schotelt de klant de maaltijd voor *the waiter dishes out the customer the meal*	strong (DO biased)	congruent
de ober serveert de klant de maaltijd *the waiter serves the customer the meal*	weak (DO biased)	congruent
de ober schotelt de maaltijd voor aan de klant *the waiter dishes out the meal to the customer*	strong (DO biased)	incongruent
de ober serveert de maaltijd aan de klant *the waiter serves the meal to the customer*	weak (DO biased)	incongruent

Four lists of stimuli were created to counterbalance structure and verb bias, so that each item appeared in a different condition in each list. Within lists, there were seven different items in each of the four conditions. In addition to the 28 target items, there were 90 filler sentences. Filler sentences mainly included intransitive verbs (e.g., *The legendary treasure really existed*) and consisted of five to nine words. Ten filler sentences were used at the beginning of the experiment as practice items.

Norming

To evaluate whether the target sentence frames featuring different verbs (i.e., strong vs. weak bias verbs) were equally plausible, 40 participants not taking part in the main experiment were asked to rate the semantic/pragmatic plausibility of the target sentences on a scale from 1 (implausible) to 7 (very plausible). The item norming consisted of 45 verb-pairs, which were used to create 180 sentences: each sentence frame appeared with both verbs (strong and weak bias), and as PO and DO dative. In the norming, sentences were randomly assigned to one of four item lists, so that each sentence frame appeared in all four conditions across lists and each verb appeared exactly once per list. Hence, each item was rated by ten participants.

Based on these plausibility ratings, 28 verb-pairs were selected that exhibited an average plausibility of above 3.5 and that did not show large differences between strong and weak verb bias conditions. In incongruent verb-structure pairings we expected lower average plausibility, especially for strong bias verbs. Table 3 shows a summary of the norms for the selected sentence frames per condition.

Table 3. Mean plausibility ratings and standard deviations (in parentheses) for strong and weak bias verbs by sentence structure and verb-structure congruence.

	DO dative		PO dative	
Verb bias	Congruent	Incongruent	Congruent	incongruent
Strong bias	5.60 (1.70)	5.06 (1.94)	5.67 (1.60)	5.18 (1.94)
Weak bias	5.96 (1.45)	5.35 (1.80)	5.51 (1.87)	5.67 (1.85)

As shown in Table 3, all sentences received high plausibility ratings. In line with our predictions, paired-samples *t*-tests revealed a main effect of congruence, $t(1,27) = 2.05$,with sentences in congruent verb-structure configurations being judged as more plausible than sentences in incongruent verb-structure configurations. The interaction between verb bias and congruence was marginally significant, $F(1,27) = 3.27$, $p = .08$, suggesting that verb-structure congruence had a different effect on plausibility in the two verb bias conditions. Pairwise comparisons showed that strong bias verbs in the incongruent condition were rated as less plausible than the other sentence types ($t[1,27] = -1.88$, $p = .07$). To control for possible effects of plausibility, we planned to carry out a complementary analysis including the plausibility ratings as a continuous factor (see below ...).

Procedure

Participants were randomly assigned to one of the four item lists. Instructions for the experiment appeared on the screen and participants received ten practice sentences before the experiment started. Sentences were presented in rapid serial visual presentation (RSVP). The sequence of events for each trial is illustrated in Table 4 (adapted from Konopka & Bock, 2009). The experiment was programmed using Presentation® software (Version 16.3, www.neurobs.com).

Table 4.Sequence of events during each experimental trial.

Duration	Event
200 ms	+
100 ms	De
100 ms	ober
100 ms	serveert
100 ms	de
100 ms	maaltijd
100 ms	aan
100 ms	de
100 ms	klant.
100 ms	##########
533 ms	4 5 2 9 1
100 ms	[screen blanked]
500 ms	twee
10 ms	[screen blanked]
5000 ms (max)	Nee Ja
500 ms	☺ or ☹
	De ober….

After presentation of a 200 ms fixation cross, participants read a sentence, which was presented one word at a time. Word presentation time was 100 ms, similar to earlier English RSVP studies. Although average word length in Dutch is higher than in English (Hagoort & Brown, 2000), a pilot study revealed that participants could process words with a presentation time of 100ms. Participants were instructed to read the sentence silently and to remember the content (i.e., the message) as they would have to reproduce it later.

They then performed a distractor task, in which they first saw a display of five digits and then had to judge whether a digit (written out in letters: e.g., *twee* 'two') had been part of this array of five digits. They responded by pressing the left (*yes*) or right (*no*) mouse button and had a maximum of five seconds to do so. They were given immediate feedback in the form of a happy face for a correct answer and a sad face for an incorrect answer. On 50% of the trials the correct answer to the distractor task was '*yes*' and on the remaining half the answer was '*no*'. On critical trials the correct answer to the distractor task was always '*yes'*, while on filler trials this could vary.

After the distractor task, participants were prompted to repeat the sentence they had read at the beginning of the trial. They were instructed to produce the sentence as quickly as they could without making any mistakes or producing disfluencies. The first noun phrase of that sentence appeared on the screen (e.g., *The driver*) and participants were asked to complete the sentence with the verb, direct object and indirect object (i.e., without reading the subject noun phrase of the sentence out loud). They then pressed a button to proceed to the next trial. Responses were recorded and the speech output was transcribed by the experimenter offline. Later, speech onsets were coded manually in Praat (Boersma & Weenink).

Participants were debriefed at the end of the experimental session about the goal of the main experiment. In addition, they were asked to describe which strategy they used to remember the sentences and to what extent they remembered the exact form of the sentences. In line with findings from sentence recall studies (e.g., Bock & Brewer, 1974; Potter & Lombardi, 1990, 1998), participants reported they often tried to reconstruct new surface structures from memory of underlying sentence meanings.

Assessment of selective and nonselective inhibition

After the main experiment, participants carried out a flanker task (measuring selective inhibition) and a stop-signal task (measuring nonselective inhibition). In the flanker task (Eriksen & Eriksen, 1974), participants had to react to a target stimulus (< or >) as fast as possible, while ignoring irrelevant flanker stimuli. In a row of five symbols, participants had to categorize the middle symbol as a leftward or rightward pointing arrow by pressing the left or right button on a button box. Flankers (on either side of the target) were congruent (<<<<<), incongruent (>><>>), or neutral (- - < - -) with respect to the target (here: <). Each trial started with a beep and a fixation cross, which was presented for 250 ms. After the fixation cross, the symbol string was presented for 1500 ms. The interstimulus interval was 1000 ms. Six practice trials preceded the total of 72 experimental trials.

Mean accuracy of the responses (pooled over participants) was 95% correct (*SD* = 3). Mean response times (for correct responses only) in the three conditions were 447 ms in the congruent condition (*SD* = 134), 545 ms in the incongruent condition (*SD* = 154), and 433 ms (*SD* = 121) in the neutral condition. Selective inhibition ability was measured by regressing reaction times in the neutral condition from reaction times in the incongruent condition. In contrast to subtraction scores, this yields residuals which are uncorrelated with the general processing speed in the neutral condition (cf. DeGutis, Wilmer, Mercado, & Cohan, 2012). These residuals reflected flanker cost after statistically removing neutral condition variance, with higher scores indicating poorer selective inhibition.

To assess nonselective inhibition, we used a computerized version of the stop-signal task designed by Verbruggen, Logan, and Stevens (2008). In the stop-signal task, participants have to identify a visual stimulus but withhold their response when an auditory signal is presented. Here, the primary task was a shape judgment task that required participants to discriminate between a square (1.5 x 1.5 cm) and a circle (1.5 cm in diameter). On go-trials (75%), participants had to respond to the stimulus as fast and accurately as possible by pressing the "*z"* key for *square* and the "*/"*key for *circle* (on an English/Dutch keyboard) Each trial started with a fixation cross that remained on the screen for 250 ms. After the fixation cross, the shape stimulus was presented until the participant responded, with a maximum duration of 1250 ms. On stop-signal trials (25% of all trials), an auditory stop signal (a beep at 750Hz, presented for 75 ms) was presented after the shape stimulus with a variable delay (SSD). Participants were instructed to withhold their response on these trials. SSD was initially set at 250 ms and was continuously adjusted: when inhibition was successful SSD increased by 50 ms and when it was unsuccessful it decreased by 50 ms. The interstimulus interval was 2000 ms. The task consisted of two phases: a practice phase of 32 trials and an experimental phase of 64 trials in each of three blocks. Participants received feedback about their performance on the task (e.g., number of incorrect responses on no-signal trials) after each block.

Performance was measured with the stop-signal RT (SSRT). SSRT was estimated by subtracting the mean SSD across all trials from the mean RT on no-signal trials. A short SSRT indicates that participants can withhold their response relatively late during response preparation and is thus indicative of good nonselective inhibition.

Scoring and Analysis

Sentences were scored as having either double object (DO) dative or prepositional object (PO) dative syntax (e.g., *proposes the professor the plan*; *proposes the plan to the professor*). Sentences with intransitive syntax or other constructions were excluded, as were sentences with repeated sentential subjects (e.g., *The student proposes...)*, omitted direct or indirect objects (e.g., *pays back the mother*), verb substitutions (e.g., *asks* instead of *proposes*), or noun substitutions (except for minor substitutions, e.g., *grandma* instead of *grandpa*, or the use of plural instead of singular nouns). Additionally, sentences following a wrong answer to the distractor task were also excluded. This was done to control for possible effects of response feedback in the distractor task on subsequent response latencies. One item (verb pair: *distribute/feed*) was excluded from analysis because of a programming error.

Finally, we eliminated responses with onsets latencies longer than 3000 ms and onsets more than 2.5 standard deviations away from the grand mean. The final dataset consisted of 720 responses (371 PO sentences, 349 DO sentences), equivalent to ten scorable responses per participant in each verb bias condition.

All data were analyzed using R (R Development Core Team, 2013) and the R packages *lme4* (Bates, Maechler, & Dai, 2009) and *languageR* (Baayen, 2008b). Analyses on structure choice and error rates were carried out with mixed logit models (coefficients are given in log-odds). Onset latencies (RT's) were analysed with linear mixed effects models (coefficients are given in milliseconds). Model factors included Verb bias as a categorical factor, Congruence, and Sentence structure, after they were centred. In all analyses, we used a backwards elimination procedure, starting from an initial model containing all experimental factors and their interactions and random by-subject and by-item intercepts (Baayen, 2008a; Baayen, Davidson, & Bates, 2008; Jaeger, 2008). Non-significant effects were removed, starting from the highest-order interactions and going back to a basic additive model with only main effects. Since MCMC sampling (Baayen, et al., 2008) is not implemented in *lme4* for models with random slopes, p values were obtained by treating the t-statistic as a z-statistic, with effects considered to be reliable at the $\alpha = 0.05$ level, $z = 1.96$. For the remaining fixed effects structure, random slopes were added if they improved model fit as indicated by likelihood ratio tests.

Complementary analyses were carried out a) using verb bias as a continuous factor (i.e., collostructional strength), and b) including the predictor Plausibility (according to the item norming) to assess the effects of the experimentally manipulated factors above and beyond the effect of the plausibility ratings. These analyses confirmed the pattern of results obtained from analyses using only experimentally manipulated factors and are therefore not reported. Scores obtained in the stop-signal task and flanker task did not correlate significantly ($r = -.19$), suggesting that they measure different types of inhibition (e.g., De Jong, Coles, & Logan, 1995).

3.3 Results

We report four sets of results. *First* we examine error rates across sentence types, verb-structure congruence and verb bias in the full dataset (n = 972). Responses were coded as an error when they contained word substitutions or other constructions than the ditransitive, followed an incorrect answer to the distractor task, and when onset latencies exceeded the outlier threshold. In a *second* analysis, we compare the proportions of DO and PO datives produced across conditions after excluding errors (n = 720). The *third* analysis tests whether verb bias, congruence, and sentence structure predicted verb onset latency on trials where the speakers produced the intended structure (i.e., the sentence structure that was presented to them, n = 697). *Finally*, we describe how interactions obtained with Flanker and Stop task scores clarify how speakers differed in their performance on the production task.

Error rates

Descriptively, speakers produced most errors by using a different sentence construction (7% of all trials), substituting the verb (5%) and by omitting a noun (4%, i.e., *laat het lied horen* instead of *laat het publiek het lied horen* 'let the song hear' instead of 'let the audience hear the song'). For the analysis of error rates, we collapsed across the different types of errors and compared errors across sentence structures, verb bias and congruence conditions.

Error rates showed an effect of congruence with sentences featuring congruent verb-structure pairings being produced with fewer errors (21%) than incongruent sentences (31%; β = .57, z = 3.62, in the additive model). Verb bias and Sentence structure did not predict error rates (z < 1.28).

Production of DO and PO datives

Overall, speakers produced slightly more PO datives (52%) than DO datives (48%). Sentence structure as presented to the participants was the most important predictor of the produced sentence type (binomial dependent variable, levels "DO" and "PO"), with speakers largely repeating the structure of the sentences they had just read (β = 9.17, z = 11.22, in the additive model). Only on a small minority of trials (3%), speakers flipped the sentence structure. In a separate model (with flipping as a binomial dependent variable, levels "yes" and "no"), flipping probability was primarily predicted by the congruence of the structural preference of the verb and the presented sentence structure (β = 2.62, z = 2.94): Speakers were more likely to flip sentence structure in incongruent (6% in all incongruent trials) than in congruent verb-structure pairings (0.5% in all congruent trials) and thus followed the preference of the verb in their structure choice. Flipping also occurred more often for DO (into a PO structure; 5%) than for PO structures (1%; β = -1.56, z = -2.25).

Verb onsets

Onset analyses were carried out on trials where speakers produced sentences with the intended structure (i.e., excluding 'flipped' structures, n = 697), making use of linear mixed effects models. Analyses of verb onset revealed a main effect of Verb bias (β = -38.78, *SE* =17.15, t = -2.26 in the additive model). In line with a competition-based account, speakers were faster to produce datives containing strong bias verbs (*M* = 814, *SD* =264) than datives containing weak bias verbs (*M* = 851, *SD* = 302). Hence, initiating a sentence with a verb that is used with the two structures equally often takes longer than initiating an utterance with a verb that has a strong bias towards one structure over the other. This suggests that structural alternatives compete during production, with stronger competition leading to longer reaction times. Congruence (of verb preference and sentence type) and Sentence structure did not reliably predict verb onsets (ts < .66).

Verb onsets and the role of inhibitory control

Adding Flanker score to a model predicting verb onsets, yielded only a marginally significant interaction between Verb bias, Congruence, and Flanker score, β = -1.09, *SE*=0.66, t =-1.66, p = .10 (with random by-item slope for Flanker score). To examine the effect of selective inhibition more closely, we ran a separate model for congruent and incongruent verb-structure pairings. A model predicting verb onsets for congruent trials (n = 381) revealed a main effect for Verb bias only (β = -45.04, *SE* = 22.52, t =-2.00), whereas for incongruent trials (n = 316) the effect of Verb bias was weakly modulated by Flanker score but the interaction did not reach significance (β = -0.77, *SE* = 0.51, t =-1.53). For illustrative purposes, we aggregate trials on the subject level and show a scatterplot displaying the relation between flanker cost and verb bias effect (weak bias RT – strong bias RT) by subjects in Figure 1.

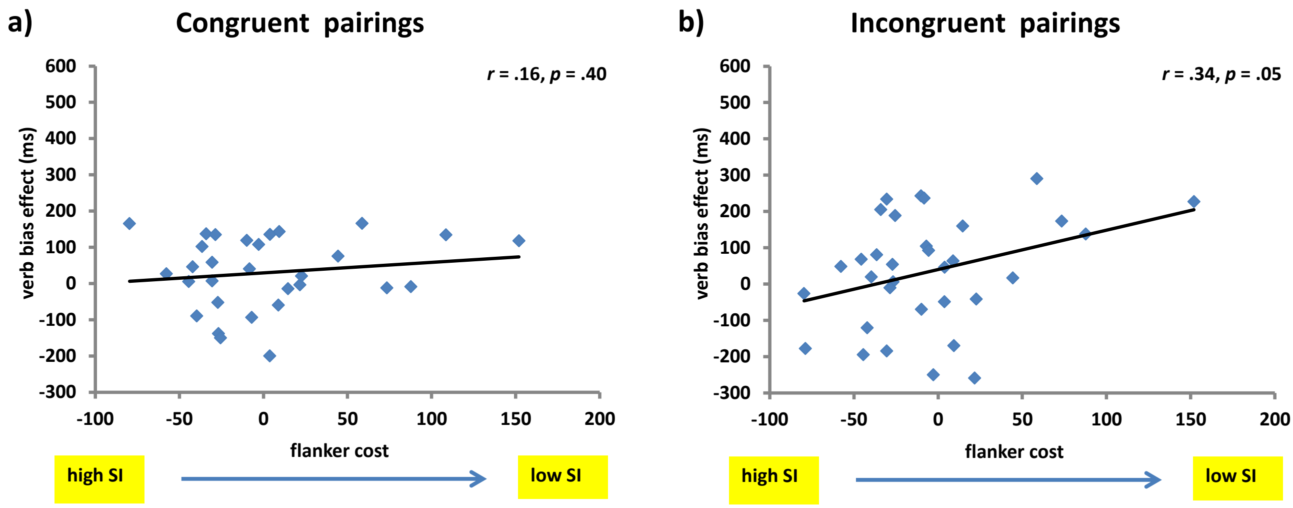

Figure 1.Scatterplots of the relationship between flanker cost (residuals) and the magnitude of the verb bias effect (weak bias RT - strong bias RT) for a) items with congruent verb-structure pairings *n* = 30, and b) items with incongruent verb-structure pairings, *n* = 33. For these plots and the associated correlation index, outliers were excluded by calculating Cook's distance and removing extreme values. high SI = high Selective Inhibitory control, low SI = low Selective Inhibitory control.

As shown in Figure 1, the weak interaction between Verb bias, Congruence and Flanker score was driven by the positive correlation between Flanker score and the Verb bias effect in items with incongruent verb-structure pairings on the one hand and the weakly positive relation between Flanker score and the Verb bias effect for congruent verb-structure items on the other hand. For incongruent trials, the RT difference between sentences with weak versus strong bias verb increased as speakers' selective inhibitory control decreased. In other words, inhibitory control might have helped speakers to rapidly overcome competition between structural frames in incongruent verb-structure items such that that they did not show a difference in onset latency when producing sentences with a strong and a weak verb bias.

Adding Stop score instead of Flanker score revealed a marginally significant interaction between Verb bias, Sentence structure and Stop score ($\beta = -1.63$, $SE = 0.90$, $t = -1.80$, $p = .07$). Separate analyses were performed per sentence structure (PO vs. DO dative). The analyses of PO datives showed a main effect of Verb bias and a marginally significant effect of Stop score ($\beta = -56.03$, $SE = 24.51$, $t = -2.29$, and $\beta = -1.51$, $SE = 0.79$, $t = -1.92$, $p = 0.06$ respectively), whereas the onset of DO datives was predicted by an interaction between Verb bias and Stop score ($\beta = 1.89$, $SE = 0.63$, $t = 3.02$). Figure 2 shows the relation between

Stop score and the Verb bias effect for DO and PO structures, while aggregating data on the subject level.

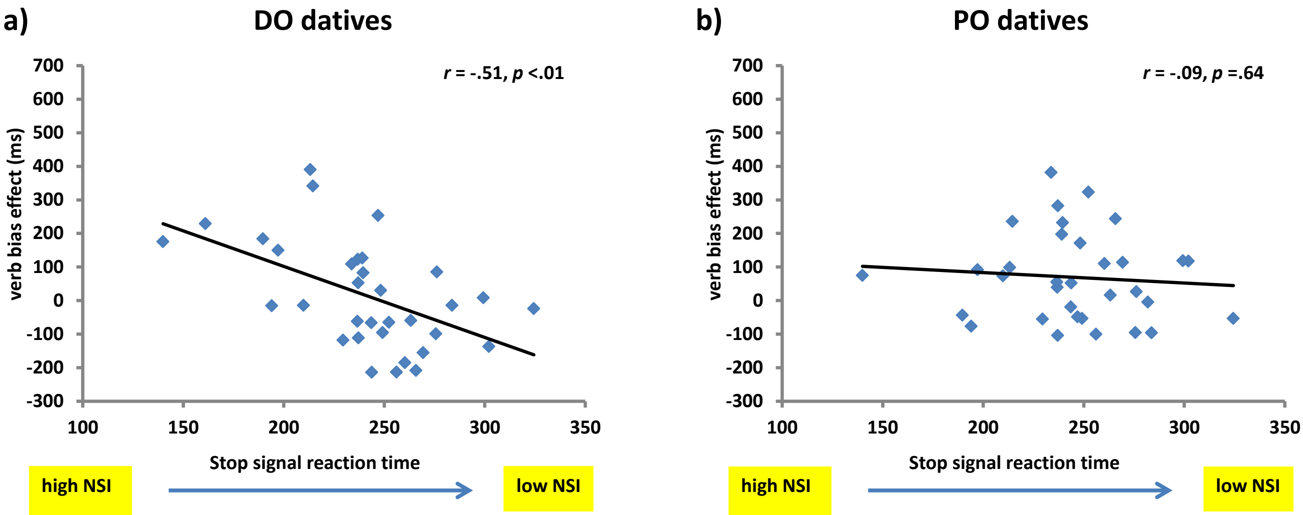

Figure 2. Scatterplots of the relationship between stop signal reaction time and the magnitude of the verb bias effect (weak bias RT - strong bias RT) for a) items with DO structure $n = 33$, and b) items with PO structure, $n = 33$. For these plots and the associated correlation index, outliers were excluded by calculating Cook's distance and removing extreme values. high NSI = high NonSelective Inhibitory Control, low NSI = low NonSelective Inhibitory control.

Figure 2 shows that the Verb bias effect in DO datives was negatively correlated with Stop score. Producing DO datives, good inhibitors were faster to initiate strong bias than weak bias verbs, while poor inhibitors were faster to initiate weak bias verbs. PO datives did not show a significant relation between the Verb bias effect and Stop score.

3.4 Discussion

The goal of this study was to investigate whether syntactic alternatives can compete for selection and if so, how competition is resolved by the speaker. We predicted that the production of datives with strong versus weak bias verbs would reveal syntactic competition by showing longer onset latencies for sentences where multiple, equipotent syntactic alternatives are available (i.e., for the weak verb bias condition). In an experiment using RSVP we found that participants were indeed slower to initiate sentences featuring weak bias verbs than sentences with strong bias verbs. This finding suggests that syntactic flexibility led to the simultaneous activation of both dative sentence frames.

But if this is true, the question arises as to why speakers only make use of this flexibility to produce the alternative sentence structure (i.e., 'flip') in 3% of the cases. The nature of the RSVP paradigm used in this study might explain why speakers did not frequently flip sentence structure. Potter and Lombardi (1990) were the first to use this paradigm to investigate the nature of verbatim short-term memory for sentences. They argued that short-term recall in RSVP relies on regeneration from a sentence's meaning instead of recall of its surface structure. According to their theory, recall is (near) verbatim because speakers are likely to reuse recently activated (unordered) lexical entries and syntactic structures (Lombardi & Potter, 1992; Potter & Lombardi, 1998). Likewise, the relative scarcity of structure flipping in the current study can be explained by a priming-account: an implicit memory trace of the recently activated structure promotes re-usage of that structure when regenerating a sentence from its meaning. However, even though speakers may be primed to re-use a presented sentence structure, competition between syntactic frames should still arise when a weak bias verb is retrieved during sentence regeneration. As already mentioned above, Melinger & Dobel (2005) found that presenting a non-alternating verb in isolation already primed the re-use of its subcategorization frame during picture description. Similarly, retrieval of a verb in isolation could be sufficient to activate the weighted links between verb lemmas and their subcategorization frames to the degree specified by the verb's bias.

A priming account also explains why speakers reproduced the syntactic surface form even in the incongruent verb-structure condition. In priming studies, the 'inverse preference effect' causes the activation of a less preferred structure to lead to greater internal weight changes and hence to more priming than a preferred syntactic structure (Ferreira & Bock, 2006). Structural preferences and verb bias interact in determining priming strength, such that priming a structure containing a verb with an opposite preference (e.g., a DO structure with a PO preferring verb) will lead to larger priming effects (Bernolet & Hartsuiker, 2010). Hence, when presenting participants with a structure featuring a verb that has an opposite structural preference, priming effects will be larger than for structures featuring a verb with congruent verb bias.

In addition to hypothesizing that the absence of a verb preference would cause competition between syntactic frames, we proposed that a selective inhibitory control mechanism would enable speakers to resolve this competition efficiently. Selective inhibitory control was assessed with the flanker task. Results indicated that selective inhibition weakly mediated structure selection under syntactic flexibility. In items with incongruent verb-structure pairings, poor selective inhibitors were faster to initiate sentences with strong bias verbs than sentences with weak bias verbs, whereas onset latencies of good selective inhibitors did not differ much between verb bias conditions. This finding suggests that good selective inhibitors are better (faster) at overcoming competition from equipotent syntactic alternatives when the verb does not bias towards one structure, especially for sentences with incongruent verb-structure pairings. Although this pattern of results is in line with the predictions from a competition model, the effects are too weak to draw strong conclusions from. A possible reason for the absence of a significant interaction between selective inhibition and verb bias is that sample size in the current experiment was relatively small. Yet, the observed trend showing involvement of selective inhibition in dative sentence production is suggestive of interference from competing representations during structure selection.

For comparison, we also examined the effect of a presumably different construct of inhibition; nonselective inhibition. As expected, a different pattern of results was found for nonselective inhibition than for selective inhibition. We hypothesized that good nonselective inhibition may facilitate production speed (or have no effect on production latencies), based on results from studies using pictures with single objects (with unrelated or semantically related competitors). Instead, in a full model including an interaction between Stop score, Verb bias and Structure, good nonselective inhibitors were slightly slower to initiate sentences than poor nonselective inhibitors ($\beta = -1.41$, $SE = 0.77$, $t = -1.83$, $p = .07$). A possible explanation for this effect is that poor nonselective inhibitors employed a different strategy than good nonselective inhibitors to produce sentences in the specific paradigm that we used.

Production speed, planning scope, and effects of syntactic flexibility

In a study examining grammatical planning scope using picture-word interference tasks, slow speakers (i.e., speakers with long speech onset latencies) were found to have a more extensive planning scope than fast speakers (Wagner, Jescheniak, & Schriefers, 2010). In the present study, the fast onsets of poor nonselective inhibitors may also be due to a limited planning scope. Myachykov et al. (2013) argue that a limited planning scope does not allow for weighing of global syntactic options and that therefore the manifestation of competition between syntactic frames may depend on task demands such as time pressure. Similarly, speakers' planning strategies should also influence the degree to which competition between frames arises. In incremental (piecemeal) planning, competition between syntactic frames under syntactic flexibility may not arise because lexical accessibility drives structure choice. In fact, incremental planning could lead to an inverse effect of verb bias: incremental planners may be faster to produce sentences featuring weak than strong bias verbs because in the latter case they are constrained to using one specific frame (and therefore order of lexical items), while under syntactic flexibility they have more possibilities to fill sentence slots (cf. Ferreira, 1996). This would explain why the fast, poor non-selective inhibitors showed the reverse verb bias effect: faster onsets for sentences with weak than strong bias verbs.

This hypothesis can be put to test by adding Production speed to a model predicting verb onsets from Verb bias, Congruence, Sentence structure and Stop score. To obtain a neutral index of production speed, we measured onset latencies on filler trials. For the scoring of the filler trials we used the same criteria as for target trials: responses were excluded that a) did not have the correct wording or structure, b) followed a wrong answer to the distractor task, or c) had onset latencies longer than 3000 ms and onsets more than 2.5 standard deviations away from the grand mean. Based on the remaining trials (83% of all trials, equivalent to 66 trials per participant on average) we computed an average production speed per participant. This continuous measure was included in a model predicting verb onsets from Verb bias, Congruence, Sentence structure and Stop score. According to our hypothesis, when speakers initiate sentences more slowly, verb bias effects should get stronger, such that sentences with strong bias verbs are initiated faster than sentences with weak bias verbs. In line with this prediction, a linear mixed effects model analysis yielded a significant interaction between Verb bias and average Production speed ($\beta = -0.33$, $SE = 0.11$, $t = -2.96$).

Figure 3 shows the direction of this interaction in a scatterplot: Slower speakers showed larger effects of verb bias.[1]

[1]Slow responders also showed larger verb bias effects in a model including Flanker score. For congruent verb-structure pairings, all slow responders showed larger verb bias effect. In contrast, for incongruent verb-structure pairings, only the poor selective inhibitors showed an increase of the verb bias effect in slow responders. This was especially the case in incongruent PO structures (resulting in a five-way interaction between Flanker score, Verb bias, Congruence, Structure, and Response speed, $\beta = -.03$, $SE = 0.01$, $t = -2.64$). This result is in line with word production studies showing involvement of selective inhibition in the slowest segment of responses (e.g., Shao et al., 2013).

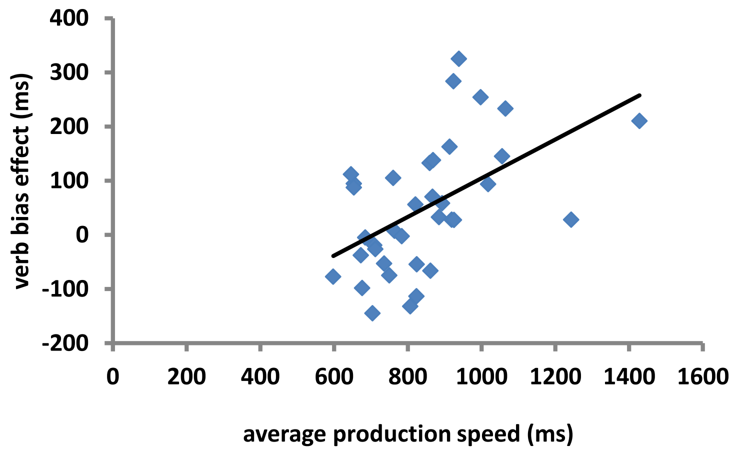

Figure 3. Scatterplot of the relationship between average production speed and the magnitude of the verb bias effect (weak bias RT - strong bias RT).

This result suggests that production speed and the associated grammatical planning scope determine to what degree competition effects are manifested. The reason why we do not find a three way interaction between Verb bias, Production speed and Stop score is that the interactions Production speed by Verb bias and Stop score by Verb bias may be mediated by the same underlying factor. This factor may be as simple as task motivation. There are studies suggesting that high task motivation can decrease stop signal reaction time (Pessoa, 2009; Slusarek, Velling, Bunk, & Eggers, 2001). Similarly, one can imagine that motivation for producing fluent sentences in the RSVP task leads speakers to plan their sentences more carefully. If motivation is low, speakers may choose to start their sentence as fast as possible, perhaps at the expense of their production accuracy. To summarize, highly motivated participants may have performed better on the stop-signal task, planned their sentences more carefully in the production task (leading to longer onset latencies) and therefore shown effects of competition between syntactic frames.

Sentence structure, planning scope, and effects of syntactic flexibility

Although the above account viably explains differences between good and poor non-selective inhibitors in their susceptibility to syntactic flexibility, it does not explain the discrepancy in results for DO vs. PO structures in these groups of speakers (see Figure 2; leading to the three-way interaction between Verb bias, Structure and Stop Score). One possibility is that the difference in Verb bias effects across sentence structures is also mediated by planning scope differences.

Previous research has shown that not only discourse factors (e.g., time pressure) and speaker characteristics (e.g., production speed), but also ease of structural processing, influences planning strategy (Konopka, 2012; Van de Velde, Meyer, & Konopka, 2014). Easy-to-assemble structures (e.g., because they are primed or preferred) have a larger grammatical planning scope than more difficult sentence structures. In Dutch, DO datives are more frequent than PO datives and may therefore be easier to produce (e.g., Colleman, 2009; Kaschak, 2007). As a consequence, planning scope for DO datives may be larger than for PO datives, especially for speakers flexible in adjusting their planning scope–slow speakers.

Another reason why syntactic planning of PO and DO datives may differ is that producing a PO structure involves more word ordering choices. Firstly, some theories do not regard PO structures as dative alternates but as monotransitives with an optional adjunct PP (cf. Brown, Savova, & Gibson, 2012; Dowty, 2003). Speakers may thus choose to leave out the indirect object and produce a monotransitive construction, with the direct object as the only obligatory argument (e.g., *submits the plan* instead of *submits the plan to the professor*). In DO datives, the indirect object can also be omitted but this requires an extra step, namely placing the direct object in immediate post-verbal position (e.g., *serves the meal* instead of *serves the customer the meal*). In the present study, the omission of the indirect object was one of the most frequently occurring errors (4% of all responses). Omissions were most frequent in sentences with PO preferring verbs presented in DO sentence frames.

Secondly, placement of the verb particle for separable verbs provides another degree of freedom in the production of PO datives. A verb is separable when it consists of a lexical core and a separable particle (e.g., *<u>terug</u>betalen* 'pay back'). This factor was matched between verb bias conditions and was therefore not included in the analyses up to now. Unlike English, Dutch has two options for placing the verb particle (1a-b)

1a) Het kind betaalt het geld terug aan de moeder
The child pays the money back to the mother
1b) Het kind betaalt het geld aan de moeder terug
The child pays the money to the mother back

Additionally, PO datives have a third variant which is the non-canonical word ordering in 1c.[2]

1c) Het kind betaalt aan de moeder het geld terug
The child pays to the mother the money back

Thus, besides the additional non-canonical variant in 1c, separable verbs provide two more word ordering options in PO datives (1a and 1b). Whether the availability of multiple word ordering options delays production (just as the availability of equipotent syntactic alternatives does) can be examined by assessing the influence of verb separability on verb onsets in both sentence structures. If verb separability affects verb onsets in DO and PO sentences, then verb retrieval at sentence onset includes preparation of the verb particle in both sentence types. However, if verb separability only affects (or has a greater influence on) the initiation of PO sentences, then this may indicate that the increased flexibility in word ordering options caused by the placement variations of the verb particle causes extra production difficulty in PO datives. A third option is that verb separability does not predict verb onsets, meaning that the verb particle is only prepared later on in the sentence.

In a linear mixed effects model with Structure and Verb separability and an interaction between these two predictors as fixed factors, verb onsets were predicted by a main effect of Verb separability: speakers were faster to initiate nonseparable verbs than separable verbs ($\beta = 55.50$, $SE = 22.50$, $t = 2.47$ in the additive model). Therefore, regardless of the sentence structure, verb separability delayed sentence onset, indicating that verb particles are planned early on in both DO and PO sentences. Although the interaction between Structure and Verb separability was not significant ($\beta = 13.11$, $SE = 34.98$, $t = 0.38$), the onset difference between separable and non-separable verbs was larger in PO datives (63 ms) than in DO datives (28 ms). The numerical difference in effect size suggests that increased flexibility due to multiple placement options for the indirect object in PO datives led to a stronger effect of verb separability in this sentence type.

Planning scope and syntactic flexibility

In sum, both the association between production speed and verb bias effects and the influence of verb separability on sentence onsets indicate that global pre-planning can be quite

[2] DO datives also have a non-canonical word ordering variant in Dutch, but only when replacing the direct object by a personal pronoun, e.g. the *het* ‘it’ or *ze* ‘them’. For example: Het kind betaalt het de moeder terug (*The child pays it the mother back,* ‘The child pays it back to the mother’.)

elaborate and can differ across speakers. Findings also suggest that the scope of planning determines to what degree effects of syntactic flexibility are manifested: A broad planning scope may be a prerequisite for structures to be able to engage in abstract syntactic competition. Vice versa, syntactic flexibility may extend grammatical planning scope by promoting a more global planning strategy. Myachykov et al. (2013) found support for this hypothesis by comparing eye-voice spans of English and Russian speakers describing pictures displaying transitive events. They found that speakers from the more flexible language (i.e., Russian) showed larger eye-voice spans for the subject rather than the object noun when describing pictures of transitive events, suggesting parallel production. In contrast, speakers of the less-flexible language (i.e., English) showed longer eye-voice spans for object than subject nouns, indicative of more incremental production. Similarly, in the present study the existence of two equipotent syntactic frames for formulating sentences with weak bias verbs might have promoted elaborate pre-planning. Namely, with two feasible syntactic structures, both the indirect and the direct object might be activated at verb onset, since both are candidates for becoming the next produced increment (the indirect object is the immediate post-verbal element in the DO dative and the direct object is the immediate post-verbal element in the PO dative). Future research may further evaluate the hypothesis that syntactically flexible sentences involve elaborate pre-planning by manipulating the accessibility of post-verbal material in such a way that effects on planning scope become manifest (cf. Wagner, et al., 2010; Wheeldon, Ohlson, Ashby, & Gator, 2013). Another possibility is to use eye-tracking during sentence planning to examine sentence-initial planning load more extensively (e.g., by looking at eye-voice spans as in Myachykov et al., 2013).

Summary

Taken together, our results provide an indication that equipotent syntactic alternatives can compete during sentence production and suggest that inhibitory control can play a role in resolving this competition. These results contrast with those from the study by Ferreira (1996). As discussed in the introduction, there are a couple of reasons for this discrepancy.

Firstly, in Ferreira's study verb frequency and sentence structure were not matched between conditions (i.e., non-alternator verbs were less frequent than alternator verbs and in all non-alternator/constraining conditions PO datives had to be produced, cf. Myachykov et al., 2013).

Secondly, different experimental paradigms were used (i.e., fragment completion and RSVP). In Ferreira's fragment completion task, participants were first presented with the subject and the verb of the to-be-produced sentence (e.g., *I gave*) and subsequently (after 1500 ms) with the remaining sentence fragments (e.g., *toys/children/to*). They were instructed to create, as quickly as possible, fluent and well-formed utterances when the post-verbal materials were shown. Upon speech onset, a timing bar was shown on screen urging speakers to complete their utterance quickly. Onset latency was measured from the onset of presenting the post-verbal words until the onset of the initial constituent (i.e., *I gave*). Instead, in the RSVP paradigm, speakers were given a preamble consisting of the sentential subject which they were instructed not to repeat (e.g., *the guard…*) and sentence onsets were measured at the verb (e.g., ... *gave the banana*). Speakers were supposed to regenerate the sentence after its rapid serial presentation from the underlying conceptual structure instead of using constituents that were presented on screen during sentence formulation. These procedural differences may have led speakers to employ different strategies. Time pressure in Ferreira's task might have caused speakers to engage in piecemeal planning (cf. Myachykov et al., 2013; Ferreira & Swets, 2002).

Thirdly, in the current experiment we contrasted the production of sentences with weak vs. strong bias verbs. Investigating syntactic flexibility via this graded contrast might give other results than when comparing the production of verbs that have two subcategorization frames vs. only one frame. Finally, the languages under investigation are different (English vs. Dutch). Unlike English, Dutch has an extra non-canonical word order variant for each dative construction (Van der Beek, 2004). As well as this, Dutch has fewer non-alternating verbs than English. Whether these procedural and language-related differences have an influence on dative sentence production is a question for further research.

3.5 Conclusion

In conclusion, the present experiment shows that syntactic alternatives can compete for selection and suggests that selective inhibitory control may facilitate rapid selection between them. Results are in line with comprehension and production theories including a competition component (Kempen, 2013) and with findings from a cross-linguistic study by Myachykov and colleagues (2013). More research is needed to examine the link between syntactic flexibility and planning scope in more detail.

Acknowledgements

We would like to thank Agnieszka Konopka and Gerard Kempen for helpful discussions and Annelies van Wijngaarden and student assistants from the Psychology of Language Department for help with data collection and processing.

Appendix A

verb pair (strong/weak)	verb pair english	coll. strength	sentence frame Dutch	sentence frame English
PO biased				
voorleggen/voorstellen	submit/propose	33.13/2.72	de student **legt/stelt** het plan **voor** aan de professor	the student submits/proposes the plan to the Professor
uitdelen/nalaten	distribute/leave	18.54/1.07	de opa **deelt/laat** de spullen **uit/na** aan de kleinkinderen	the grandfather distributes/leaves his things to his grandchildren
overmaken/terugbetalen	transfer/pay back	23.29/2.72	het kind **maakt/betaalt** het geld **over/terug** aan de moeder	the child transfers/pays back the money to the mother
slijten/verpatsen	sell/sell out	6.14/2.56	de juwelier **slijt/verpatst** de ketting aan de oma	the jeweler sells/sells out the necklace to the grandmother
schenken /cadeaudoen	donate	13.28/0.86	de directeur **schenkt/doet** een cheque cadeau aan de werknemer	the manager donates the employee the voucher
afleggen/zweren	swear	9.73/0.57	de getuige **legt/zweert** de eed **af** aan de rechter	the witness swears his oath to the judge
melden/verkondigen	report/proclaim	5.33/0.65	de president **meldt/verkondigt** de leugens aan het volk	the president reports/proclaims the lies to the people
overdragen/teruggeven	transfer/return	41.39/2.85	de koningin **draagt/geeft** de ketting **over/terug** aan de burgemeester	the queen transfers/returns the post chain to the Mayor
versturen/zenden	send	2.36/1.11	de man **verstuurt/zendt** een telegram aan de advocaat	the man sends a telegram to the lawyer
bekend_maken/verklappen	make known/reveal	6.14/0.43	de goochelaar **maakt/verklapt** de truc **bekend** aan de clown	the magician makes/reveals the trick known to the clown
verklaren/uitleggen	declare/explain	11.61/0.33	de dief **verklaart/legt** de diefstal **uit** aan de politie	the villain declares/explains the theft to the police
verstrekken/verdelen	provide/divide	5.81/1.02	de organisatie **verstrekt/verdeelt** het voedsel aan de bevolking	the organization provides/divides food packages to the population
verlenen/bieden	grant/offer	13.67/0.71	de zakenman **verleent/biedt** hulp aan de oma	the businessman gives/offers assistance to the grandmother
verkopen/lenen	sell/borrow	66.26/3.63	de man **verkoopt/leent** de boeken aan het meisje	the man sells/borrows the books to the girl
DO biased				
leren/voorlezen	teach/read out	32.83/0.20	de lerares **leert/leest** de kinderen het verhaal voor	the teacher teaches/reads out the children the story
aanraden/aanprijzen	recommend	14.39/0.20	de verkoper **raadt/prijst** het gezin de tv **aan**	the seller recommends the television to the family
voorhouden/tonen	show	14.12/0.27	de ontwerper **houdt/toont** het model het jurkje **voor**	the sewer shows the model the dress
verzoeken/vragen	request/ask	10.49/0.99	de conducteur **verzoekt/vraagt** de passagiers stilte	the conductor requests/asks the passengers silence
adviseren/aanbevelen	advise/recommend	8.64/0.48	de arts **adviseert/beveelt** de vrouw de zalf aan	the beautician advises/recommends the woman the cream
verbieden/weigeren	forbid/refuse	8.35/0.62	de moeder **verbiedt/weigert** het kind de chocolade	the mother forbids/refuses the child the chocolate
garanderen/bevestigen	guarantee/confirm	7.85/0.34	de handelaar **garandeert/bevestigt** de koper de winst	the presenter guarantees/confirms the player the profit
wijzen/latenzien	point/show	7.05/0.76	de chauffeur **wijst/laat** de fietser de route zien	de driver points/shows the biker the route
besparen/sparen	save	6.89/0.32	de korting **bespaart/spaart** de dame geld	the discount will save the lady money
geven/aanreiken	give/hand	5.56/0.84	de bewaker **geeft/reikt** de aap de banaan **aan**	the guard gives/hands the monkey the banana
ontnemen/benemen	obstruct/deprive	5.25/0.48	het hek **ontneemt/beneemt** de buurman het uitzicht	the fence obstructs/deprives the view of the neighbor
ontfutselen/aftroggelen	pilfer	2.40/0.64	de artiest **ontfutselt/troggelt** de man het horloge af	the artist pilfers the man his watch
voorschotelen/serveren	serve	2.72/0.48	de ober **schotelt/serveert** de klant de maaltijd **voor**	the waiter serves the customer the meal
uitdelen/voeren*	distribute/feed	18.54/0.51	de kinderen **delen/voeren** het brood **uit** aan de eenden	the children distribute/feed the bread to the ducks

* excluded

References

Altmann, L. J. P., & Kemper, S. (2006). Effects of age, animacy and activation order on sentence production. *Language and Cognitive Processes, 21*(1-3), 322-354.

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2 ed.): Linguistic Data Consortium.

Baayen, R. H. (2008a). *Analyzing linguistic data* (Vol. 505): Cambridge University Press Cambridge, UK.

Baayen, R. H. (2008b). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". *R package version 0.953*.

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

Bates, D., Maechler, M., & Dai, B. (2009). lme4: linear mixed-effects models using S4 classes. 2009. R Package, version 0.999375–31.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition, 114*(3), 455-461.

Bjorklund, D. F., & Harnishfeger, K. K. (1990). The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Developmental review, 10*(1), 48-71.

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

Bock, K. (1986). Syntactic persistence in language production. *Cognitive Psychology, 18*(3), 355-387.

Bock, K., & Brewer, W. F. (1974). Reconstructive recall in sentences with alternative surface-structures. *Journal of Experimental Psychology, 103*(5), 837-843.

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? *Journal of Experimental Psychology-General, 129*(2), 177-192.

Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations - bridging the syntactic cleft. *Psychological Review, 99*(1), 150-171.

Bock, K. J., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition, 21*(1), 47-67.

Boersma, P., & Weenink, D. Praat: doing phonetics by computer [computer program]. Version 5.3. 39. 2013.

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. *Cognitive foundations of interpretation*, 69-94.

Brown, M., Savova, V., & Gibson, E. (2012). Syntax encodes information structure: Evidence from on-line reading comprehension. *Journal of Memory and Language, 66*(1), 194-209.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science, 26*(5), 609-651.

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

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass, 6*(5), 259-278.

Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition, 98*(2), 105-135.

Clifton, C., & Staub, A. (2008). Parallelism and competition in syntactic ambiguity resolution. *Language and Linguistics Compass, 2*(2), 234-250.

Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences, 31*(5), 593-611.

De Jong, R., Coles, M. G., & Logan, G. D. (1995). Strategies and mechanisms in nonselective and selective inhibitory motor control. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 498.

De Smedt, K. (1990). IPF: An incremental parallel formulator. *Current research in natural language generation*, 167-192.

De Smedt, K. (1994). Parallelism in incremental sentence generation. *Parallel natural language processing*, 421-447.

DeGutis, J., Wilmer, J., Mercado, R. J., & Cohan, S. (2012). Using regression to measure holistic face processing reveals a strong link with face recognition ability. *Cognition*.

Dowty, D. (2003). The dual analysis of adjuncts/complements in categorial grammar. In E. L. C. M. C.Fabricius-Hansen (Ed.), *Modifying Adjuncts* (pp. 33-66).

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & psychophysics, 16*(1), 143-149.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715-736.

Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language, 46*(1), 57-84.

Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*(5), 724-755.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. *Language and Cognitive Processes, 21*(7-8), 1011-1029.

Forstmann, B. U., Jahfari, S., Scholte, H. S., Wolfensteller, U., van den Wildenberg, W. P., & Ridderinkhof, K. R. (2008). Function and structure of the right inferior frontal cortex predict individual differences in response inhibition: a model-based approach. *The Journal of Neuroscience, 28*(39), 9790-9796.

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.

Green, M. J., & Mitchell, D. C. (2006). Absence of real evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 55*(1), 1-17.

Gropen, J., Pinker, S., Hollander, M., Goldberg, R., & Wilson, R. (1989). The learnability and acquisition of the dative alternation in English. *Language, 65*(2), 203-257.

Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38*(11), 1531-1549.

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.

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.

Jones, L. B., Rothbart, M. K., & Posner, M. I. (2003). Development of executive attention in preschool children. *Developmental Science, 6*(5), 498-504.

Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. *Memory & Cognition, 35*(5), 925-937.

Kelly, M. H., Bock, J. K., & Keil, F. C. (1986). Prototypicality in a linguistic context: Effects on sentence structure. *Journal of Memory and Language, 25*(1), 59-74.

Kempen, G. (2013). Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics*.

Kempen, G., & Harbusch, K. (2004). A corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann & C. Habel (Eds.), *Multidisciplinary approaches to language production* (pp. 173-181). Berlin: Mouton De Gruyter.

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

Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science, 1*(3), 273-290.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *Journal of Memory and Language, 66*(1), 143-162.

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

Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language, 31*(6), 713-733.

McRae, K., Spivey-Knowlton, M. J., & Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language, 38*(3), 283-312.

Melinger, A., & Dobel, C. (2005). Lexically-driven syntactic priming. *Cognition, 98*(1), B11-20.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*(1), 49-100.

Mohamed, M. T., & Clifton, C. (2011). Processing temporary syntactic ambiguity: the effect of contextual bias. *Quarterly Journal of Experimental Psychology, 64*(9), 1797-1820.

Mostofsky, S. H., & Simmonds, D. J. (2008). Response inhibition and response selection: two sides of the same coin. *Journal of Cognitive Neuroscience, 20*(5), 751-761.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: the case of English and Russian. *Quarterly Journal of Experimental Psychology, 66*(8), 1601-1619.

Nigg, J. T. (2000). On inhibition/disinhibition in developmental psychopathology: Views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological Bulletin, 126*(2), 220-246.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass, 4*(10), 906-924.

Pessoa, L. (2009). How do emotion and motivation direct executive control? *Trends in Cognitive Sciences, 13*(4), 160-166.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*(6), 633-654.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38*(3), 265-282.

Prat-Sala, M. (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.

R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Ridderinkhof, R. K. (2002). Micro-and macro-adjustments of task set: Activation and suppression in conflict tasks. *Psychological research, 66*(4), 312-323.

Segaert, K., Menenti, L., Weber, K., & Hagoort, P. (2011). A paradox of syntactic priming: why response tendencies show priming for passives, and response latencies show priming for actives. *Plos One, 6*(10), e24209.

Shao, Z., Meyer, A. S., & Roelofs, A. (2013). Selective and nonselective inhibition of competitors in picture naming. *Memory & Cognition, 41*(8), 1200-1211.

Shao, Z., Roelofs, A., & Meyer, A. S. (2012). Sources of individual differences in the speed of naming objects and actions: The contribution of executive control. *Quarterly Journal of Experimental Psychology, 65*(10), 1927-1944.

Slusarek, M., Velling, S., Bunk, D., & Eggers, C. (2001). Motivational effects on inhibitory control in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*(3), 355-363.

Spaulding, T. J. (2010). Investigating mechanisms of suppression in preschool children with specific language impairment. *Journal of Speech, Language and Hearing Research, 53*(3), 725.

Tomlin, R. S. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. *Language and conceptualization*, 162-189.

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 der Beek, L. (2004). *Argument order alternations in Dutch.* Paper presented at the LFG'04 Conference.

Van Gompel, R. P., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 52*(2), 284-307.

Van Gompel, R. P., Pickering, M. J., & Traxler, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language, 45*(2), 225-258.

Verbruggen, F., & Logan, G. D. (2008). Response inhibition in the stop-signal paradigm. *Trends in Cognitive Sciences, 12*(11), 418-424.

Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40*(2), 479-483.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75*(2), 105-143.

Vosse, T., & Kempen, G. (2009). In defense of competition during syntactic ambiguity resolution. *Journal of Psycholinguistic Research, 38*(1), 1-9.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the Flexibility of Grammatical Advance Planning During Sentence Production: Effects of Cognitive Load on Multiple Lexical Access. *Journal of Experimental Psychology-Learning Memory and Cognition, 36*(2), 423-440.

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

4

SYNTACTIC FLEXIBILITY AND PLANNING SCOPE

The effect of verb bias on advance planning during sentence recall

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.

Abstract

In sentence production, grammatical advance planning scope depends on contextual factors (e.g., time pressure), linguistic factors (e.g., ease of structural processing), and cognitive factors (e.g., production speed). The present study tests the influence of the availability of multiple syntactic alternatives (i.e., syntactic flexibility) on the scope of advance planning during the recall of Dutch dative phrases. We manipulated syntactic flexibility by using verbs with a strong bias or a weak bias towards one structural alternative in sentence frames accepting both verbs (e.g., strong/weak bias: *De ober schotelt/serveert de klant de maaltijd [voor]* 'The waiter dishes out/serves the customer the meal'). To assess lexical planning scope, we varied the frequency of the first post-verbal noun (N1, Experiment 1) or the second post-verbal noun (N2, Experiment 2). In each experiment, 36 speakers produced the verb phrases in a Rapid Serial Visual Presentation (RSVP) paradigm. On each trial, they read a sentence presented one word at a time, performed a short distractor task, and then saw a sentence preamble (e.g., *De ober…*) which they had to complete to form the presented sentence. Onset latencies were compared using linear mixed effects models. N1 frequency did not produce any effects. N2 frequency only affected sentence onsets in the weak verb bias condition and especially in slow speakers. These findings highlight the dependency of planning scope during sentence recall on the grammatical properties of the verb and the frequency of post-verbal nouns. Implications for utterance planning in everyday speech are discussed.

4.1 Introduction

In sentence production, words are retrieved from the mental lexicon and combined into grammatical sequences. This is done at an impressive rate. Dutch speakers have an average conversational speaking rate of 4.23 syllables per second (Verhoeven, De Pauw, & Kloots, 2004). According to Indefrey and Levelt (2004) encoding of a single word and preparing the articulation of its first syllable takes 600 ms on average. For sentences to be produced fast and fluently, it is therefore inevitable that part of the speech plan is prepared before speech onset.

Bock and Levelt (1994) proposed that speech planning occurs in three stages. First a message is constructed that specifies the intended meaning of the utterance. In a second stage, the message is grammaticalized via functional and positional encoding processes. Functional encoding comprises the retrieval of lexical concepts and lemmas (i.e., grammatical representations of words) and the assignment of these lemmas to grammatical roles (e.g., the subject role). During positional encoding, lexical items are given a serial order and syntactic structures are built. The third stage is the construction of the sound form of the utterance during phonological encoding. Thus, in this model the constituent structure of sentences is generated in two stages. Other authors have argued for direct, single-stage mapping between the message and the constituent structure (Cai, Pickering, & Branigan, 2012; Pickering, Branigan, & McLean, 2002). The models differ, among other things, in their predictions about lexical influences on grammatical encoding processes. According to single-stage models, lexical accessibility can directly influence word order–and thus syntactic choices–such that highly accessible (e.g., high frequency) units are prioritized. In contrast, multiple-stage models postulate that the thematic structure of the message is first mapped to a functional-grammatical structure, thereby driving syntactic choice independently of lexical influences (Chang, 2002; Chang, Dell, & Bock, 2006).

Earlier research on the advance planning of sentences with varying phrase structures has found evidence for the phrase as the default planning unit, for instance by showing that speakers take longer to initiate sentences starting with complex (e.g., *the dog and the hat move*) than simple noun phrases (e.g., *the dog moves*) (Martin, Crowther, Knight, Tamborello, & Yang, 2010; Smith & Wheeldon, 1999; Wheeldon, Ohlson, Ashby, & Gator, 2013). However, in other studies using the same sentence types, lexical planning scopes ranging from single lexical items to entire clauses have been found (Griffin, 2001; Meyer, 1996). The production of more complex sentences (such as descriptions of transitive events) has also proven to be flexible, with speakers sometimes prioritizing the encoding of single linguistic elements and sometimes encoding an abstract plan of the utterance before speaking (Gleitman, January, Nappa, & Trueswell, 2007; Kuchinsky & Bock, 2010; Van de Velde, Meyer, & Konopka, 2014). Together, these findings suggest that planning scope is not fixed but variable.

Several studies have attempted to identify the conditions under which speakers decrease or extend their scope of planning. Ferreira and Swets (2002) examined how time pressure affects grammatical planning scope in two experiments using two-digit sums (e.g., 9 + 7 = ?). Speakers were instructed to formulate their answers to these sums as follows: (a) "sixteen," (b) "sixteen...is the answer," or (c) "The answer is... sixteen." Besides utterance type, the difficulty of the arithmetic problem was manipulated. In a second experiment, participants were prompted to start speaking as quickly as possible through the use of a deadline procedure. Speech onsets and utterance durations were measured. In both experiments, speech onset latencies were similar for all three utterance types but speakers initiated their utterances later when problem difficulty increased. Only in Experiment 2 did utterance *duration* also depend on problem difficulty such that answers to difficult arithmetic sums took longer to formulate. This suggests that speakers planned and spoke simultaneously when they were prompted to start their utterance immediately. In contrast, when there was no pressure to start speaking quickly, speakers made use of more extensive advance planning when generating complex constructions and hence showed longer onset latencies when problem difficulty increased.

In addition to contextual factors, speaker characteristics may also have an influence on the amount of advance planning. Wagner, Jescheniak and Schriefers (2010) found large differences in planning scope related to speaking rate. They used a picture-word interference paradigm to measure grammatical planning scope: speakers had to produce simple sentences describing objects on a screen (e.g., *the frog is next to the mug*). Unrelated or semantically related auditory distractors were presented at picture onset. Onset latencies were measured to examine semantic interference effects on the first and second object, which indexed the scope of grammatical encoding. Dividing their sample into a group of slow speakers and a group of fast speakers (based on their average naming latencies in an unrelated distractor condition), they found that slow speakers showed larger interference effects on the second noun, suggesting that they engaged in more advance planning than fast speakers. A similar difference was found by Wheeldon, Ohlson, Ashby, and Gator (2013), who used the same type of sentences, but made use of a picture preview to increase the accessibility of the second noun. When it was known which sentence position the previewed object would occupy, slow speakers showed a larger preview benefit than fast speakers. The authors concluded that slow speakers have a larger lexical processing scope than fast speakers.

Planning scope is also sensitive to linguistic factors such as ease of structural assembly. In the study by Wagner et al. (2010) described above, speakers had to produce simple sentences (e.g., *the frog is next to the mug*) and more complex structures (e.g., *the red frog is next to the red mug*). In two of their experiments, speakers only used one sentence type (simple or with adjectives) while in another experiment they had to switch between sentence types. Wagner et al. found that speakers made use of more parallel planning, showing interference from the semantically related auditory distractors on both the first and the second noun, when they only had to produce one sentence type. Opperman et al. (2010) showed the same pattern for phonological advance planning. These results suggest that when structures are easy to produce (because they are repeated), lexical planning scope may be expanded. Konopka (2012) tested this hypothesis by comparing the production of sentences beginning with conjoined noun phrases (e.g., *the saw and the axe are above the cup*) after manipulating structural and lexical processing ease using structural priming and a lexical frequency manipulation. Conjoined noun phrases either contained semantically related or unrelated objects. Lexical planning scope was measured as the degree of semantic interference from the first onto the second object. Semantic interference was only found for structurally primed sentences beginning with easily accessible words. This result indicates that speakers engage more in advance planning when sentence structures are easy to assemble than when structures are more difficult to construct. Relatedly, Konopka and Meyer (2014) found that if relational encoding in picture naming (i.e., the encoding of causal relations instead of individual characters) is facilitated through the use of easily apprehensible events and/or priming the to-be-used structure, speakers shift from a piecemeal towards a more parallel planning strategy.

Another linguistic factor that may influence planning scope is syntactic flexibility (Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013; see also Ferreira, 1996). Syntactic flexibility refers to the availability of syntactic choices during grammatical encoding. According to some models, syntactic flexibility leads to competition between syntactic alternatives (Dell, Chang, & Griffin, 1999; Ferreira, 1996; Kempen, 2013; Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013). For example, when producing a dative sentence a speaker can choose between a prepositional object (PO, e.g., *Steve gives the apple to Susan*) and a double object (DO, e.g., *Steve gives Susan the apple*) frame. Under some circumstances, structural alternatives may be equipotent, meaning that they are activated to the same degree. At the choice point of the dative sentence, i.e., after the verb (*give*), when a structural alternative is uniquely determined, the grammatical encoding system can either activate the indirect object or the direct object slot. The degree to which the indirect object and direct object are prepared while the system decides which structure to produce is reflected in the planning scope. If both objects are lexically encoded, the lexical advance planning scope is wider than when only one option is prepared.

Myachykov et al. (2013) examined whether the number of available syntactic options used in a picture naming task in one experiment could predict sentence-initial processing load in a second picture naming experiment. In Experiment 1, speakers of English and Russian were encouraged to produce as many structural alternatives for a set of depicted events as possible within 15 seconds per picture. Since Russian grammar offers more options for building transitive constructions than English grammar, Russian participants were expected to use a wider range of syntactic options than English participants. In Experiment 2, a new group of English and Russian participants described the same set of pictures while their eye-gaze was measured. The goal of the study was to link syntactic flexibility in the first experiment to sentence initial processing load in the second experiment, indexed by sentence onset latency and eye-voice span (i.e., the temporal lag between the offset of the last fixation to a referent and producing that referent's name). As expected, Russian speakers used a wider range of syntactic alternatives than English speakers. In Experiment 2 they were also slower to initiate their sentences and showed longer eye-voice spans for sentence-initial subject nouns than for object nouns. This eye-voice span pattern suggests a planning strategy involving the (partial) preparation of non-initial increments or an abstract framework before sentence onset (Bock, Irwin, Davidson, & Levelt, 2003). In contrast, English speakers showed shorter onset latencies and longer eye-voice spans for object nouns than subject nouns, indicative of more incremental production. More importantly, in both groups syntactic flexibility in Experiment 1 predicted sentence-initial latency effects in Experiment 2, indicating that more syntactic alternatives led to longer onset latencies after accounting for the effect of language (i.e., English vs. Russian).

The above findings indicate that syntactic flexibility may lead to a higher sentence-initial processing load and longer eye-voice spans for subject than object nouns. The authors interpret these results as demonstrating competition between syntactic frames. They propose that competition between frames leads to an expansion of planning scope, as a larger part of the syntactic plan needs to be prepared prior to speech onset because of the necessity of making a syntactic choice. However, Myachykov's study does not provide a direct link between syntactic flexibility and grammatical planning scope; instead the study only shows that syntactic flexibility and sentence initial processing load are linked.

The current study aimed to connect syntactic flexibility and planning scope directly by adding a manipulation of lexical accessibility. The question was whether increased syntactic flexibility lead to increasing planning scope. In an experiment using rapid serial visual presentation (RSVP; Lombardi & Potter, 1992; Potter & Lombardi, 1990; 1998; see below for details) participants produced dative sentences featuring verbs with a bias towards the PO or DO dative. Syntactic flexibility was manipulated by using verbs varying in the strength of bias for one of the structural alternatives (see 1a and 1b).

1a) *De ober serveert de koning het feestmaal.* (weak DO bias)
'The waiter serves the king the banquet.'

1b) *De ober schotelt de koning het feestmaal voor.* (strong DO bias)
'The waiter dishes the king the banquet out'
['The waiter dishes out the banquet to the king']

Verb bias is the preference of a verb for a syntactic structure, based on its frequency of co-occurrence with the verb (Colleman, 2009; Stallings, MacDonald, & O'Seaghdha, 1998). The strength of these verb biases determines the likelihood that structural alternatives will be selected when the verb is used. Therefore, verbs without a significant preference for one sentence structure will support the selection of both alternatives to a similar degree, which leads to syntactic flexibility in the grammatical encoding system.

In Experiment 1, the frequency of the first object noun (N1) was manipulated (high vs. low frequency) in addition to syntactic flexibility. The second object noun (N2) was the same in each sentence frame across conditions and had low frequency (see 2a and 2b).

2a) *De ober serveert/schotelt de monarch het feestmaal voor*. (low frequency N1)
'The waiter serves/dishes out the monarch the banquet.'

2b) *De ober serveert/schotelt de koning het feestmaal voor.* (high frequency N1)
'The waiter serves/dishes out the king the banquet.'

In Experiment 2, we manipulated the frequency of the second object noun (N2; high vs. low frequency). The first object noun always had high frequency (see 3a and 3b).

3a) *De ober serveert/schotelt de klant het feestmaal voor.* (low frequency N2)
'The waiter serves/dishes out the customer the banquet.'

3b) *De ober serveert/schotelt de klant de maaltijd voor.* (high frequency N2)
'The waiter serves/dishes out the customer the meal.'

By examining how both frequency manipulations affected speech onsets, we can make inferences about the lexical planning scope for sentences with a high and low degree of syntactic flexibility. There are a number of different possibilities. First, speakers may engage in strictly incremental planning and prepare only the verb when they initiate their utterance (Recall that the subject noun phrase was provided at the beginning of the trial, and the verb was the first word the speakers had to produce). If this is true, there should be no effect of noun frequency at all. If speakers consistently encode their utterance up to and including the first noun before speech onset, there should be a main effect of N1 frequency, with faster onsets for utterances starting with high frequency than lower frequency nouns. If their lexical planning scope reaches, by default, up to the second noun, we expect to find a main effect of N2 frequency. More interestingly, if planning scope changes with the syntactic flexibility of the verb, there should be an interaction between noun frequency and syntactic flexibility (i.e., weak vs. strong verb bias). In line with findings of Myachykov et al. (2013) we hypothesized that syntactic flexibility would expand planning scope and that noun frequency would only have an effect on speech onsets in sentences with verbs that have no significant bias towards one syntactic frame (i.e., under high syntactic flexibility).

In addition to linguistic factors, we also examined the effect of variations in the speakers' speed of initiating utterances on their planning scope. As noted, in studies by Wagner et al. (2010) and Wheeldon et al. (2013) speakers with long utterance onset latencies were found to have a broader planning scope than speakers with shorter latencies. We examined whether this would also be the case in the current study by computing the participants' average speech onset latency on filler trials and adding average speech rate as a factor to a model predicting sentence onsets. By using production speed as a continuous predictor, we avoided the use of a median split procedure and obtained a more fine-grained measure of production speed.

As noted, we used the RSVP paradigm to elicit utterances with fixed wording (verbs and nouns) and structure. In this paradigm, participants are presented with a sentence in a word-by-word fashion at a high speed (100 ms per word). Subsequently they perform a short distractor task, and then see a sentence preamble (in our case the subject noun phrase), which they have to complete to form the presented sentence. It is assumed that the fast presentation of the sentence and the intervening distractor task lead to the formation of a conceptual representation of the sentence, the wording of which has to be reconstructed during later recall (Potter & Lombardi, 1990). Thus, as in everyday speech production, a conceptual message needs to be translated into a sequence of words (e.g., Bock & Levelt, 1994).

Potter and Lombardi tested their paradigm by presenting five lure words prior to the recall phase (i.e., the sentence preamble). On half of the trials, one of the lure words was conceptually related to one of the words in the to-be-recalled target sentence. After the word list, one probe word (never the conceptually related lure) was presented and participants had to judge whether this word had been part of the previously presented word list. Potter and Lombardi found that during recall, participants tended to exchange target words (verbs and nouns) for lures, but only when they were in line with the conceptual message conveyed by the target sentence (Potter & Lombardi, 1990). Verb exchanges even occurred when the categorization frames of the intruding verb were not compatible with the surface structure of the presented target sentence. Participants restored the grammaticality of the sentences by using a frame congruent with the selected verb (Lombardi & Potter, 1992). These results suggest that the RSVP paradigm indeed taps sentence reconstruction process rather than retrieval of an episodic memory representation of the (linearly ordered) surface structure.

The RSVP paradigm has later been used in several sentence production studies examining constraints on structural priming (e.g., Chang, Bock, & Goldberg, 2003; Griffin & Weinstein-Tull, 2003; Konopka & Bock, 2009; Tooley & Bock, 2014). Here RSVP prime trials preceded structurally matching or mismatching target trials. Priming was measured as the extent to which speakers re-used the prime trial structure on target trials. Critically, the paradigm produced priming effects that were comparable in magnitude to priming effect in picture description tasks (Chang, et al., 2003). This finding further supports the view that the RSVP task taps structural mechanisms of sentence production.

However, there are also obvious differences between sentence recall via the RSVP paradigm and everyday sentence production. Most importantly, the participants do not generate the message based on the thoughts they wish to express, but instead read a sentence and store its content in working memory. Based on this memory representation, the sentence has to be regenerated. The role of verbal working memory in the RSVP task manifests itself in the finding that sentence recall is often (near) verbatim. Potter and Lombardi (1998) explain the verbatim recall with the fact that speakers are likely to re-use the recently activated lexical entries, but stress that these entries are unordered in memory. Hence, during reconstruction in the recall phase, regular sentence production mechanisms are used to linearize lexical items retrieved from memory. Studies of the relationship between verbal working memory and language production are consistent with this view (Acheson & MacDonald, 2009a, 2009b; Bock, 1996; Slevc, 2011).

In sum, RSVP offers a way of studying the generation of sentences that are otherwise not easy to elicit. Although the paradigm has its shortcomings, previous studies have shown that it can be used to tap certain aspects of normal sentence production. For the present purposes it is most important that the way retrieved lexical items are combined into sentences must be based on the participants' general lexico-syntactic knowledge and their prior linguistic experience. We are interested in the processes involved in the retrieval of the verb and its arguments. Retrieval of a weak bias verb should result in the automatic activation of two equipotent syntactic frames (high syntactic flexibility), whereas retrieving a strong bias verb should lead to the activation of one dominant syntactic frame (low syntactic flexibility). By varying the frequency of the post-verbal nouns along with the syntactic flexibility of the verb and examining the consequences for the verb onset times, we investigate how syntactic flexibility influences the activation of upcoming lexical material prior to speech onset.

4.2 Experiment 1

We investigated the effects of verb bias and noun frequency on speakers' planning scope when producing dative verb phrases during sentence recall. In addition, we examined the role of participants' response speed in explaining individual differences in advance planning. We used an RSVP paradigm to elicit sentences with fixed wording and structure: participants constructed dative sentences from a preamble (the first noun phrase, e.g., *The jeweler*) after a rapid word-by-word presentation of the entire sentence (*The/jeweler/sells/the/necklace/to/the/grandmother*). Without repeating the preamble, participants started their utterance by producing the verb, which could have a strong or a weak bias towards one dative alternative.

In addition to varying verb bias, we also manipulated the frequency of the first noun following the verb (i.e., the direct object in PO datives, and the indirect object in DO datives) The second noun was the same for each sentence frame across conditions and had low frequency. Consequently, noun frequency differences were always congruent with the verb's preference and the presented sentence structure. Here we focus only on facilitatory or interfering effects of noun frequency on sentence production. In line with earlier findings, we expect an extension of planning scope, indexed by an effect of N1 frequency, only for sentences with weak bias verbs. In other words, we expect an interaction between verb bias and noun frequency difference on the RTs.

4.2.1 Method

Participants

36 adult native speakers (ages 18 -30 years) of Dutch gave informed consent and participated in the experiment for payment. All participants had normal or corrected-to-normal vision. Consent for conducting the study had been obtained from the Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen.

Materials

Dative verbs with weak and strong biases towards the prepositional dative and double object dative structures were selected based on a corpus analysis of the Dutch dative alternation (Colleman, 2009). In this corpus analysis, collostructional strength was identified for 252 alternating dative verbs. Collostructional strength is the degree of association between one lexical item (in this case a verb) and two or more functionally similar abstract constructions. The degree of association is based on the frequencies of the verb occurring in each of these constructions and on the overall frequencies of the construction in the corpus and is computed using the Fisher exact test. An index of distinctive collostructional strength was calculated as -log(Fisher exact, 10). The higher the index, the stronger the preference is of the verb for the construction. For example, *give* has a strong preference for the DO structure and *show* has a weak preference, with collostructional strengths of 5.56 and 0.27, respectively. Although *give* is an example of a high frequency verb, degree of strength is not correlated with lexical frequency. The corpus study revealed a wide range of collostructional strength for verbs preferring DO and PO constructions, with PO dative preferring verbs displaying a wider range (0.17 – 69.07) than DO dative preferring verbs (0.16 – 40.8).

A total of 28 verb pairs with low (M = 0.57, SD = 0.28) and high (M = 13.35, SD = 10.90) collostructional strength were selected from Colleman (2009), resulting in a weak bias and a strong bias set (e.g., *serveren* 'serve' and *voorschotelen* 'dish out'). For each verb pair, one sentence was constructed which could accept both verbs (e.g., *De ober serveert de klant de maaltijd* 'The waiter serves the customer the meal' vs. *De ober schotelt de klant de* *maaltijd voor* 'The waiter dishes out the customer the meal') in the sentence frame (DO or PO) of the verb's preference. Sentences consisted of one main clause and were written in the present tense. Verb bias conditions and dative structures were matched for verb log lemma frequency, syllable count and separability (CELEX Lexical Database, Baayen, Piepenbrock, & Gulikers, 1995). Separability refers to the possibility of separating the verb core and its particle, as in for example *terug*betalen 'pay back'. Dutch has many separable verb and two placement options for the particle in PO datives: before or after the canonical position of the indirect object. In our experimental stimuli the particle always preceded the indirect object, e.g., het kind betaalt het geld terug aan de moeder 'the child pays the money back to the mother'. Verb lemma frequency was uncorrelated to verb collostructional strength (r = .18). Table 1 shows the descriptive statistics for the verbs organized by structural preference.

Table 1. Descriptive statistics for DO and PO dative preferring verbs used in Experiment 1.

Verb preference		Collostructional strength	Log Lemma frequency	Log N1 frequency	Log N2 frequency
DO dative	Average	6.83	1.38	1.44	0.69
	Stdev	10.32	0.73	0.69	0.21
	Range	.16-41	0.0 - 3.11	0.30 - 2.66	0.30 - 1.00
PO dative	Average	7.09	1.08	1.30	0.56
	Stdev	9.71	0.69	0.60	0.20
	Range	0.17-33	0.0 - 2.64	0.30 - 2.44	0.30 - 0.90

Note: collostructional strength represents the bias for a DO structure for DO preferring verbs and the bias for a PO structure for PO preferring verbs.

In half of the items the first noun had high frequency (M = 1.94, SD = 0.34) and in the other half of the items it was low in frequency (M = 0.81, SD = 0.29). Nouns were matched on other characteristics affecting lexical accessibility: number (plural vs. singular), length, number of syllables, and animacy. The experiment used a 2 (Verb bias) x 2 (N1 frequency) within-participant and within-item factorial design. Four lists of stimuli were created to counterbalance verb bias and N1 frequency, so that each item appeared in a different condition in each list (see Appendix A for a complete list of the stimuli used in the experiment). Within lists, there were seven items in each of the four conditions. In addition, there were ten practice items and 84 filler items used to separate target items.

Norming

We carried out a norming study to evaluate whether target frames carrying different noun combinations and verbs were equally plausible. 40 participants who did not participate in the main experiments were asked to rate the plausibility of the sentences on a scale from 1 (implausible) to 7 (very plausible). Items were randomly assigned to one of four item lists, such that each sentence frame appeared in all four conditions across lists and each verb appeared exactly once per list. Table 2 shows a summary of the norms for the selected sentence frames per condition.

Table 2. Mean plausibility ratings and standard deviations (in parentheses) for strong and weak bias verbs by sentence structure and N1 frequency (High N1 vs. Low N1).

	DO dative		PO dative	
Verb bias	High N1	Low N1	High N1	Low N1
Strong bias	5.97 (1.48)	5.94 (1.45)	5.61 (1.61)	5.25 (1.97)
Weak bias	5.82 (1.48)	6.05 (1.19)	5.52 (1.65)	5.30 (1.83)

Since verbs always showed a bias for the structure they occurred in and noun frequency differences were congruent too, we expected high plausibility ratings across all conditions and sentence types. As expected, ratings did not differ across verb bias $t(1,27) = 0.39$ or noun frequency conditions ($t[1,27] = 0.57$). More importantly, there was no interaction between verb bias and N1 frequency ($F[1,26] = 1.06$). DO datives were rated as more plausible than PO datives ($F[1,26] = 4.57$, $p < .05$). This difference might be due to the fact that DO datives occur more often in Dutch than PO datives (e.g.,Colleman, 2009).

Procedure

Participants were randomly assigned to one of the four item lists. Instructions for the experiment appeared on the screen and participants received ten practice sentences before the experiment started. Sentences were presented in rapid serial visual presentation. The sequence of events for each trial is illustrated in Table 3 (adapted from Konopka & Bock, 2009). The experiment was programmed using Presentation® software (Version 16.3, www.neurobs.com).

Table 3. Sequence of events during each experimental trial.

Duration	Event
200 ms	+
100 ms	De
100 ms	ober
100 ms	serveert
100 ms	de
100 ms	maaltijd
100 ms	aan
100 ms	de
100 ms	klant.
100 ms	##########
533 ms	4 5 2 9 1
100 ms	[screen blanked]
500 ms	twee
10 ms	[screen blanked]
5000 ms (max)	Nee Ja
500 ms	☺ or ☹
	De ober….

After presentation of a 200 ms fixation cross, participants read a sentence, which was presented one word at a time. Word presentation time was 100 ms, similar to earlier English RSVP studies. Although average word length in Dutch is higher than in English (Hagoort & Brown, 2000), a pilot study revealed that participants could process words with a presentation time of 100 ms. Participants were instructed to read the sentence silently and to remember the content (i.e., the message) as they would have to reproduce it later.

They then performed a distractor task, in which they first saw a display of five digits and then had to judge whether a digit (written out in letters: e.g., *twee* 'two') had been part of this array of five digits. They responded by pressing the left (*yes*) or right (*no*) mouse button and had a maximum of five seconds to do so. They were given immediate feedback in the form of a happy face for a correct answer and a sad face for an incorrect answer. On 50% of the trials the correct answer to the distractor task was '*yes*' and on the remaining half the answer was '*no*'. On critical trials the correct answer to the distractor task was always '*yes',* while on filler trials this could vary.

After the distractor task, participants were prompted to repeat the sentence they had read at the beginning of the trial. They were instructed to produce the sentence as quickly as they could without making any mistakes or producing disfluencies. The first noun phrase of that sentence appeared on the screen (e.g., *The waiter*) and participants were asked to complete the sentence with the verb, direct object, and indirect object (i.e., without reading the subject noun phrase of the sentence out loud). They then pressed a button to proceed to the next trial. Responses were recorded and the speech output was transcribed by the experimenter offline. Later, speech onsets were measured manually in Praat (Boersma & Weenink).

Participants were debriefed at the end of the experimental session about the goal of the experiment. They were also asked to describe which strategy they used to remember the sentences and to what extent they remembered the exact form of the sentences. In line with findings from sentence recall studies (e.g., Bock & Brewer, 1974; Potter & Lombardi, 1990, 1998), participants reported that they often tried to reconstruct new surface structures from memory of underlying sentence meanings.

Scoring and analysis

Utterances were scored as having either double object (DO) dative or prepositional object (PO) dative syntax (e.g., *proposes the professor the plan*; *proposes the plan to the professor*). Utterances with intransitive syntax or other constructions were excluded, as were sentences with repeated sentential subjects (e.g., *The student proposed...)*, omitted direct or indirect objects (e.g., *pays back the mother*), verb substitutions (e.g., *asked* instead of *proposed*), or noun substitutions (e.g., *manuscript* instead of *plan*). Additionally, utterances following a wrong answer to the distractor task were also excluded. This was done to control for possible effects of response feedback in the distractor task on subsequent response latencies.

Finally, we eliminated responses with onset latencies longer than 3000 ms or with onset latencies more than 2.5 standard deviations away from the grand mean. The final dataset consisted of 650 responses (316 PO sentences, 334 DO sentences), equivalent to nine scorable responses per participant in each verb bias condition.

All data were analyzed using R (R Development Core Team, 2013) and the R packages *lme4* (Bates, Maechler, & Dai, 2009) and *languageR* (Baayen, 2008b). Analyses on error rates were carried out with mixed logit models (coefficients are given in log-odds). Onset latencies (RTs) were analyzed with linear mixed effects models (coefficients are given in milliseconds). Model factors included Verb bias, N1 frequency, and Sentence structure as categorical factors, after they were centred. In all analyses, we used a backwards elimination procedure, starting from an initial model containing all experimental factors and their interactions and random by-subject and by-item intercepts (Baayen, 2008a; Baayen, Davidson, & Bates, 2008; Jaeger, 2008). Non-significant effects were removed, starting from the highest-order interactions going back to a basic additive model with only main effects. For the remaining fixed effects structure, random slopes were included where mentioned; they were added only if they improved model fit as indicated by likelihood ratio tests (models with maximal random structures showed similar results and are therefore not listed; cf. Barr, Levy, Scheepers, & Tily, 2013). Since MCMC sampling is not implemented in *lme4* for linear mixed effects models with random effects, p values reported in the results were computed based on the t-distribution using the Satterthwaite approximation in the *lmerTest* package (Baayen, et al., 2008; Kuznetsova, Brockhoff, & Christensen, 2013).

4.2.2 Results

We report results of three sets of analyses. First we examined error rates across Sentence structures, N1 frequency and Verb bias conditions in the full dataset (n = 1008). Responses were coded as errors when they contained word substitutions or other constructions than the ditransitive, followed an incorrect answer to the distractor task, and when onset latencies exceeded the outlier threshold. In a second analysis, we tested whether Verb bias, N1 frequency and Sentence structure predicted verb onset latency after excluding errors (n = 650). Finally, we examined how whether the speakers' production speed interacted with the effects of Verb bias and/or Noun frequency.

Error rates

Speakers' accuracy in reproducing target sentences with the correct wording and structure was predicted by N1 frequency: Speakers were more likely to correctly reproduce a sentence when the first noun had high frequency. Furthermore, there was an interaction between Verb bias and Sentence structure (respectively β = -0.36 SE = 0.14, z =-2.51 and β =

-0.70, *SE* = 0.29, z =-2.44). Figure 1 depicts the interaction: Whereas error rates were higher for PO structures with weak bias verbs than strong bias verbs (β = -0.57, *SE* = 0.20, *z* = -2.91), there was no effect of verb bias in DO structures (*z* = 0.59).

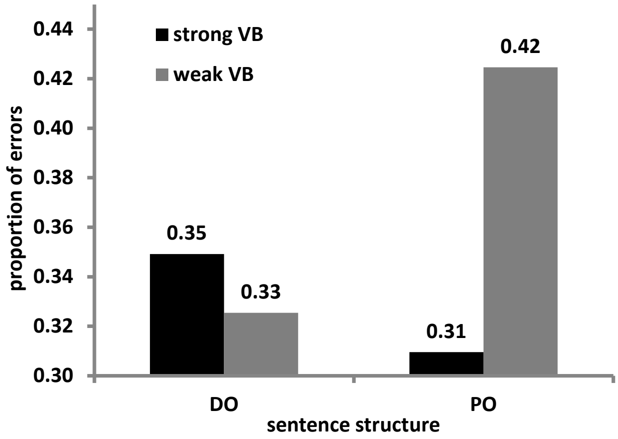

Figure1. Error rates (proportions of errors) as a function of verb bias and sentence structure.

Verb onsets

Unlike error rates, verb onsets did not show an effect of N1 frequency (*t* = -0.71). Verb bias and Sentence structure did not produce any main effects or interactions either (*t*s < -0.63).

Fast and slow speakers

To examine how differences in speech onset latencies were related to the performance on the RSVP task, we added average production speed to a model predicting verb onsets from Sentence structure, N1 frequency and Verb bias. To this end, we measured onset latencies on the filler trials as a neutral index of production speed. For the scoring of the filler trials we used the same criteria as for target trials: responses were excluded that a) did not have the correct wording or structure, b) followed a wrong answer to the distractor task, or c) had onset latencies longer than 3000 ms and onsets more than 2.5 standard deviations away from the grand mean. Based on the remaining trials (81% of all trials, equivalent to 61 trials per participant on average) we computed the average production speed per participant. Adding this continuous factor to the full model predicting verb onsets yielded a main effect of Production speed (β = 953.49, *SE* = 132.65, *t* =7.19, in a model with by-item random slopes for Production speed). There were no interactions between Production speed and the experimentally manipulated factors.

4.2.3 Discussion

The results of Experiment 1 only show suggestive evidence of extended planning as a result of increased syntactic flexibility. Results of the analysis on error rates suggest that planning was harder for sentences featuring weak bias verbs than for those with strong bias verbs, but only in PO datives.

One reason why syntactic planning of PO and DO datives may differ is that producing a PO structure involves more word ordering choices. First of all, speakers may choose to leave out the indirect object—which is the reason some theories regard PO structures as transitives with an optional adjunct (cf. Brown, Savova, & Gibson, 2012; Dowty, 2003)—leading to the production of an ordinary transitive with only one obligatory argument (e.g., *submits the plan* instead of *submits the plan to the professor*). Secondly, placement of the verb particle for separable verbs (e.g., *terugbetalen* [pay back]) is also flexible for PO datives. Unlike English, Dutch has two options for placing the verb particle in a PO dative: a) before the indirect object and b) after the indirect object. These additional word ordering options add yet another degree of flexibility to the production of PO structure. In addition, the PO dative is less common than the DO dative (31 vs. 69% in Colleman, 2009). Altogether, the increased difficulty of structural processing for PO datives may have led speakers to build PO structures more incrementally than DO structures (Konopka, 2012). Consequently, any effect of syntactic flexibility in PO datives may not be visible at sentence onset, but only in errors. Conversely, DO datives are the easier structures and therefore speakers are not bound to a strictly incremental production strategy. Here, syntactic flexibility can have an effect on sentences initial processing load. Indeed, DO sentences featuring a strong bias verb (M = 751 ms, SD = 249 ms) had numerically shorter onset latencies than DO sentences with a weak bias verb (M = 777 ms, SD = 266 ms), but this difference was not reliable (t = 0.82).

The second experimental manipulation, N1 frequency, only showed an effect on error rates: Sentences with a high frequency N1 were better remembered and therefore were produced with fewer errors than sentences with a low frequency N1. The absence of N1 frequency effects on onset latencies may indicate that the first noun was not planned at verb onset. However, it could also indicate that planning scope incorporates not only the first noun, but a wider range of words. Consequently, sentence initiation times may depend on the time needed to prepare both N1 and N2, especially in syntactically flexible sentences where planning scope is hypothesized to be broader than in inflexible sentences. Since the frequency of N2 was always low in Experiment 1 so that N2 required longer preparation time than N1 (e.g., Miozzo & Caramazza, 2003), any facilitating effect of high frequency N1 might have been concealed in sentences featuring weak-bias verbs.

This hypothesis can be put to test by evaluating the effect of N2 frequency (as a between-item factor) on onset latencies for syntactically flexible sentences. If the lexical item with the longest preparation time determines the sentence onset latency in syntactically flexible sentences, then the frequency of N2 should be able to predict verb onsets in weak-verb bias sentences. Although the frequency of the two nouns in an item was carefully controlled, there were differences in N2 frequency *between* items (ranging from 0.30 to 1). Adding N2 frequency as a continuous factor to a linear mixed effects model predicting verb onsets (see section 4.2), yielded a significant interaction between Sentence structure, Verb bias and N2 frequency (β = -243.61, *SE* = 87.41, *t* = -2.79). Examining this interaction more closely revealed that the interaction between Verb bias and N2 frequency was especially apparent in PO datives (β = 437.20, *SE* = 126.26, *t* =3.46) and not in DO datives (*t* = -0.83). When producing PO datives, speakers were faster to initiate sentences featuring weak bias verbs as N2 frequency increased (β = -323.16, *SE* = 150.06, *t* = 3.46). In contrast, N2 frequency did not produce an effect in sentences with strong bias verbs (*t* = 0.89). Taken together, these results provide an explanation for the absence of an interaction between Verb bias and N1 frequency in Experiment 1: any effect of N1 frequency on onset latencies was disguised by the parallel retrieval of the lower frequency N2.

To obtain further experimental support for the possibility that planning scope in syntactically flexible sentences includes N2, we carried out a second experiment in which we manipulated the frequency of the second noun (the direct object in DO structures and the indirect object in PO structures), while keeping N1 frequency constant, i.e. all N1s had high frequency.

4.3 Experiment 2

The goal of Experiment 2 was to test whether syntactic flexibility could lead to an increased grammatical planning scope by manipulating the frequency of N2. We hypothesize that syntactic flexibility expands planning scope up to and including N2 during the recall of dative verb phrases. We thus expect to see an influence of N2 frequency on onset latencies only for syntactically flexible sentences (i.e., the weak verb bias condition). In other words, there should be an interaction between Verb bias condition and N2 frequency.

4.3.1 Method

Participants

A different group of 36 adult native speakers (ages 18 – 30 years) of Dutch gave informed consent and participated in the experiment for payment. All participants had normal or corrected-to-normal vision.

Materials

Verb pairs were mostly similar to those used in Experiment 1. Four new verb pairs were added. Table 4 shows the descriptive statistics for the set of 32 verb pairs.

Table 4. Descriptive statistics for DO dative and PO preferring verbs used in Experiment 2.

		Collostructional strength	Log Lemma frequency	Log N1 frequency	Log N2 frequency
Verb preference					
DO dative	Average	6.30	1.30	1.69	1.05
	Stdev	9.80	0.74	0.47	0.70
	Range	0.16-41	0.0 - 3.11	0.78 - 2.66	0.00 - 2.40
PO dative	Average	7.39	1.07	1.60	0.93
	Stdev	10.22	0.68	0.47	0.67
	Range	0.17-41	0.0 - 2.64	0.78 - 2.44	0.00 - 2.37

Note: collostructional strength represents the bias for a DO structure for DO preferring verbs and the bias for a PO structure for PO preferring verbs.

In half of the items the second noun had low frequency (*M* = 0.42, *SD* = 0.28) and in the other half of the items the frequency of the second noun matched the first noun's frequency (i.e., high frequency: *M* = 1.57, *SD* = 0.44). As in Experiment 1, nouns were matched on other characteristics affecting lexical accessibility. By varying the frequency of the second noun, the direction of frequency differences was always congruent with verb bias. Hence, in sentences with PO biasing verbs, the direct object was more frequent than, or equally frequent as, the indirect object. In sentences with DO biasing verbs, the indirect object was more frequent than, or equally frequent as, the direct object (see example 3a and 3b in the Introduction).

The experiment used a 2 (Verb bias) x 2 (N2 frequency) within-participant and within-item factorial design. Four lists of stimuli were created to counterbalance verb bias and noun-frequency conditions, so each item appeared in a different condition in each list (see Appendix B for a complete list of the stimuli used in the experiment). Within lists, there were eight different items in each of the four conditions. 94 fillers were included to separate target items, with ten used at the beginning of the experiment as practice items.

Norming

A new norming study with 60 participants confirmed that target frames carrying different noun combinations and verbs were equally plausible. Table 5 shows a summary of the norms for the selected sentence frames per condition.

Table 5. Mean plausibility ratings and standard deviations (in parentheses) for strong and weak bias verbs by sentence structure and N2 frequency (High N2 vs. Low N2).

	DO dative		PO dative	
Verb bias	High N2	Low N2	High N2	Low N2
Strong bias	5.42 (1.80)	5.85 (1.53)	5.55 (1.71)	5.34 (1.83)
Weak bias	5.27 (1.88)	5.34 (1.73)	5.18 (1.84)	5.34 (1.73)

Ratings did not differ across Noun frequency conditions and Sentence structure (*t*s < 1.31). Importantly, there was no interaction between Verb bias and N2 frequency ($F[1,30] = .00$). However, weak bias verb sentences were rated to be slightly less plausible than sentences with strong bias verbs, $t(1,31) = 1.82$, $p = 0.08$.

Procedure

The same procedure was used as in Experiment 1.

Scoring and analysis

Scoring was the same as in Experiment 1. Again, analyses were carried out using mixed logit models and linear mixed effects models. Fixed factors included Verb bias, N2 frequency, and Sentence structure. Additional analyses were carried out including the predictor Plausibility (according to the item norming) to assess the effects of the experimentally manipulated factors above and beyond the effect of the plausibility ratings. These analyses confirmed the pattern of results obtained from analyses using only experimentally manipulated factors and are therefore not reported.

4.3.2 Results

Error rates

Speakers made more errors in reproducing sentences with weak bias verbs (M = 0.36, SD = 0.48) than strong bias verbs (M = 0.31, SD = 0.46), but the difference was only marginally significant (β = -0.25, SE = 0.13, z = -1.90, p = 0.06). Error rates did not differ across Noun frequency conditions or Sentence structures (zs < 1.17).

Verb onsets

Onset analyses were carried out on trials where speakers produced sentences with the intended structure (n = 754), making use of linear mixed effects models. Verb bias, N2 frequency and Sentence structure were fixed predictors in these models.

In line with our predictions, we found a significant interaction between Verb bias and N2 frequency (β = -55.43, SE = 27.99, t = -1.98, p < .05). Speakers were faster to produce weak-verb bias sentences with a high frequency N2 than with a low frequency N2. In sentences with strong bias verbs, onsets were not predicted by N2 frequency (t = -0.80). To examine this effect more closely, we ran a second model with N2 frequency as a continuous predictor. Table 6 summarizes the fixed effects of the best model fit.

Table 6. Summary of fixed effects in the linear mixed effects model predicting verb onset latencies in Experiment 2. N = 754, log-likelihood = -5044.

Predictor	Coefficient	*SE*	*t*	*Pr(>|t|)*
(Intercept)	765.38	24.47	31.27	< 2e-16
Structure	-29.45	19.03	-1.55	0.13
Verb bias	15.46	15.93	0.97	0.34
N2 frequency (continuous)	-6.01	12.88	-0.47	0.64
Verb bias by N2 frequency	70.72	23.82	2.97	<0.01

Note: by-subject random slopes are included for the interaction between Verb bias and N2 frequency

In this model, Verb bias showed a (now stronger) interaction with N2 frequency (p < .01. Figure 2 shows the interaction between Verb bias and N2 frequency.

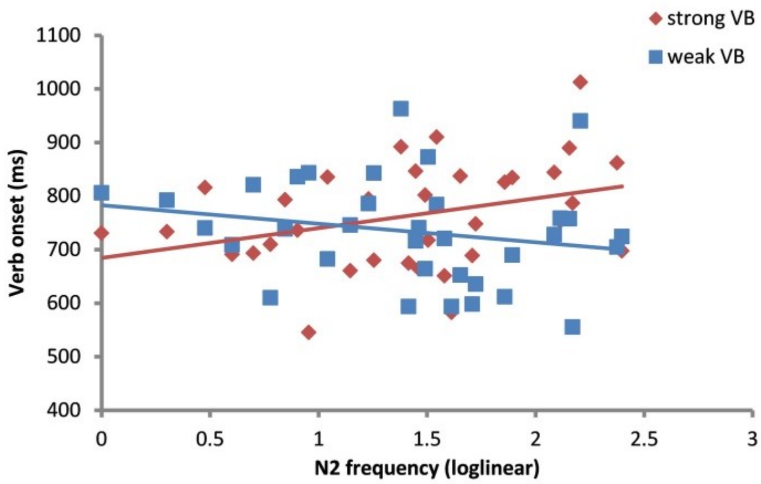

Figure 2. Scatterplot of the relationship between N2 frequency and Verb bias, collapsing across sentence type.

Speakers were faster to initiate weak-verb bias sentences featuring nouns with a high frequency N2 than a low frequency N2 (β = -39.63, *SE* = 18.04, *t* = -2.20). In contrast, N2 frequency did not produce an effect in sentences with strong bias verbs (*t* =1.53).

Fast and slow speakers

The fact that by-subjects random slopes for the interaction between Verb bias and N2 frequency improved model fit in a model predicting verb onsets (see above) already suggests that there was substantial subject-level variability in the strength of this interaction. One possible source for these individual differences is production speed (Wagner, et al., 2010). Therefore, we measured average production latencies on the filler trials per participant as an index of production speed. After excluding incorrect responses and outliers as on the target trials (see above), we added this factor to a model predicting verb onsets from Verb bias, N2 frequency (continuous), and Sentence structure. The final model included a significant two-way interaction between Verb bias and Production speed (β = -0.29, *SE* = 0.11, *t* = -2.61). Figure 3 shows the interaction between Verb bias and Production speed.

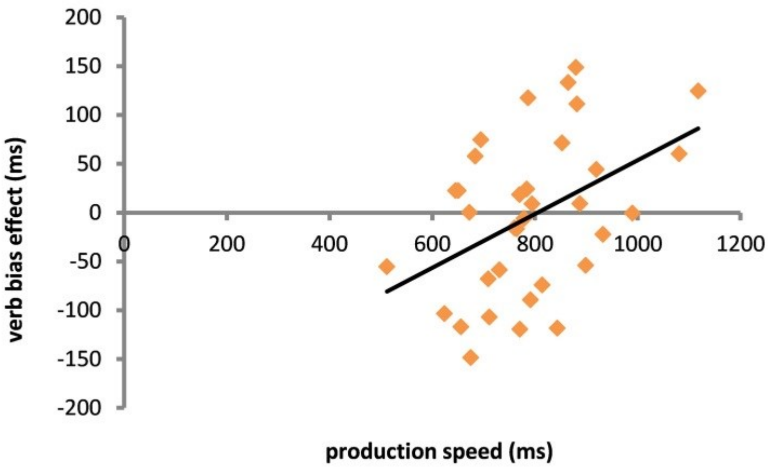

Figure 3. Scatterplot of the relationship between production latency and the magnitude of the verb bias effect (weak bias RT - strong bias RT). For this plot, outliers (n = 2) based on extreme values of Cook's distance were excluded.

A longer average speech onset latency on the filler items was associated with a larger verb bias effect on experimental items, such that sentences with weak bias verbs were produced more slowly than sentences with strong bias verbs. High production speed on filler trials was associated with an effect of verb bias in the opposite direction; weak verb bias sentences were initiated slightly faster than strong verb bias sentences.

There was a trend towards a three-way interaction between Verb bias, N2 frequency and Production Speed ($\beta = 0.27$, $SE = 0.16$, $t = 1.65$, $p = 0.10$); in the weak bias condition, slower production speed led to a larger delay in producing utterances with low frequency N2 compared to a high frequency N2. In utterances with strong bias verbs there were no differences in onset times for low and high frequency N2s, nor was there a relationship with production speed.

4.3.3 Discussion

The results of Experiment 2 provide evidence for extensive planning at the lexical level in sentences with syntactic flexibility. The significant interaction between N2 frequency and Verb bias suggests that primarily in sentences with weak bias verbs, lexical planning scope included N2. This effect was especially apparent in slower speakers who showed delayed onsets for weak-verb bias sentences with low frequency N2s relative to weak-verb bias sentences with high frequency N2s.

4.4 General discussion

The aim of this study was to examine the influence of the availability of multiple syntactic frames (i.e., syntactic flexibility), and hence multiple placement options of post-verbal material, on the lexical planning scope during recall of dative verb phrases. Syntactic flexibility was varied by using verbs with different degrees of bias towards each dative alternative. In addition, lexical accessibility was varied by using post-verbal nouns with low and high frequency. Assessing the effect of increased lexical accessibility on speech onsets allowed us to make inferences about the scope of lexical planning during sentence recall. In Experiment 1, the frequency of the first post-verbal noun was varied (i.e., the direct object in PO and the indirect object in DO datives). Results only provided suggestive evidence that planning scope was wider for weak than strong bias verbs.

In Experiment 2, we manipulated the frequency of the second post-verbal noun (i.e., the indirect object in PO and the direct object in DO datives). Results provided evidence for lexical planning up to and including the second post-verbal noun in syntactically flexible sentences. A significant interaction was found between verb bias and N2 frequency; N2 frequency only mattered during the recall of weak bias sentences. Weak bias sentences with a high frequency N2 were initiated faster than sentences with a low frequency N2. There was substantial inter-subject variability and the interaction between verb bias and N2 frequency was primarily driven by the speakers with longer average speech onset latencies on filler trials–the slow speakers. Previous research has shown that slow speakers (i.e. speakers who initiate their utterances with relatively long latencies) engage in more extensive advance planning than fast speakers (Wagner, et al., 2010; Wheeldon, et al., 2013). The present results suggest that slow speakers may be more flexible in extending their planning scope than fast speakers. Wagner et al. (2010) found that speakers decrease their scope of planning under cognitive load induced by a secondary conceptual decision task. Hence, extensive advance planning is cognitively more demanding than piecemeal, incremental planning. Slow speakers might thus have more cognitive capacity (e.g., cognitive control), allowing them to engage in flexible advance planning than faster speakers. Further experimentation is needed to investigate this possibility.

There are two possible mechanisms through which syntactic flexibility could influence the activation of lexical material and planning scope. According to a first hypothesis, syntactic flexibility induced by weak verb bias may give rise to the activation of both post-verbal nouns in the dative verb phrase: When retrieving a weak bias verb, two dative structural alternatives featuring different object orders become activated to roughly the same degree. Consequently, both post-verbal nouns (i.e., the direct and the indirect object) may become activated, as they could both fill the position immediately after the verb. The insertion of the direct object would lead to the production of a PO dative, while the insertion of the indirect object would lead to the production of a DO dative. This hypothesis is in line with a single-stage view of grammatical encoding, suggesting that upon verb retrieval, the processor automatically tries to select the constituent immediately adjacent to the verb in the linear structure (Cai, et al., 2012; Pickering, et al., 2002).

On a different account, syntactic flexibility may lead to competition between abstract dative sentence frames (Ferreira, 1996; Hwang & Kaiser, 2013; Myachykov, et al., 2013). During the time needed to resolve competition and select a target frame, post-verbal objects may be retrieved in parallel. In the absence of competition (i.e., in sentences with strong bias verbs), the utterance may be initiated immediately after retrieval of the verb (and possibly the first noun). This account is in line with a multiple-stage view of grammatical encoding, in which an abstract (i.e., lexically independent) hierarchical structure intermediates the mapping from functional-level input to a linear structure (e.g., Bock & Levelt, 1994).

Our data do not distinguish between these two accounts. Both predict that under syntactic flexibility N2 should be activated, be it immediately after verb retrieval as a candidate to fill the post-verb position, or during the resolution of competition between the two dative candidate frames. Importantly though, both accounts imply that upon the retrieval of a weak-bias verb, syntactic flexibility offers a speaker the choice to insert an indirect or direct object into the developing verb phrase/ to construct a PO or a DO dative frame.

The process of choosing between the insertion of the direct or indirect object in the post-verbal slot may influence onset latencies for sentences with weak bias verbs in two different ways. On the one hand, frequency differences between to-be-inserted nouns may support quick settling on the ‘winning’ syntactic alternative. On this *relative* frequency view, noun frequency differences may help the speaker in choosing a structure, by promoting the insertion of the higher frequency noun into the sentence structure first (cf. Stallings & MacDonald, 2011; Stallings, et al., 1998). Consequently, if the direct object NP has higher frequency than the indirect object NP, the selection of a PO structure will be promoted, while an indirect object NP with higher frequency promotes the selection of a DO structure. Note that in both experiments, noun frequency differences were always congruent with the presented sentence structure and the preference of its verb, i.e., N1 always had higher frequency than N2. However, differences still existed in the *degree* of consistency between the presented sentence structure, its verb bias and noun frequency differences. That is, a noun ordering in which N1 has higher frequency than N2 (i.e., HF-LF ordering) is *more* consistent than an ordering with a high frequency N1 and a high frequency N2 (HF-HF, or in Experiment 1: LF-LF). Therefore, a *relative* frequency view predicts quick settling on a structure with a weak bias verb (and therefore shorter onset latencies) for a) the high frequency N1 condition in Experiment 1 (i.e., HF-LF ordering), and b) the low frequency N2 condition in Experiment 2 (i.e., HF-LF ordering). Although our results did not show any effect of N1 frequency, the opposite pattern was observed for N2 frequency. Noun frequency differences did not facilitate structure choice. Instead, speakers were *slower* to initiate weak-bias sentences with a HF – LF ordering than with a HF – HF ordering.

Therefore it seems that the *absolute* frequency of the nouns within speakers' scope of planning matters for the onset latencies. On this *absolute* frequency view, structure choices are made rather independently from the lexical frequency of the nouns and effects of frequency on onset latencies only reflect the ease of retrieving the nouns that are within the scope of planning. This view thus predicts that sentences with weak bias verbs can only be initiated when the to-be-produced noun with the lowest frequency is prepared. Findings from both experiments support this view by showing that onset latencies were longer for weak-verb bias sentences with low frequency than high frequency N2s (and high frequency N1s, i.e., HF - LF ordering).

Although syntactic flexibility affected verb onset latencies, it did not affect the speakers' production choices. In Experiment 1 and 2 together, only eight occasions of flipping (i.e., producing the alternative structure to the presented one) were found. There are two reasons why syntactic flexibility did not affect production choice. Firstly, verb biases and differences in the frequency of the first and second post-verbal noun were always congruent with the sentence structure presented to the participant. For instance, the verb *voorschotelen* 'dish out' which has a strong bias towards the DO object dative, was only presented in a DO frame and with an indirect object (e.g., *king*) that was more frequent than the direct object (e.g., *banquet*). These circumstances all promoted the usage of the structure as it was presented to the participant. Only in sentences with weak-bias verbs and nouns with equal frequency (e.g., low frequency N1 with a low frequency N2 in Experiment 1 and high frequency N1 with a high frequency N2 in Experiment 2), conditions supported flipping of sentence structure–although speakers might have been primed to re-use the presented structure through the paradigm that we used. Indeed, flipping of sentence structure primarily occurred in the weak-verb bias condition (6 out of 8 occurrences) and when the frequency of

N1 and N2 was (almost) equal (6 out of 8 occurrences). Secondly, as many authors have pointed out, the RSVP paradigm promotes re-usage of the same structure and lexical material (Chang, et al., 2003; Potter & Lombardi, 1998; Tooley & Bock, 2014). Structural priming effects have been observed in many sentence generation studies, and these effects tend to increase in strength when prime and target share lexical content (Cleland & Pickering, 2003; Pickering & Branigan, 1998). Thus, it is not surprising to observe them in the RSVP paradigm as well.

One might see the absence of structural differences between the target sentences and the sentences participants reproduced as suggesting that the participants did not reconstruct the sentences based on the conceptual structure but instead retrieved the fully specified string of words from working memory. However, in the post-experimental debriefing, participants reported that they often forgot the precise content and wording of the sentence due to the intervening distractor task and had to reconstruct the wording from their memory of the underlying sentence meaning. The nature of their errors provides converging evidence for this observation. In both experiments, the majority of the errors (56 and 70 % of the errors in Experiment 1 and 2, respectively) were substitutions of the target verb and/or a noun with often conceptually similar words, e.g., the use of 'entrepreneur' instead of 'buyer'. Most importantly, it is difficult to see how the observed interaction between Verb bias and Noun frequency could arise if the linearized sequence of words were retrieved from working memory: Since verbs and nouns in the target sentences were carefully matched on characteristics influencing their accessibility (e.g., frequency of the verb and non-manipulated object noun, plausibility, length and syllable count), no systematic effects on onset latencies would be expected by a strictly episodic view. Instead, the interaction suggests that verb retrieval during recall involves the activation of associated subcategorization frames to the degree specified by the verb's bias, which in turn influences the degree to which upcoming lexical material is (re-)activated at verb onset (cf. Melinger & Dobel, 2005).

In sum, results of Experiment 1 and 2 suggest that syntactic flexibility expands planning scope by promoting the early activation of lexical material during sentence recall. This is in line with findings from Myachykov et al. (2013), who found that speakers of a less flexible language (English) showed more strictly incremental sentence planning than speakers of a syntactically flexible language (Russian). The current study extends these findings by manipulating syntactic flexibility within one language (Dutch), using a different sentence structure (datives) and a different paradigm.

Acknowledgements

We would like to thank Agnieszka Konopka and Gary Dell for helpful discussions and Annelies van Wijngaarden and student assistants from the Psychology of Language Department for help with data collection and processing.

Appendix A

verb pair (strong/weak)	**verb pair English**	**sentence frame Dutch** (*with high vs. low frequency N1*)	**sentence frame English** (*with high vs. low frequency N1*)
PO biased			
versturen/zenden	send	De aanklager verstuurt *een bericht/telegram* aan de raadsman	The client sends *a message/telegram* to the lawyer
overmaken/uitbetalen	transfer/pay out	De boer maakt *het geld/de som* over aan de smid	The farmer transfers *the money/sum* to the smith
schenken/cadeau doen	donate	De directeur schenkt *het horloge/de vulpen* aan de griffier	The director donates *the watch/pen* to the clerk
bekend maken/verklappen	reveal	De goochelaar maakt *het geheim/de truc* bekend aan de concurrent	The magician reveals *the secret/trick* to the competitor
slijten/verkwanselen	sell/bargain away	De juwelier slijt *de steen/het goud* aan de knul	The jeweler sells *the stone/gold* to the boy
uitdelen/voeren	distribute/feed	De kinderen delen *het eten/de appels* uit aan de geiten	The children distribute *the food/apples* to the goats
afstaan/nalaten	relinquish/bequeath	De oma staat *de foto's/sieraden* af aan de verwanten	Grandma relinquishes *the photos/jewelry* to the relatives
terugbetalen/noemen	pay back/mention	De ondernemer betaalt *het bedrag/de lening* terug aan de financier	The merchant pays back *the amount/loan* to the financer
melden/verkondigen	report/proclaim	De president meldt *de theorie/het verzinsel* aan de adviseur	The president reports *the theory/fabrication* to the advisor
lenen/opdragen	lend/dedicate	De schrijver leent *het boek/debuut* aan de fan	The author lends *the book/debut* to the fan
doorspelen/doorschuiven	pass	De spelleider speelt *de kaarten/fiches* door aan de amateurs	The leader passes *the cards/chips* to amateurs
voorleggen/voordragen	submit/recite	De student legt *het verhaal/pleidooi* voor aan de docent	The student submits *the story/argument* to the lecturer
betalen/uitbetalen	pay/pay out	De verzekeraar betaalt *de schade/onkosten* aan de huurder	The insurer pays *the damages/expenses* to the tenant
verlenen/bieden	give/offer	De zakenman verleent *hulp/bijstand* aan het oudje	The businessman gives *help/assistance* to the oldie
DO biased			
wijzen/laten zien	point/showing	De agent wijst *de chauffeur/bestuurder* het sein	The officer points *the driver/operator* the signal
adviseren/aanbevelen	advise/recommend	De arts adviseert *de patient/bejaarde* het medicijn	The doctor advises *the patient/senior* the ointment
garanderen/bevestigen	guarantee/confirm	De bankier garandeert *de meneer/cliente* het rendement	The dealer guarantees *the sir/client* the returns
verbieden/weigeren	prohibit/refuse	De barman verbiedt *het meisje/de puber* de likeur	The bartender prohibits *the girl/adolescent* liquor
opleveren/verschaffen	provide	De loterij levert *de leden/spelers* tonnen op	The lottery provides *the members/players* thousands
beletten/ontzeggen	prevent/deny	De douane belet *het gezelschap/de misdadiger* de doorgang	Customs prevents *the group/criminal* the passage
leren/voorlezen	teach/read out	De korpschef leert *de groep/het team* de instructie	The police chief teaches *the group/team* the instruction
verzoeken/vragen	request/ask	De man verzoekt *de vriend/kelner* de gunst	The man requests *the friend/waiter* the favor
meegeven/toegooien	give along/toss	De moeder geeft *het kind/ventje* de peer mee	The mother gives *the child/the little boy* the pear
voorschotelen/serveren	dish out/ serve	De ober schotelt *de koning/monarch* het feestmaal voor	The waiter dishes out *the king/monarch* the banquet
bezorgen/toewerpen	deliver/throw	De postbode bezorgt *het gezin/echtpaar* het pakket	The postman delivers *the family/couple* the package
ontnemen/benemen	deprives/take away	De stank ontneemt *de jongen/keizer* de eetlust	The smell deprives *the boy/emperor* the appetite
aanraden/aanprijzen	advise/recommend	De verkoper raadt *de jongeman/afnemer* de stofzuiger aan	The seller advises *the young man/customer* the vacuum cleaner
geven/aanreiken	give/hand	De verzorger geeft *de hond/aap* het voer	The guard gives *the dog/monkey* the food

Appendix B

verb pair (strong/weak)	verb pair English	sentence frame Dutch (*with high vs. low frequency N2*)	sentence frame English (*with high vs. low frequency N2*)
PO biased			
betalen/uitbetalen	pay/pay out	De bewoner betaalt de schade aan *de eigenaar/verhuurder*	The resident pays the damages to *the owner/landlord*
versturen/zenden	send	De cliënt verstuurt een bericht aan *de advocaat/raadsman*	The client sends a message to *the lawyer/adviser*
overmaken/uitbetalen	transfer/pay out	De dader maakt het bedrag over aan *het slachtoffer/de gedupeerde*	The offender transfers the amount to *the victim/duped*
schenken/cadeau doen	donate	De directeur schenkt het horloge aan *de werknemer/diplomaat*	The director donates the watch to *the employee/diplomat*
afleggen/zweren	swear	De getuige legt de eed af aan *de jurist/rechtszaal*	The witness swears the oath to *the jurist/courtroom*
bekend maken/verklappen	reveal	De goochelaar maakt de truc bekend aan *de toeschouwer/toehoorder*	The magician reveals the trick to *the audience/hearer*
terugbetalen/noemen	pay back/mention	De handelaar betaalt de kosten terug aan *de directeur/aannemer*	The merchant pays back the costs to *the director/contractor*
slijten/verkwanselen	sell/bargain away	De juwelier slijt de ketting aan *de koper/diva*	The jeweler sells the necklace to *the buyer/diva*
uitdelen/voeren	distribute/feed	De kinderen delen het brood uit aan *de honden/zwanen*	The children distribute the bread to *the dogs/swans*
doorspelen/doorschuiven	pass	De leider speelt de kaarten door aan *de spelers/gokkers*	The leader passes the cards to *the players/gamblers*
uitdelen/nalaten	distribute/bequeath	De opa deelt het geld uit aan *de familie/verwanten*	Grandpa distributes the money to *the family/relatives*
melden/verkondigen	report/proclaim	De president meldt de leugen aan *de natie/meute*	The president reports the lie to *the nation/mob*
lenen/opdragen	lend/dedicate	De schrijver leent het boek aan *het meisje/de minnares*	The author lends the book to *the girl/mistress*
voorleggen/voordragen	submit/recite	De student legt het gedicht voor aan *de leraar/docent*	The student submits the poem to *the teacher/lecturer*
overdragen/uitloven	transfer/offer	De winkelier draagt de prijs over aan *het personeel/de winnaar*	The retailer transfers the prize to *the staff/winner*
verlenen/bieden	give/offer	De zakenman verleent hulp aan *de vriend/het oudje*	The businessman gives assistance to *the friend/oldie*
DO biased			
wijzen/laten zien	point/show	De agent wijst de chauffeur *het bord/sein*	The officer points the driver *the sign/signal*
adviseren/aanbevelen	advise/recommend	De arts adviseert de patient *het middel/zalfje*	The doctor advises the patient *the remedy/ointment*
verbieden/weigeren	prohibit/refuse	De barman verbiedt de leerling *de alcohol/tequila*	The bartender prohibits the student *alcohol/tequila*
geven/aanreiken	give/hand	De bewaker geeftt de aap *het voer/de biet*	The guard gives the monkey *the food/beet*
ontfutselen/aftroggelen	pilfer	De dief ontfutselt de soldaat *het horloge/uurwerk*	The thief pilfers the soldier *the watch/clock*
beletten/ontzeggen	prevent/deny	De douane belet de misdadiger *de doorgang/doortocht*	Customs prevents the criminal *the passage/transit*
garanderen/bevestigen	guarantee/confirm	De handelaar garandeert de cliënt *het voordeel/rendement*	The dealer guarantees the client *the advantage/returns*
leren/voorlezen	teach/read out	De lerares leert de groep het *verhaal/relaas*	The teacher teaches the children *the story/tale*
besparen/sparen	save	De lift bespaart de jongen *moeite/het gedoe*	The elevator saves the boy *the effort/trouble*
opleveren/verschaffen	provide	De loterij levert de deelnemers *prijzen/tonnen* op	The lottery provides the participants *prizes/thousands*

voorschotelen/serveren	dish out/ serve	De ober schotelt de klant *de maaltijd/het feestmaal* voor	The waiter dishes out the customer *the meal/banquet*
bezorgen/toewerpen	deliver/throw	De postbode bezorgt het gezin *de krant/het pakket*	The postman delivers the family *the newspaper/package*
meegeven/toegooien	give along/toss	De vader geeft het kind *het boek/de peer* mee	The father gives the child *the book/pear*
aanraden/aanprijzen	advise/recommend	De verkoper raadt de klanten *het apparaat/de stofzuiger* aan	The seller advises the customers *the apparatus/vacuum cleaner*
verzoeken/vragen	request/ask	De voorzitter verzoekt de minister om *een verklaring/opheldering*	The Chairman requests the Minister *a statement/clearing*
ontnemen/benemen	take away	Het hek ontneemt de jongeman *het uitzicht/de uitkijk*	The fence takes away the young man *the view/lookout*

References

Acheson, D. J., & MacDonald, M. C. (2009a). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language, 60*(3), 329-350.

Acheson, D. J., & MacDonald, M. C. (2009b). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*(1), 50.

Baayen, H., Piepenbrock, R., & Gulikers, L. (1995). The CELEX lexical database (release 2 ed.): Linguistic Data Consortium.

Baayen, R. H. (2008a). *Analyzing linguistic data* (Vol. 505): Cambridge University Press Cambridge, UK.

Baayen, R. H. (2008b). languageR: Data sets and functions with "Analyzing Linguistic Data: A practical introduction to statistics". *R package version 0.953*.

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

Bates, D., Maechler, M., & Dai, B. (2009). lme4: linear mixed-effects models using S4 classes. 2009. R Package, version 0.999375–31.

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

Bock, K. (1996). Language production: Methods and methodologies. *Psychonomic Bulletin & Review, 3*(4), 395-421.

Bock, K., & Brewer, W. F. (1974). Reconstructive recall in sentences with alternative surface-structures. *Journal of Experimental Psychology, 103*(5), 837-843.

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

Boersma, P., & Weenink, D. Praat: doing phonetics by computer [computer program]. Version 5.3. 39. 2013.

Brown, M., Savova, V., & Gibson, E. (2012). Syntax encodes information structure: Evidence from on-line reading comprehension. *Journal of Memory and Language, 66*(1), 194-209.

Cai, Z. G., Pickering, M. J., & Branigan, H. P. (2012). Mapping concepts to syntax: Evidence from structural priming in Mandarin Chinese. *Journal of Memory and Language, 66*(4), 833-849.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science, 26*(5), 609-651.

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

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

Cleland, A. A., & Pickering, M. J. (2003). The use of lexical and syntactic information in language production: Evidence from the priming of noun-phrase structure. *Journal of Memory and Language, 49*(2), 214-230.

Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences, 31*(5), 593-611.

Dell, G. S., Chang, F., & Griffin, Z. M. (1999). Connectionist models of language production: Lexical access and grammatical encoding. *Cognitive Science, 23*(4), 517-542.

Dowty, D. (2003). The dual analysis of adjuncts/complements in categorial grammar. In E. L. C. M. C.Fabricius-Hansen (Ed.), *Modifying Adjuncts* (pp. 33-66).

Ferreira, F., & Swets, B. (2002). How Incremental Is Language Production? Evidence from the Production of Utterances Requiring the Computation of Arithmetic Sums. *Journal of Memory and Language, 46*(1), 57-84.
Ferreira, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*(5), 724-755.
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.
Griffin, Z. M. (2001). Gaze durations during speech reflect word selection and phonological encoding. *Cognition, 82*(1), B1-B14.
Griffin, Z. M., & Weinstein-Tull, J. (2003). Conceptual structure modulates structural priming in the production of complex sentences. *Journal of Memory and Language, 49*(4), 537-555.
Hagoort, P., & Brown, C. M. (2000). ERP effects of listening to speech compared to reading: the P600/SPS to syntactic violations in spoken sentences and rapid serial visual presentation. *Neuropsychologia, 38*(11), 1531-1549.
Hwang, H., & Kaiser, E. (2013). Having a syntactic choice is not always better: the effects of syntactic flexibility on Korean production. *Language, Cognition and Neuroscience*(ahead-of-print), 1-17.
Indefrey, P., & Levelt, W. J. (2004). The spatial and temporal signatures of word production components. *Cognition, 92*(1-2), 101-144.
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.
Kempen, G. (2013). Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics*.
Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology, 58*(1), 68-101.
Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology, 73*, 1-40.
Kuchinsky, S. E., & Bock, J. K. (2010). *From seeing to saying: Perceiving, planning, producing.* Paper presented at the 23rd meeting of the CUNY human sentence processing conference.
Kuznetsova, A., Brockhoff, P., & Christensen, R. (2013). lmerTest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). *R package version*, 2.0-0.
Lombardi, L., & Potter, M. C. (1992). The regeneration of syntax in short term memory. *Journal of Memory and Language, 31*(6), 713-733.
Martin, R. C., Crowther, J. E., Knight, M., Tamborello, F. P., 2nd, & Yang, C. L. (2010). Planning in sentence production: evidence for the phrase as a default planning scope. *Cognition, 116*(2), 177-192.
Melinger, A., & Dobel, C. (2005). Lexically-driven syntactic priming. *Cognition, 98*(1), B11-20.
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.
Miozzo, M., & Caramazza, A. (2003). When more is less: A counterintuitive effect of distractor frequency in the picture-word interference paradigm. *Journal of Experimental Psychology: General, 132*(2), 228-252.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: the case of English and Russian. *Quarterly Journal of Experimental Psychology, 66*(8), 1601-1619.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*(4), 633-651.

Pickering, M. J., Branigan, H. P., & McLean, J. F. (2002). Constituent structure is formulated in one stage. *Journal of Memory and Language, 46*(3), 586-605.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*(6), 633-654.

Potter, M. C., & Lombardi, L. (1998). Syntactic priming in immediate recall of sentences. *Journal of Memory and Language, 38*(3), 265-282.

R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(6), 1503.

Smith, M., & Wheeldon, L. (1999). High level processing scope in spoken sentence production. *Cognition, 73*(3), 205-246.

Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language, 39*(3), 392-417.

Tooley, K. M., & Bock, K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition, 132*(2), 101-136.

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.

Verhoeven, J., De Pauw, G., & Kloots, H. (2004). Speech rate in a pluricentric language: A comparison between Dutch in Belgium and the Netherlands. *Language and Speech, 47*, 297-308.

Wagner, V., Jescheniak, J. D., & Schriefers, H. (2010). On the Flexibility of Grammatical Advance Planning During Sentence Production: Effects of Cognitive Load on Multiple Lexical Access. *Journal of Experimental Psychology-Learning Memory and Cognition, 36*(2), 423-440.

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

5

DATIVE ALTERNATION AND PLANNING SCOPE IN SPOKEN LANGUAGE

A corpus study on effects of Verb bias in VO and OV clauses of Dutch

Van de Velde, M., Kempen, G., & Harbusch K. (under revision). Dative alternation and planning scope in spoken language: A corpus study on effects of verb bias in VO and OV clauses of Dutch.

Abstract

The syntactic structure of main and subordinate clauses is determined to a considerable extent by verb biases. For example, some English and Dutch ditransitive verbs have a preference for the prepositional object dative, whereas others are typically used with the double object dative. In this study, we compare the effect of these biases on structure selection in (S)VO and (S)OV dative clauses in a corpus of spoken Dutch (CGN). This comparison allowed us to make inferences about the size of the advance planning scope during spontaneous speaking. If the verb is an obligatory component of clause-level advance planning scope, as is claimed by the hypothesis of *hierarchical incrementality*, then biases should exert their influence on structure choices, regardless of early (VO) or late (OV) positioning of the verb in the clause. Conversely, if planning proceeds in a piecemeal fashion, strictly guided by lexical availability, as claimed by *linear incrementality*, then the verb and its associated biases can only influence structure choices in VO sentences. We tested these predictions by analyzing structure choices in the Corpus of Spoken Dutch using mixed logit models. Our results support a combination of linear and hierarchical incrementality, showing a significant influence of verb bias on structure choices in VO, and a weaker (but significant) effect in OV clauses.

5.1 Introduction

During grammatical encoding, conceptual messages are transformed into linguistically structured word sequences. Most models of sentence production, starting with Garrett (1976), hold that grammatical encoding consists of two sub-processes: functional and positional encoding. Functional encoding comprises the retrieval of lemmas (i.e., lexical entries specifying the syntactic properties of a word) and the assignment of grammatical roles (e.g., subject, direct object, indirect object, head) to these lemmas. During positional encoding, lemmas are embedded in syntactic constituents and receive a serial order. There is considerable debate on what factors drive the ordering of constituents in linguistic structures. One group of theories assumes that the mapping from a conceptual to a linguistic structure is lexically mediated. According to these theories, lemma accessibility can directly influence word order so that highly accessible units are prioritized, i.e., receive early placement in the clause. For example, consider a speaker describing a *dog–chases–mailman* event. If, at the outset of sentence formulation, the concept with the patient role ("undergoer") is more accessible to the speaker than the actor, it will tend to be placed early in the clause, giving rise to a passive structure (e.g., *the mailman is being chased by the dog)* (Bock & Levelt, 1994; Kempen & Hoenkamp, 1987).

Other theories postulate a direct mapping from the conceptual message to a hierarchical sentence structure (Chang, 2002; Chang, Dell, & Bock, 2006). Given the conceptual structure of the message, in particular the thematic roles, functional-hierarchical structures are formed with terminal nodes that can be unified with lexical items. Linearization of the hierarchical structure and unification with lexical items takes place during the subsequent positional stage. Sentence production can still proceed incrementally but is driven by an overarching structural framework rather than by activation levels of individual lexical items.

In the literature, these two accounts of sentence production have been referred to as (a) *linearly* (or lexically/radically) incremental, and (b) *hierarchically* (or structurally/ weakly) incremental (Bock, Irwin, Davidson, & Levelt, 2003; Lee, Brown-Schmidt, & Watson, 2013). The accounts do not only differ in their ideas on how structure choices come about, they also make different assumptions about the scope of pre-planning needed to initiate an utterance. Linear incrementality posits that as little as one lexical item needs to be encoded before the formulation and pronunciation of a sentence can be initiated. In contrast, according to hierarchical incrementality, the retrieval of a sentence-initial increment encompasses the (partial) preparation of later increments and/or of an overarching abstract syntactic framework. More specifically, the number of lemmas retrieved prior to speech onset is constrained by the grammatical structure of the sentence (Lee, et al., 2013).

Within the hierarchical structure of *clauses*, the verb is often granted a central role. In particular, the subcategorization frame of the head verb of the clause is assumed to play an important role in organizing functional role assignment, leading to preparation of clause-size planning units centered around the verb lemma (Bock & Levelt, 1994). The ditransitive verb *give*, for instance, requires three arguments: the subject, and two objects (one direct, one indirect). In versions of Tree Adjoining Grammar (cf. Ferreira, 2000; Joshi, 1987; Kroch & Joshi, 1985), the lexical entries for verbs have the form of “elementary trees” with one branch whose terminal node is bound to (unified with) the verb itself, and additional branches that are to be bound to constituents specifying arguments and adjuncts of the verb. In accounts of syntactic production based on a TAG-like formalism, this implies that binding the non-head constituents of a clause needs to wait until after the head verb of the clause has been selected from the lexicon. This holds even if some of these constituents, in particular the subject, precede the verb in the clause that is finally uttered.

The aim of the present study is to test this claim directly. We examine how strongly the preferred subcategorization frame of the head verb can influence the syntactic shape of the clause depending on its position in the clause: early (VO) or late (OV). Many verbs display a bias for one subcategorization frame (i.e., elementary syntactic tree) over another. With respect to the so-called dative alternation, for instance, some verbs prefer the double object dative (DO; e.g., *Anna gives David the apple*) while others prefer the prepositional object dative (PO; e.g., *Anna gives the apple to David*). According to some models of sentence production, verb biases exert their influence via weighted links between verb lemmas and structural representations (Chang, 2002; Chang, Janciauskas, & Fitz, 2012). Hence, a verb that is typically used with (i.e., has a bias towards) the DO dative—e.g., Dutch *opleveren* 'yield'—will have a more highly weighted connection to the DO than to the PO structure. Several studies have found effects of verb bias on structure choices or word-order choices in sentence production (Colleman, 2009; Ferreira, 1994; Stallings, MacDonald, & O'Seaghdha, 1998) and in sentence comprehension (Gahl & Garnsey, 2004; Garnsey, Pearlmutter, Myers, & Lotocky, 1997; Wilson & Garnsey, 2009).[1]

[1] Verb bias effects have been found in sentence *comprehension* studies as well. These studies have mainly focused on the parsing of sentences that contain verbs with a bias toward a direct-object NP (e.g., *we confirmed the date of our visit*) as opposed to a sentential complement bias (e.g., *we confirmed the date was correct*).

Verb preferences thus play a role in syntactic decisions. However, in all of the cited production studies the verb preceded the object argument(s) (i.e., VO). This raises the question of whether verb biases are also able to constrain syntactic choices if, in the resulting clause, they occupy a late position in that clause (i.e., OV). In the current study, we examine the influence of verb bias on syntactic choices in subordinate clauses of Dutch, that have obligatory OV order.

In Dutch subordinate clauses, the verb is placed after the canonical position of direct and indirect object. For example, in the embedded DO dative in (1), the verb *gaf* 'gave' is produced at the end of the relative clause.

(1)	de	jongen,	die	de	apen	de	bananen	gaf,	rende	weg
	the	boy	who	the	monkeys	the	bananas	gave	ran	away

'the boy who gave the monkeys the bananas ran away'

Geven 'give' has a strong bias towards the DO dative. If the verb lemma is retrieved before sentence formulation starts (as is assumed by verb-centered hierarchical models), then the bias of the verb may promote the selection of a DO structure in the subordinate clause. However, if the objects in an OV dative clause are planned before the verb is retrieved (as is assumed by strictly linear incrementality), verb bias cannot exert any influence on structure choices in OV clauses.

Before turning to our empirical study, we review the available evidence on how far ahead verbs are planned, i.e., their “pre-planning scope”.

5.1.1 The pre-planning scope for clauses

Early studies have focused on advance planning of simple subject-verb-object clauses. In a study by Lindsley (1975), participants described pictures displaying transitive events (e.g., a girl greeting a boy) with referents and actions that were either “old”, (i.e. known to the participant before trial onset) or “new”. Response latencies were measured as the participants orally produced subject-only (*the girl*), subject-verb (*the girl is greeting*), verb-only (*greeting*) and subject-verb-object utterances (*the girl is greeting the boy)*. Onset latencies were then compared among these utterance types to test the predictions of three models with diverse assumptions regarding the amount of preparation of the verb at sentence onset. (All models tacitly assumed that the subject is retrieved prior to sentence initiation, and that verb selection takes longer than subject selection.) According to the Pre-Predicate model, speakers initiate their utterance immediately after subject selection. Hence, characteristics of the verb do not contribute to onset latency at all. This also implies that verb-only utterances should always take longer to initiate than subject-first sentences. The Post-Predicate model posits that speakers retrieve both the subject and the verb before initiating their utterance. Subject-only utterances should thus be initiated faster than subject-verb utterances. Verb-only utterances should also have shorter onset latencies than subject-verb utterances, except when the subject is known to the speaker, in which case latencies should be similar because they are determined solely by the time it takes to complete verb selection. The third, Semi-Predicate model assumes that speakers initiate their utterance after selecting the subject and completing some (pre-lexical, i.e., visual and/or conceptual) preparation of the verb. Similarly to the Post-Predicate model, it predicts faster onsets for subject-only than subject-verb utterances, but in contrast to the Post-Predicate model it makes a different prediction for subject-verb utterances with known subjects versus verb-only utterances. Regarding this contrast, the model predicts longer latencies for verb-only utterances because the verb here needs to be completely prepared, against only partial preparation in a subject-verb utterance.

Lindsley found that subject-only utterances were initiated faster than subject-verb clauses, and that subject-verb clauses with old subjects were initiated faster than verb-only utterances. These results were taken as evidence for the Semi-Predicate model in which the verb is at least partly prepared (i.e., pre-lexically) before sentence onset. Kempen and Huijbers (1983) replicated and extended these findings for Dutch.

Vigliocco and Nicol (1998) found evidence for planning of the verb in a pre-linear stage by examining subject-verb agreement errors in English declarative (SVO) and interrogative (VSO) sentences. In their study, participants had to complete sentence beginnings, such as *The helicopter of the flights*. Since the head noun of the subject NP (*helicopter*) and the local noun (*flights*) differ in number, there is an increased probability of subject-verb agreement errors (specifically, so-called *attraction errors*) in which the verb agrees with the local noun rather than with the head noun of the sentence (leading to the production of: ... *are safe* instead of the correct ... *is safe*). However, in questions (e.g., *Is the helicopter for the flights safe?*) there is no linear proximity between the verb and the local noun. Still, participants produced similar amounts of agreement errors during the formulation of interrogative and declarative clauses. The authors interpreted this finding as evidence for computation of agreement during a stage in which syntactic structure is built prior to a linearization (positional) stage.

The above studies suggest that the verb is (at least partly) planned before overt production of a sentence is initiated. However, there is also evidence against pre-planning for the verb. Schriefers, Teruel and Meinshausen (1998) used a semantic interference paradigm to examine whether the verb is an obligatory part of the grammatical advance planning scope. In their study, participants had to describe pictures using either VO or OV clauses. In German, the language targeted here, subordinate clauses have OV word order, as in Dutch. The to-be-produced word order was induced by lead-in fragments. For instance, the fragment *auf dem nächsten Bild sieht man wie* 'on the next picture one sees how' needs a completion in the form of an OV clause, whereas the fragment *auf dem nächsten Bild* 'on the next picture' elicits a VO completion. The verbs to be used in the picture descriptions included transitive and intransitive verbs. After each lead-in fragment, participants were auditorily and visually presented with a distractor verb. Distractor verbs were either semantically related, semantically unrelated, syntactically deviant (i.e., had a different subcategorization frame), or identical to the target verb. A neutral condition (without distractor) was also included. Results indicated that, for transitive sentences, semantically related and syntactically deviant distractor verbs lead to interference only when the verb occurs in sentence-initial position. For sentences with intransitive verbs, no interference effects were obtained. The authors conclude that the verb lemma and its associated subcategorization frame is not necessarily planned ahead of overt sentence initiation.

5.2 The present study

In the present study, we investigate whether speakers plan the head verb of a clause before utterance initiation even if the verb is clause-final (OV). To address this question, we compare the effect of verb bias (DO versus PO) on structure choices in VO and OV dative clauses within *spoken language*. Verb biases represent the relative strengths of the connections between verbs and their associated subcategorization frames. In order to assess these biases, we analyze a syntactically annotated corpus of spoken Dutch (CGN; see Section 5.2.2.1).

Our main analysis is based on two important presuppositions. Firstly, we assume, in line with results from priming studies (Cleland & Pickering, 2006; Hartsuiker & Westenberg, 2000), that syntactic representations are shared between the written and spoken modalities. To check this assumption, we compare verb biases in written text (the Alpino and CONDIV corpora; see beginning of Section 5.3) with those obtained from a spoken corpus (CGN).

Secondly, we presuppose that the direction (DO or PO) and the strength (lemma-to-structure weight) of verb biases are represented in the mental lexicon. By implication, verb biases must be invariant with respect to the linear position the verb takes in the clause of which it is the head. In other words, the bias of a verb is based on the same underlying representation, irrespective of whether it ends up in VO or OV position. The most veridical measure of a verb's bias is obtained from *written* text, considering that writing situations, which usually allow extensive editing and revision, are relatively immune to time pressure, distraction, cognitive load and other factors that may obscure the underlying bias (Akinnaso, 1982). Consequently, if during *spoken* language production a verb shows different biases in different positions in the clause, this must be due to such factors. In order to establish the invariance of verb biases with respect to within-clause position, we therefore compare verb biases in VO and OV position in *written* language.

5.2.1 Hypotheses

Theories of linear and hierarchical incrementality yield opposite predictions regarding the effect of verb biases in OV clauses in *spoken language*. Linear incrementality assumes piecemeal planning guided by lexical availability. According to this view, conceptual information may cause activation of lemmas that are going to play a functional role in a clause, prior to activation of the head verb of that clause. This provides conceptually easily accessible lemmas the opportunity to be inserted into a generic clausal structure—i.e., a structure not yet shaped by a head verb—before the head verb has been inserted. In datives, the choice between the DO vs. PO alternatives must take place prior to placement of the first object (direct or indirect) noun phrase. Therefore, the verb can only influence the structure choice in VO structures, i.e., before the placement of the first object. This means that if sentences are planned in linearly incremental fashion, we do not expect to find an influence of verb preferences on structure choices in OV clauses. In other words, there should be an interaction between verb bias and verb position, such that verb bias significantly predicts structure choices in VO clauses, whereas the degree to which verb bias predicts structure choices in OV clauses should not exceed chance level.

On the other hand, hierarchical incrementality predicts that conceptual information is mapped directly onto a clausal structure unified with, and shaped by, the head verb. Thematic roles or event roles (i.e., agent, theme, recipient) are directly assigned a functional role (i.e., subject, direct object, indirect object) and a syntactic shape (NP, PP, AP), under control by the head verb. Since structure choices are made before the sentence is linearized, the position of the verb in the to-be-produced clause should not be a factor determining whether or not its bias can exert an effect on structure choices. Hence, if sentence planning proceeds in hierarchically incremental fashion, and the verb is a necessary part of the clause-initial planning scope, then the direction and strength of verb preferences should be a strong predictor of structure choices in OV clauses. In statistical terms, on a strictly hierarchically incremental account there should be a main effect of verb bias but no interaction between verb bias and verb position.

Finally, there is the possibility that linear and hierarchical incrementality each account for *part* of the sentences produced by language users. In particular, while hierarchical incrementality might hold when the head verb of a clause is easily accessible, when the head verb is hard to access, more easily accessible nonverbal constituents may already be inserted into the generic clausal structure earlier than the verb. Such a course of events may serve to prevent speech pauses and to promote fluency if verb-final word order is mandatory in the clause under construction. However, it may also give rise to mismatches between the linear order of direct and indirect objects on the one hand, and the DO/PO bias of the finally chosen head verb on the other. A combination of hierarchical and linear incrementality predicts that verb biases are capable of exerting their influence on structure choices in OV clauses but do so to a lesser extent than in VO clauses. This may result in an interaction between verb bias and verb position, such that verb bias is a strong predictor of structure choices in VO clauses and a weaker predictor in OV clauses.

5.2.2 Materials and methodology

To assess verb bias effects on structure choices in VO and OV dative clauses, we use corpus data and data obtained from a written sentence completion task (see Section 5.2.2.2). We introduced this test to obtain more reliable estimates of DO/PO verb biases in VO clauses—more reliable than the estimates obtained from the relatively small number of VO clauses in the CGN corpus (which comprise only 32% of the spoken clauses). We judged a written (instead of a less convenient spoken) task would be adequate, given that a recent study (Van Bergen, Van Lier, & De Swart, 2014) yielded written VO biases that were generalizable to spontaneous speech.

In the present section, we first describe how data were extracted from the corpora we investigated (5.2.2.1). This is followed by a description of how the sentence completion task was set up and analyzed (5.2.2.2). Lastly, we explain how we applied distinctive collexeme analysis in order to compute verb biases from the corpus and sentence completion data (5.2.2.3).

5.2.2.1 Corpora

The present study is based on two syntactically annotated corpora ("treebanks"): the syntactically annotated part of the Corpus of Spoken Dutch (CGN 2.0; 130595 corpus graphs and about 1.1 million words; Van Eerten, 2007), and the Alpino Treebank (the cdbl newspaper part of the Eindhoven Corpus; 7153 sentences and about 126 000 words; Uit den Boogaard, 1975; Van der Beek, Bouma, Malouf, & Van Noord, 2002). In both corpora, the sentences are annotated with dependency graphs—annotations that are relatively theory-neutral (Hoekstra, Moortgat, Schuurman, & Van Der Wouden, 2001; Van der Beek, et al., 2002). Dependency graphs specify functional-dependency relations between constituents and subconstituents. Additionally, the graphs contain specifications of lexical entry, string position, and syntactic category of every (sub)constituent. With these data at hand, we collected occurrences of the Dutch ditransitive construction using a list of ditransitive verbs compiled by Colleman (2009), and by querying additional ditransitive verbs occurring in dative sentence constructions.[2]

The queries yielded 1042 and 261 dative clauses for CGN and Alpino, respectively. These results were manually filtered and analyzed. From both data sets, we removed passive clauses which did not contain a direct object. We also excluded idiomatic expressions where the verb in isolation is not ditransitive but only takes a dative in combination with another PP or NP (e.g., *de hand schudden* 'shake hands', *het leven redden* 'save the life'), and cases that did not include the combination of a direct and an indirect object, or had been misclassified as ditransitive. These criteria led to the exclusion of 16 (2%) and 52 (20%) sentences from the CGN and Alpino output, respectively. The higher exclusion rate for the Alpino corpus was mainly caused by the higher passivization tendency in written language use (e.g., O'Donnel, 1974).

---

[2] CGN was searched with TIGERSearch (König & Lezius, 2003), applying special measures to recognize compound verbs with separable verb prefixes. Alpino was searched with a JAVA program of our own making, also able to retrieve separable verbs. Notice that, in Dutch, the separable part of a verb (the 'particle') always follows the direct and indirect object NPs—even in main clauses, where the finite form of the verb occupies "verb-second" position (SVO), e.g., *De vader geeft het kind de peer mee* 'The father gives the child the pear along' In PO structures, the indirect object (like any other type of PP) may follow the particle, e.g., *De vader geeft de peer mee aan het kind* 'The father gives the pear along to the child'.

For each of the clauses in the final data set (i.e., 1026 and 209 clauses in CGN and Alpino, respectively), we determined the type of construction they embody (DO or PO), their citation form (infinitive), and their verb position (VO or OV). The classification of verb infinitives yielded a total of 162 and 83 different ditransitive verbs in CGN and Alpino, respectively. Sentences were classified as PO datives if they consisted of a noun phrase (NP) and a prepositional phrase (PP, with *aan* ‘to’ or *voor* ‘for’ as head). Clauses were classified as DO datives when their objects featured two noun phrases. Both CGN and Alpino contain more DO than PO datives: In the final data set, 77% (789 CGN clauses) and 68% (142 Alpino clauses) featured a DO dative. Table 1 displays the number and percentage of sentences in each structural configuration for both corpora.

Additionally, we classified the position of the verb as preceding or following the canonical position of the (in)direct objects (i.e., VO vs. OV). In Dutch, verbs are placed immediately following the first constituent (“verb-second”, i.e. before any objects) when they are finite and occur in a declarative main clause. In most other cases, verbs are placed verb-finally, that is, after the (in)direct object NPs (Haeseryn, Romijn, Geerts, Rooij, & Toorn, 1997; Koster, 1975; Zwart, 1993). In both corpora, OV order occurs more frequently than VO order, with 68% (CGN) and 75% (Alpino) of the total number of sentences displaying OV order (see Table 1).

Table 1. Number and percentage of sentences displaying each configuration (DO vs. PO and VO vs. OV order) in Alpino and CGN.

	Alpino		CGN	
Structure	VO order	OV order	VO order	OV order
DO dative	37 (26%)	105 (74%)	257 (33%)	532 (67%)
PO dative	15 (22 %)	52 (78%)	67 (28%)	170 (72%)

In PO dative clauses with OV order, the verb can also be placed immediately after the direct object, preceding the prepositional phrase (*aan/voor* + the indirect object). For example, in the CGN clause *dat u kredieten moet geven aan de glastelers* 'that you must give loans to the greenhouse-horticulturists'*,* the verb *geven* is placed before the PP *aan de glastelers.* When comparing verb bias effects in VO and OV position in spoken language, we excluded these cases (n = 80 in CGN) as they may be argued to be only “halfway” OV.

5.2.2.2 The sentence completion task

Participants

40 adult native speakers of Dutch (ages 18 -30 years) from Radboud University Nijmegen participated in the study. They received payment for their participation. All participants had normal or corrected-to-normal vision. Permission to conduct the study had been obtained from the Ethics Board of the Social Sciences Faculty of the Radboud University Nijmegen.

Materials (see Appendix A)

We selected 16 verb pairs from the CGN corpus. Each verb pair consisted of a DO- (e.g., *uitleggen* 'explain') and a PO-biased (e.g., *opleggen* 'impose') verb, as determined on the basis of the distinctive collexeme analysis we had applied to all ditransitive verbs in CGN. The selected verbs occurred in a ditransitive construction within the corpus at least two times. For each verb pair, one main clause (as in (2a/b)) was constructed which could accept both verbs and which was based on materials used earlier in our lab (Van de Velde & Meyer, 2014) and in Van Bergen, Van Lier, and De Swart (2014).

(2a)	Het	Schoolhoofd	legt	de	leerlingen	de	regels	uit
DO bias	The	Headmaster	explains	the	students	the	rules	
	'The headmaster explains the rules to the students'							
(2b)	Het	Schoolhoofd	legt	de	leerlingen	de	regels	op
PO bias	The	Headmaster	imposes	the	students	the	rules	
	'The headmaster imposes the rules on the students'							

These sentences were broken down into a preamble *Het schoolhoofd ...*, and three sentence fragments: recipient, theme and verb (infinitive). Each verb pair was presented with two fragment orders: recipient-verb-theme (RT order) and theme-verb-recipient (TR order), e.g., [*leerlingen uitleggen regels*] and [*regels uitleggen leerlingen*], respectively.

Four lists of stimuli were created to counterbalance verb bias and fragment order, so that each item appeared in a different condition (PO vs. DO bias, and recipient-verb-theme vs. theme-verb-recipient order) in each list. Target sentences were each separated by two filler sentences (34 fillers in total). The fillers were monoclausal sentences, usually including intransitive verbs (e.g., *The elderly man slept through the ceremony*). They were presented in a manner similar to the presentation of the target items, e.g., *The elderly man* ... [*ceremony slept through*]. The verb was always presented as the second fragment, whereas the order of the remaining fragments was randomized with approximately 50% displaying an order congruent with the original sentence order (e.g., [*through slept ceremony*]), the other half incongruent (i.e. [*ceremony slept through*]).

Within each list, target items were organized such that (a) no two consecutive target items involved a verb with the same preference direction (PO or DO), and (b) no more than two target items with similar fragment orders were shown consecutively. All sentences were printed in a pen-and-paper questionnaire.

Procedure

Participants were randomly assigned to one of the four item lists and tested in small groups. They received a sheet with written instructions, and two sheets containing the sentence fragments, after giving informed consent. Instructions emphasized that sentence fragments had to be completed to grammatically correct sentences and that present tense had to be used whenever possible. Furthermore, the participants were instructed to work quickly, not to use many extra words, and to write the first thing that came to mind. The task consisted of 16 experimental and 34 filler items. At the end of the session, participants were debriefed about the goal of the task. It took participants about 20 minutes to complete the task.

Scoring and analysis

Sentences were scored as embodying either double object (DO) dative or prepositional object (PO) dative syntax. Sentences with intransitive syntax or other constructions were discarded, as were sentences with the direct or indirect object omitted, with verb substitutions, or with noun substitutions (except for minor substitutions, e.g., *oma* 'grandma' instead of *opa* 'grandpa', or the use of plural instead of singular nouns). The final dataset consisted of 569 responses (253 PO sentences, 316 DO sentences), equivalent to fourteen scorable responses per participant.

To evaluate the sentence completion task, we tested whether verb bias (i.e., PO- vs. DO-preference as estimated on the basis of the CGN data) was a significant predictor of structure choices in the completion task, independent of the order in which theme and recipient were presented. To this end we performed a mixed logit model predicting the logit-transformed likelihood of a PO-response (Jaeger, 2008). The model included Verb preference and Fragment order as fixed categorical factors, after they had been centered. We used a backwards elimination procedure, starting from an initial model containing all experimental factors and their interactions, as well as a maximum random structure, to arrive at the model that best fits the data (Baayen, 2008a; Baayen, Davidson, & Bates, 2008; Jaeger, 2008).

Results

Figure 1 shows the mean proportion of PO responses per condition. On average, PO datives were produced in 44% of the trials, which is more often than in the CGN speech corpus (23%). Previous studies have also found that PO datives occur more frequently in experimental studies (e.g., Bernolet & Hartsuiker, 2010) than in natural corpus data (e.g., Colleman, 2009). This difference has been attributed to the lack of pronominal arguments in experimental studies. Pronominal indirect objects tend to be formulated as NPs, thereby boosting the production of DO datives in natural corpus data (Colleman & Bernolet, 2012).

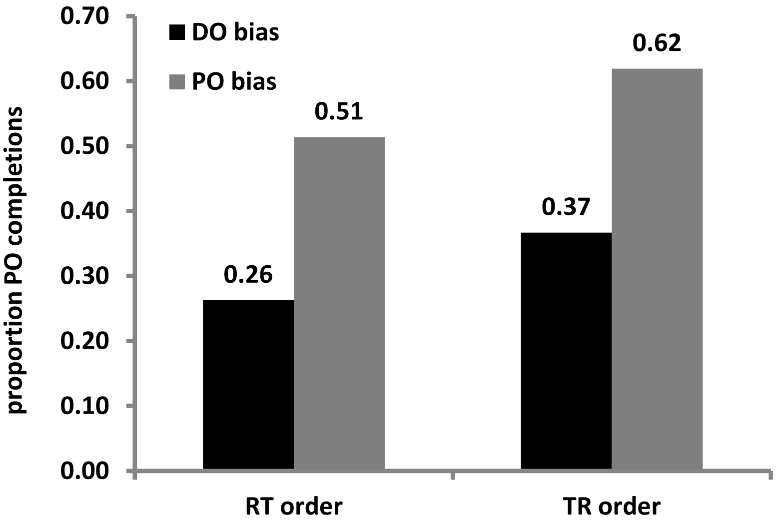

Figure 1. Proportion of PO completions per condition. “RT” and “TR” denote order of presentation of theme and recipient fragments in the experimental items.

To test for effects of Fragment order and Verb preference, we ran a mixed logit model. The final model yielded significant main effects of Verb preference and Fragment order (with by-subject and by-item random intercepts, and a by-subject random slope for Fragment order)[3]. The interaction between Verb preference and Fragment order was not significant and was thus excluded from the model. The results of the analysis are summarized in Table 2. The significant main effect of Verb bias confirms the finding of van Bergen et al.

[3] Including more by-item or by-subject random slopes (for a maximum random structure) led to non-convergence of the model.

(2014) that verb preferences are similar across corpora and controlled experiments. In particular, the direction of verb bias (i.e., DO versus PO) based on CGN data was a significant predictor of sentence structure in the completion task, above and beyond the effect of Fragment order.

Table 2. Summary of the fixed effects in the mixed logit model (N = 569; loglikelihood = -330).

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	-0.22	0.31	-0.70	0.48
Fragment order	0.68	0.27	2.56	< 0.05
Verb preference	1.36	0.22	6.13	< 0.001

5.2.2.3 Distinctive collexeme analysis

The data from the sentence completion task (all VO), as well as the finally selected CGN and Alpino sentences (classified as VO or OV), were subjected to *distinctive collexeme analysis* to determine their DO or PO verb biases (Gries, 2006; Gries, 2007; Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003, 2005). This technique allows us to establish the degree of preference of lexical items (here: ditransitive verbs) for one syntactic construction over another (here: DO versus PO dative). It does so by comparing the observed frequency of a given ditransitive verb occurring in one of the alternative syntactic constructions with the expected frequency of occurrence based on the overall distribution of the two alternative constructions in the set of ditransitive verbs.

The statistical test used for this purpose is the Fisher-Yates Exact Test (FET). This statistic does not involve distributional assumptions (e.g., normality), nor does it overestimate verb bias for rare verb–structure pairings (Stefanowitsch & Gries, 2003). The result is a *p*-value, indicating the preference of the verb for one of the two constructions. Verb bias is usually expressed as $-\log_{10}(p_{FET})$, with higher values indicating a stronger preference for one construction over the other. Tables 3a and 3b show the resulting distinctive collostructional strengths of the ten verbs with the highest preferences for a DO and a PO structure in the CGN and Alpino corpora. Although, in both corpora, DO structures occur more often than PO structures, PO-preferring verbs display stronger biases than DO-preferring verbs in both CGN (*M* = 0.93, *SD* = 1.62, *range* = 0.18 - 11.47 vs. *M* = 0.32, *SD* = 0.61, *range* = 0.12 - 5.06) and in Alpino (*M* = 0.58, *SD* = 0.33, *range* = 0.16 - 1.50 vs. *M* = 0.25, *SD* = 0.19, *range* = 0.17 - 1.20).

Table 3. Collexemes with the strongest overall bias (i.e., collapsing across verb position) for the DO and the PO dative in (a) CGN (spoken corpus) and (b) Alpino (written corpus)

(a) CGN		**(b) Alpino**	
DO preference		**DO preference**	
Dutch (English)	**coll. strength**	**Dutch (English)**	**coll. strength**
zeggen (say)	5.06	bezorgen (deliver)	1.20
vertellen (tell)	3.83	doen (do)	0.68
maken (make)	1.16	vertellen (tell)	0.68
bieden (offer)	1.00	aandoen (affect)	0.51
gunnen (award)	0.93	maken (make)	0.51
kwalijk_nemen (resent)	0.93	opleveren (yield)	0.51
beloven (promise)	0.81	aanbieden (offer)	0.39
voorhouden (hold_up)	0.81	doorgeven (pass_on)	0.34
wensen (wish)	0.81	toekennen (grant)	0.34
kosten (cost)	0.72	voorschotelen (serve)	0.34
PO preference		**PO preference**	
Dutch (English)	**coll. strength**	**Dutch (English)**	**coll. strength**
vragen (ask)	11.47	overdragen (transfer)	1.50
verkopen (sell)	3.77	overhandigen (hand)	1.50
schrijven (write)	2.48	afdragen (pay)	0.99
hebben (have)	2.17	meedelen (communicate)	0.99
afstaan (cede)	1.26	uitreiken (distribute)	0.99
laten (let)	1.26	verkopen (sell)	0.99
verlenen (grant)	1.26	schenken (donate)	0.62
zoeken (search)	1.26	zenden (send)	0.62
teruggeven (return)	0.86	berekenen (calculate)	0.49
aanbieden (offer)	0.78	beschikbaar_stellen (make_available)	0.49

Since p_{FET}-values are dependent on sample size, this measure does not lend itself to direct comparisons of degrees of association derived from samples of different sizes (Gries, 2006; Wiechmann, 2008). Therefore, in the correlation analyses to be reported below, we used $\log_{10}$ odd ratios as a measure of association strength to enable direct comparisons across corpora (CGN and Alpino) and data sets (VO and OV). The $\log_{10}$ odd ratio of a verb was defined as ($\log_{10}$[(#DO + 1)/(#PO + 1)]). We added 1 to the obtained frequency counts (e.g., #DO +1) in order to deal with zero frequencies (cf. Bernolet & Hartsuiker, 2010).

5.2.3 Data analysis

All data were analyzed using R (R Development Core Team, 2013), and the R packages *lme4* (Bates, Maechler, & Dai, 2009) and *languageR* (Baayen, 2008b). Analyses on structure choice in VO and OV dative clauses were carried out with mixed logit models (coefficients are given in log-odd ratios). Model factors include Verb-second bias (i.e. DO/PO bias in verb-second position) as a continuous factor and Verb position (VO vs. OV) as a categorical factor (after they have been centered), and a random by-item (verb) intercept (Baayen, 2008a; Baayen, et al., 2008; Jaeger, 2008). All models use a maximum random structure as justified by the model's design, with a by-item random slope for Verb position (Barr, Levy, Scheepers, & Tily, 2013). The model with an interaction between Verb bias and Verb position is compared against the model containing only main effects for Verb-second bias and Verb position, using likelihood ratio tests.

5.3 Results: Checking two presuppositions and answering the key question

In separate sections (5.3.1 through 5.3.3), we report results from three sets of analyses. Firstly, to check the assumption that overall verb biases are similar between spoken and written language, we correlate verb biases obtained from corpora of written vs. spoken language (in $\log_{10}$ odd ratios), after excluding outliers based on extreme Cook's distances (Stevens, 1984). Besides comparing biases obtained from CGN and Alpino, we add verb biases from another corpus of written language: the newspaper component of the CONDIV corpus. The CONDIV corpus, consisting of articles from three Dutch and three Belgian newspapers, was analyzed by Colleman (2009). In order to evaluate cross-corpus similarity of bias *direction*, we also compute generalized kappa for the three corpora while correcting for chance (Fleiss, 1971; King, 2004). This statistic can be interpreted as a chance-corrected measure of agreement among three or more independently rated categories (here: verb biases derived from three independent corpora).

Secondly, in order to check the assumption that verb biases in VO and OV position *in written language* are identical, we carry out a correlation analysis between VO verb biases obtained from the written sentence completion task on the one hand, and overall verb biases (VO and OV together, all in Log_{10} odd ratios) from the CONDIV corpus on the other. Because OV-order is used more frequently in written (and spoken) language, these overall biases must be largely based on OV-clauses. Hence, a high positive correlation suggests that VO and OV biases are very similar. In order to explore the similarity of VO- and OV-based biases, we can only utilize the CONDIV corpus as the other written corpus (Alpino) is too small (see Section 5.3.2).

The third and final analysis step focuses on our key question: Do verb bias effects on structure choices *in spoken language* vary depending on the VO vs. OV position of the verb in the clause? Using a mixed logit model (Jaeger, 2008) with VO-specific Verb bias (obtained from the sentence completion task) and Verb position as fixed factors, we predict the selection of DO- versus PO-structure in dative clauses of the CGN corpus. This analysis uses the full CGN dataset consisting of 666 inflected ditransitive verb tokens, without aggregating data on the verb level. (The 666 verb tokens belong to 32 different verb types, i.e. citation forms, which overlapped between CGN and the sentence completion task.)

5.3.1 Cross-corpus agreement of DO/PO biases

First we examine the overall similarity of the verb biases in the three target corpora. Table 4 shows the correlations among verb biases measured in Log_{10} odd ratios.

Table 4. Correlations among verb biases (expressed as log_{10} odd ratios) obtained from the Alpino and CGN corpus in the present study, and from the CONDIV corpus in Colleman (2009). Between parentheses are the number of different ditransitive verbs and the *total number of verb form occurrences* per analysis.

	Alpino	CGN	CONDIV
Alpino			
CGN	.49** (46 - *936*)		
CONDIV	.66** (66 - *11782*)	.60** (98 - *14811*)	

**Correlation is significant at the .01 level (two-tailed).

Correlations between verb biases in the different corpora, taking direction of preference into account, were strong. The highest correlation obtained between verbs in the CONDIV and Alpino corpus, which both contain contemporary Dutch newspaper language. The fact that verb biases obtained from the CGN corpus correlate strongly with verb biases in CONDIV and Alpino implies that verb biases affect structure choices to a similar degree in spoken and written modalities.

To assess the similarity in the *direction* of verb preferences between corpora we compute generalized kappa (Fleiss' kappa). Generalized kappa can be used to measure the degree of cross-corpus agreement with respect to the direction of verb bias, over and above the agreement that would be expected on the basis of chance alone (Fleiss, 1971). According to this statistic, agreement differs significantly from chance ($z = 2.94$, $SE_{Fleiss} = .10$, $p < .01$).The value of Kappa =.29 is characterized as fair (Landis & Koch, 1977).

5.3.2 DO/PO biases in written VO and OV clauses

In order to establish that in *written* language, verb biases do not differ when used in different clause positions, we run a correlation analysis between VO biases obtained from the sentence completion task and overall biases (expressed as $\log_{10}$odd ratios). Unfortunately, Alpino proves too small for this analysis[4], so we only utilize the verb biases compiled by Colleman (2009) over verbs in OV and VO positions together in the CONDIV corpus. However, since OV-order is far more frequent than VO-order (75 vs. 25 %) in written language, biases are largely OV-based. Results show that within written language, VO biases (sentence completion task) are highly similar to OV biases (CONDIV corpus), $r = .82$, $p < .001$, $n = 27$. (Three verbs—*brengen* 'bring', *verwijten* 'blame' and *leren* 'teach'—were excluded from this analysis due to extreme Cook's distances.)

5.3.3 DO/PO bias in spoken VO and OV clauses as indicator of planning scope

So far, we have found that overall verb biases are similar across written and spoken modalities and that within *written* language VO and OV-based biases are virtually the same. Addressing now our main question as to whether the verb is planned in advance in OV clauses, we turn to a comparison of verb bias effects in VO vs. OV clauses within *spoken* language. To this end we first calculate, through distinctive collexeme analysis, the mean VO-based DO/PO biases in the sentence completion task. Table 5 summarizes the outcome. The thus obtained verb biases, together with Verb position, are used as fixed factors in a mixed logit model predicting DO/PO choice in the CGN corpus. The dataset for this analysis consists of 666 CGN clauses that contain one of the verbs of the sentence completion task. See Table 6 for a summary of the results.

Table 5. Mean strength of the DO and PO biases (in $-\log_{10}(p_{\mathrm{FET}})$) computed from the sentence completion task.

	Verb preference	
Verb bias	**DO**	**PO**
Average	2.10	1.81
Stdev	1.41	1.77
Range	0.33 - 5.23	0.25 - 7.23

[4] Alpino contains 28 verbs occurring in VO position, but only five of these occur more than once in the corpus (such that only seven unique biases can be computed). Moreover, for the comparison between VO and OV based verb biases, only verbs that occur in both positions can be included. This would yield $n = 12$ verbs for the comparison of Alpino-based verb biases in OV and VO position, with the additional limitation that these biases would be calculated based on very few cases.

Table 6. Summary of the fixed effects in the mixed logit model predicting DO/PO choice in the CGN corpus (*n* = 666, log likelihood = -282). The model also includes a random by-item intercept. A random slope for Verb position is not included, as it leads to non-convergence.

Predictor	Coefficient	*SE*	Wald *Z*	*p*
Intercept	-1.96	0.31	-6.42	<.001
Verb bias	-0.47	0.14	-3.29	<.001
Verb position	-0.17	0.24	-0.71	0.48
Verb bias x Verb position	0.67	0.27	2.47	0.01

In the final model (Table 6), Verb bias is a significant predictor of syntactic structure of OV and VO clauses in CGN. Additionally, the model yields a significant interaction between Verb bias and Verb position. A comparison of this model with a model that only contains main effects reveals that the inclusion of the interaction between Verb Bias and Verb Position is justified ($X^2(1, N = 666) = 7.96, p < 0.01$): A significant amount of variance is explained by the interaction. Hence, the effect of verb bias on structure choices is significantly different for verbs in different clausal positions[5].

To examine the interaction more closely, we split the data into VO (*n* =198) and OV (*n* = 468) clauses and examined the degree to which Verb bias predicts syntactic structure for each verb position separately. As expected, Verb bias is a strongly significant predictor of structure choices in VO clauses, ($\beta = -0.77$, $SE = 0.28$, $z = -2.72$). In OV clauses, Verb bias is also a significant predictor of structure choices ($\beta = -0.25$, $SE = 0.11$, $z = -1.93$). Although the effect was weaker than in VO clauses, this finding supports the hypothesis that verbs are usually selected prior to the selection of a DO or PO structure, even in dative OV clauses.[6]

[5] The outcome of this analysis remains essentially the same if we include the 80 OV clauses in CGN where the indirect object PP follows the ditransitive verb (see last paragraph of Section 5.2.2.1).

[6] We also calculated the correlation between verb biases as obtained from the sentence completion task (VO position) and verb biases obtained from CGN (OV position), after removing outliers (*n* = 3; *brengen* 'bring', *doorgeven* 'pass on', and *schrijven* 'write' due to extreme Cook's distances). Although this analysis does not take overall distributional differences into account (as verb biases were expressed in $\log_{10}$odd ratios to enable comparison between corpora of different sizes), the result confirms the findings from the mixed logit model: verb-second biases obtained from the sentence completion task correlate significantly with verb-final biases obtained from CGN, $r = .37$, $p = .05$, $n = 29$. We also computed the correlation between verb biases based on VO sentences in CGN and biases based on the sentence completion task, after removing outliers (*n* = 1, *garanderen* 'guarantee' due to an extreme Cook's distance). In line with the outcome of the mixed logit model, VO-based biases in CGN are more strongly correlated to biases computed from the completion task than OV-based biases are ($r = .78$, $p < .001$, $n = 21$).

5.4 Discussion

In this study, we compared the effect of verb biases on structure choices in VO and OV dative clauses in a corpus of spoken Dutch (CGN).We used the results as a kind of litmus test, enabling inferences about the size and structure of the advance planning scope during spontaneous speaking to be made. We contrasted two views of clausal pre-planning that entail differential predictions regarding the effect of verb bias in VO versus OV clauses. According to the hypothesis of *hierarchical incrementality*, the verb is an obligatory component of clausal planning scope. Therefore, verb biases should exert an influence on structure choices regardless of the early (VO) or late (OV) position of the verb in the clause. Conversely, according to a view of *linear incrementality*, planning proceeds in a piecemeal fashion strictly guided by lexical availability. Consequently, the verb and its associated bias can only influence structure choices in VO clauses. We tested these predictions by analyzing structure choices in CGN, using mixed logit models.

Results are in line with a combination of linear and hierarchical incrementality, showing a significant effect of verb bias on structure choices in OV clauses, but an even stronger effect of verb bias in VO clauses. The significant main effect of verb bias in OV clauses shows that during spontaneous speaking, verb biases of verbs in clause-final position can drive syntactic choices, in line with the hypothesis of hierarchical incrementality. According to this hypothesis, the head verb of the clause and its associated subcategorization frame have already been activated before unification and placement of the leftmost object NP of a dative OV clause. In fact, in many subordinate CGN clauses, the verb is preceded not only by the direct and/or indirect object but also by other arguments and/or adjuncts. Still, verb biases were a significant predictor of structure choices in these spontaneously produced OV clauses.

However, the significant interaction between verb bias and verb position indicates that verb bias exerts a significantly weaker influence on structure choices in OV than in VO clauses. This finding is in line with a combination of hierarchical and linear planning strategies, suggesting that speakers engage in hierarchically incremental planning most of the time, but sometimes resort to a more linearly incremental planning strategy. For, if sentences were planned in a strictly hierarchically incremental fashion, we would expect structure choices in VO and OV clauses to be determined by verb bias equally strongly.

As discussed at the beginning of Section 5.2, when comparing verb bias effects in VO and OV position, we presupposed that lemma representations, including any verb biases, are independent from the verb's linear position in the clause. We established this by comparing experimentally elicited VO biases and overall corpus-based biases (largely based on OV clauses) within written language. The result of a correlation analysis indicated that, within written language, verb biases exerted by verbs in VO position are very similar to those in OV position. Unfortunately, due to the small size of the Alpino corpus, we could not make the direct comparison between VO and OV based biases within one corpus of written language. The corpus even proved too small to derive OV-based biases from.

In Section 5.1.1, dealing with the experimental literature on the scope of clausal pre-planning, we briefly summarized Lindsley's (1975) Semi-Predicate model, which holds that at the onset of overt production of simple SVO clauses the speaker has planned the head verb at least at a conceptual level, i.e. pre-lexically. The fact that this model received experimental support in English and Dutch raises the question whether some form of *pre-lexical* preparation of the head verb could explain the (weak but significant) effect of DO/PO bias in OV clauses. Consider the scenario in which a speaker has decided to extend an NP with a relative clause and selects OV word order. Then, given the event s/he is about to describe, s/he activates the general concept of *transfer-of-possession*. Since this concept does not uniquely select one particular verb, the PO and DO options for the dative structure are both open at this point, e.g., PO with *verkopen* 'sell'; DO with *bieden* 'offer' (see Table 2). Now, suppose the speaker encodes the constituents referring to the agent and the recipient of the transfer-of-possession event *prior to having decided on the verb*. At this pre-lexical stage of verb planning, the speaker could select the DO option, based on the fact that DO structures occur more frequently in Dutch than PO structures, and thus serve as the default option. The two dependent constituents can now be inserted into a generic clausal structure—the indirect preceding direct object NP—, followed by the clause-final verb.

This version of the Semi-Predicate hypothesis can be put to the test by comparing the predictive power of a model for PO/DO choices in OV clauses containing verb-specific biases as a predictor against a model with only an intercept—the latter representing the default bias in favor of DO. If pre-lexical selection of a DO vs. PO dative is indeed a viable scenario, then adding verb-specific biases to a model predicting PO/DO choices will not explain significantly more variance than the intercept-only model. Comparing the intercept-only model with a model that included an intercept and the factor Verb bias, we found that the latter factor yields a (marginally) significant increase of predictive power ($X^2(1, N = 468) = 3.67$, $p = 0.055$). This result supports our assumption that the data pattern we obtained above is based on lexical rather than pre-lexical pre-planning of verbs in OV position.

However, the study by Schriefers, et al. (1998), which we briefly discussed in Section 5.1.1, seems to contradict this conclusion. They interpreted their data as showing that verbs in clause-final position are not automatically part of the grammatical advance planning scope. Recall that, in a picture-word interference paradigm, participants had to describe pictures accompanied by visually and auditorily presented distractor verbs. Verbs were either semantically related, syntactically deviant, or unrelated to the action depicted. Lead-in fragments, presented before the onset of the picture, induced the production of sentences with either SOV or VSO order. Onset latencies revealed effects of semantic distractors (i.e., semantic interference) in VSO sentences only, suggesting that the verb is not prepared before sentence onset in SOV sentences. This conclusion clearly disagrees with our current findings.

What could be the reason for this discrepancy? The fact that Schriefers et al. studied a different language (German instead of Dutch) and different sentence types (intransitive and monotransitive instead of ditransitive clauses), may not be irrelevant. We suggest that the experimental paradigm they deployed may have induced a more linearly incremental planning strategy. Previous research has shown that speakers tend to plan more linearly incrementally (a) when experiencing cognitive load, (b) when under time pressure, and/or (c) when their average production speed is high (Ferreira & Swets, 2002; Wagner, Jescheniak, & Schriefers, 2010).

For example, using a semantic interference paradigm, Wagner, Jescheniak and Schriefers (2010) found that the lexical advance planning scope during the production of simple sentences decreased under cognitive load. Participants had to produce simple (e.g., *the frog is next to the mug*) and slightly more complex (e.g., *the red frog is next to the red mug*) sentences describing objects on a screen. Unrelated and semantically related auditorily presented words distracting from the first or the second noun were presented simultaneously. Onset latencies were measured to make inferences about the scope of grammatical advance planning: if sentence onsets were delayed by a distractor related to the first noun, then speakers' planning scope should include this word (*the frog)*. Similarly, if sentence onsets are susceptible to semantic interference effects on the second object, speakers' scope of planning also included the second noun phrase (*the mug)*.

In two of their experiments, speakers only used one sentence type (simple or with adjectives) while in one experiment they had to switch between sentence types. To decide which sentence type they should use, participants had to compare the size of the depicted objects to a given standard (a conceptual decision task). In one participant group, simple sentences had to be produced when depicted objects were large and complex sentences when object pairs were small, while the other group was instructed to do the reverse. Wagner et al. found that speakers made use of more hierarchically incremental planning, showing interference from the semantically related auditory distractors on both the first and the second noun, when they only had to produce one sentence type. Conversely, when they had to switch between sentence types, speakers only showed interference effects on the first noun, suggesting a smaller grammatical planning scope.

Similarly, in Schriefers et al.'s picture-word interference paradigm, cognitive load might have been induced by the simultaneous presentation of auditory or visual distractors (the latter flanking the to-be-described picture). The fact that pictures without distractors were described faster than pictures flanked by verbs identical to the to-be-produced verb suggests that speakers were indeed distracted by the presentation of any material that flanked the picture. In addition, the repetition of each picture with different distractors and lead-in fragments (each picture was seen ten times, with two different lead-in fragments in five distractor conditions) might have yielded high processing difficulty. Still, speakers initiated their sentences rather fast: 880 ms was the average onset latency for SOV sentences with unrelated distractors. Therefore, it is conceivable that, in the paradigm of Schriefers et al., speakers relied on linearly incremental planning to a large extent. Since our study used corpus data from spontaneous speech, we have no control over factors that influence planning strategies. Our results suggest, however, that in everyday spontaneous conversation speakers often plan hierarchically incrementally.

Besides verb biases, there are many other factors influencing the choice for a dative alternative such as definiteness and number of the recipient and theme NPs (Bresnan, Cueni, Nikitina, & Baayen, 2007). In fact, argument properties and verb biases may be interdependent in determining structure choice: Certain verbs may co-occur with specific argument types (e.g., *take* is over seven times more likely to have a non-given recipient than *bring*; Bresnan, et al., 2007), thereby indirectly determining structure choices. However, even though we did not look at properties of recipients and themes in our corpus, several findings indicate that verb biases exert an independent effect on structure choices.

Firstly, results from controlled sentence production tasks suggest that verb biases may affect structure choices in datives, even in the presence of careful control of argument properties (i.e., presentation order, syllable count, frequency) that might influence the accessibility of the two objects in the to-be-produced utterance. Both in the computerized sentence completion task explored by Van Bergen et al., (2014) and in the written paper-and pencil questionnaire of the present study, corpus-based verb biases generalized well to a controlled experimental setup, and vice-versa.

Secondly, the high cross-corpus agreement with respect to verb biases (in strength and direction) indicates that verb biases are robust and relatively immune to the extra-grammatical factors that differ between spoken and written language, such as memory limitations, processing load, and the speaker-hearer context (Bresnan, et al., 2007). This is in line with earlier studies confirming the robustness of effects of probabilistic information on language production (Bresnan, et al., 2007; Kuperman & Bresnan, 2012; Tily et al., 2009).

In sum, the overall pattern of results is in line with a hierarchically incremental planning mechanism underlying clausal pre-planning in many language production situations. According to hierarchical planning, lexical units are unified into hierarchically organized syntactic constituents prior to being linearized and phonologically encoded. Crucially, according to this view of sentence planning, the selection of a syntactic option from a set of alternatives takes place during the mapping from conceptual to functional-grammatical roles. Verb biases, as one of the factors driving syntactic choices, operate on this pre-linearized level as well. This allows ditransitive verbs to exert their DO or PO preference even if they do not precede the direct and indirect objects in a dative clause: The syntactic choice is made in a pre-linear stage. In some OV clauses, however, verb bias may not be the eventual predictor of structure choice. For instance, when the head verb is hard to access, more easily accessible nonverbal constituents may already be inserted into the developing sentence frame prior to the verb, serving to promote production fluency. This may give rise to mismatches between linearization driven by lexical accessibility and linearization driven by verb preferences. In some OV clauses, therefore, the balance may tip towards a choice that is not in line with the verb's preference. Moreover, besides linguistic factors contributing to the weighting of syntactic choices, extra-linguistic factors may promote a more linearly incremental planning strategy. For instance, under conditions of high cognitive load or time pressure, the speaker may prefer a linearly incremental planning strategy with limited look-ahead. The present study, however, excludes strictly linearly incremental planning as the default planning strategy in spontaneous speech.

Acknowledgements

We would like to thank Peter de Swart for helpful discussions and Annelies van Wijngaarden for help with data processing.

Appendix A

Target sentences used for constructing sentence fragments in the sentence completion task

pair nr	verb	bias	target sentence in Dutch	target sentence in English
1	beloven	DO	De organisator belooft de winnaar de prijs.	The organizer promises the winner the prize.
	teruggeven	PO	De organisator geeft de winaar de prijs terug.	The organizer gives the winner the prize back.
2	geven	DO	De militair geeft de vluchteling het voedselpakket.	The soldier gives the refugee the food package.
	brengen	PO	De militair brengt de vluchteling het voedselpakket.	The soldier brings the refugee the food package.
3	vertellen	DO	De directeur vertelt de medewerker de maatregel.	The director tells the employee about the measure.
	voorleggen	PO	De directeur legt de medewerker de maatregel voor.	The director presents the employee the measure.
4	uitleggen	DO	Het schoolhoofd legt de leerlingen de regels uit.	The headmaster explains the students the rules.
	opleggen	PO	Het schoolhoofd legt de leerlingen de regels op.	The headmaster imposes the students the rules.
5	verwijten	DO	De bedrijfsleider verwijt de kassière de fout.	The manager blames the cashier the error.
	duidelijk maken	PO	De bedrijfleider maakt de kassière de fout duidelijk.	The manager makes the cashier the error clear.
6	meegeven	DO	De voorzitter geeft de secretaris het geld mee.	The Chairman gives (along) the secretary the money.
	doorgeven	PO	De voorzitter geeft de secretaris het geld door.	The Chairman passes (through) the secretary the money.
7	garanderen	DO	De coach garandeert de aanvoerder de trofee.	The coach ensures the captain the trophy.
	toevertrouwen	PO	De coach vertrouwt de aanvoerder de trofee toe.	The coach entrusts the captain the trophy.
8	bieden	DO	De consultant biedt de ondernemer de bedrijfsstrategie.	The consultant provides the entrepreneur the business strategy.
	verkopen	PO	De consultant verkoopt de ondernemer de bedrijfsstrategie.	The consultant sells the entrepreneur the business strategy.
9	leren	DO	De professor leert de promovendus de theorie.	The professor teaches the PhD candidate the theory.
	voorstellen	PO	De professor stelt de promovendus de theorie voor.	The professor proposes the PhD candidate the theory.
10	zeggen	DO	De journalist zegt de redacteur het nieuws.	The journalist tells the editor the news.
	schrijven	PO	De journalist schrijft de redacteur het nieuws.	The journalist writes the editor the news.
11	noemen	DO	De ondernemer noemt de financier het bedrag	The entrepreneur mentions the financer the amount.
	aanbieden	PO	De ondernemer biedt de financier het bedrag aan	The entrepreneur offers the financer the amount.
12	mailen	DO	De cliënt mailt de advocaat een vraag.	The client mails the lawyer a question.
	sturen	PO	De cliënt stuurt de advocaat een vraag.	The client sends the lawyer a question.
13	betalen	DO	De bewoner betaalt de verhuurder de schade.	The tenant pays the landlord the damage.
	mededelen	PO	De bewoner deelt de verhuurder de schade mee.	The tenant informs the landlord of the damage.
14	vergeven	DO	De staatssecretaris vergeeft de minister de leugen.	The Secretary forgives the Minister the lie.
	melden	PO	De staatssecretaris meldt de minister de leugen.	The Secretary reports the Minister the lie.
15	bewijzen	DO	Het bedrijf bewijst de klant een dienst	The company renders the customer a service
	leveren	PO	Het bedrijf levert de klant een dienst	The company provides the customer a service
16	wijzen	DO	De automobilist wijst de fietser de weg	The driver points the cyclist the way.
	vragen	PO	De automobilist vraagt de fietser de weg	The driver asks the cyclist the way.

References

Akinnaso, F. N. (1982). On the differences between spoken and written language. *Language and Speech, 25*(2), 97-125.

Baayen, R. H. (2008a). *Analyzing linguistic data* (Vol. 505): Cambridge University Press Cambridge, UK.

Baayen, R. H. (2008b). languageR: Data sets and functions with “Analyzing Linguistic Data: A practical introduction to statistics”. *R package version 0.953*.

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

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

Bates, D., Maechler, M., & Dai, B. (2009). lme4: linear mixed-effects models using S4 classes. 2009. R Package, version 0.999375–31.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? *Cognition, 114*(3), 455-461.

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

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

Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. *Cognitive foundations of interpretation*, 69-94.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. *Cognitive Science, 26*(5), 609-651.

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

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass, 6*(5), 259-278.

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

Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences, 31*(5), 593-611.

Colleman, T., & Bernolet, S. (2012). Alternation biases in corpora vs. picture description experiments: DO-biased and PD-biased verbs in the Dutch dative alternation. *Frequency effects in language representation*, 87-125.

Ferreira, F. (1994). Choice of passive voice is affected by verb type and animacy. *Journal of Memory and Language, 33*(6), 715-736.

Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. *Aspects of language production*, 291-330.

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.

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76*(5), 378.

Gahl, S., & Garnsey, S. M. (2004). Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language*, 748-775.

Garnsey, S. M., Pearlmutter, N. J., Myers, E., & Lotocky, M. A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language, 37*(1), 58-93.

Garrett, M. (1976). The Analysis of Sentence Production1. *Psychology of learning & motivation: V. 9, 9*, 133.

Gries, S. T. (2006). Exploring variability within and between corpora: some methodological considerations. *Corpora, 1*(2).

Gries, S. T. (2007). Coll.analysis 3.2a. A program for R for Windows 2.x.

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective onalternations'. *International journal of corpus linguistics, 9*(1), 97-129.

Haeseryn, W., Romijn, K., Geerts, G., Rooij, J. d., & Toorn, M. v. d. (1997). Algemene Nederlandse Spraakkunst. Groningen/Deurne: Martinus Nijhoff uitgevers.

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

Hoekstra, H., Moortgat, M., Schuurman, I., & Van Der Wouden, T. (2001). Syntactic annotation for the spoken Dutch corpus project (CGN). *Language and Computers, 37*(1), 73-87.

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.

Joshi, A. K. (1987). An introduction to tree adjoining grammars. *Mathematics of language, 1*, 87-115.

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

Kempen, G., & Huijbers, P. (1983). The lexicalization process in sentence production and naming - indirect election of words. *Cognition, 14*(2), 185-209.

King, J. E. (2004). *Software solutions for obtaining a kappa-type statistic for use with multiple raters.* Paper presented at the annual meeting of the Southwest Educational Research Association.

Konig, E., & Lezius, W. (2003). *The TIGER language*: Tech. Rep., Univ. of Stuttgart.

Koster, J. (1975). Dutch as an SOV language. *Linguistic analysis, 1*(2), 111-136.

Kroch, A. S., & Joshi, A. K. (1985). The linguistic relevance of tree adjoining grammar. *Technical Reports (CIS)*, 671.

Kuperman, V., & Bresnan, J. (2012). The effects of construction probability on word durations during spontaneous incremental sentence production. *Journal of Memory and Language, 66*(4), 588-611.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics, 33*(1), 159-174.

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

Lindsley, J. R. (1975). Producing simple utterances - how far ahead do we plan. *Cognitive Psychology, 7*(1), 1-19.

O'Donnel, R. C. (1974). Syntactic differences between speech and writing. *American Speech. A Quarterly of Linguistic Usage, 49*(1-2), 102-110.

R Development Core Team. (2013). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

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

Stallings, L. M., MacDonald, M. C., & O'Seaghdha, P. G. (1998). Phrasal ordering constraints in sentence production: Phrase length and verb disposition in heavy-NP shift. *Journal of Memory and Language, 39*(3), 392-417.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics, 8*(2), 209-243.

Stefanowitsch, A., & Gries, S. T. (2005). Covarying collexemes. *Corpus Linguistics and Linguistic Theory, 1*(1), 1-43.

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin, 95*(2), 334.

Tily, H., Gahl, S., Arnon, I., Snider, N., Kothari, A., & Bresnan, J. (2009). Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition, 1*(2), 147-165.

Uit den Boogaard, P. (1975). Woordfrequenties in geschreven en gesproken Nederlands.

Van Bergen, G., Van Lier, E., & De Swart, P. (2014). *Constructional variation in Dutch ditransitives: Lexical preferences in controlled production.* Manuscript submitted for publication.

Van de Velde, M., & Meyer, A. S. (2014). *Syntactic flexibility and planning scope: The effect of verb bias on grammatical advance planning.* Manuscript submitted for publication.

Van der Beek, L., Bouma, G., Malouf, R., & Van Noord, G. (2002). The Alpino dependency treebank. *Language and Computers, 45*(1), 8-22.

Van Eerten, L. (2007). Over het Corpus Gesproken Nederlands. *Nederlandse Taalkunde, 12*(3), 194-215.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition, 68*(1), B13-B29.

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.

Wiechmann, D. (2008). On the computation of collostruction strength: Testing measures of association as expressions of lexical bias. *Corpus Linguistics and Linguistic Theory, 4*(2), 253-290.

Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language, 60*(3), 368-392.

Zwart, C. J. W. (1993). *Dutch syntax: A minimalist approach.* Rijksuniversiteit Groningen.

6

SUMMARY AND GENERAL DISCUSSION

6.1 Introduction

At this point, a (very committed) reader will have processed nearly 3000 sentences. These sentences were carefully put together to bring across information, ideas, and beliefs about sentence production. Two topics within sentence production played a central role in this thesis: incrementality and syntactic flexibility in sentence planning and formulation. The studies described in the preceding chapters were designed to find answers to two key questions concerning these topics:

(1) How do factors acting on different levels of the message-to-language mapping influence the degree of incrementality during advance sentence planning?
(2) How is syntactic flexibility implemented in the grammatical encoding system and how do speakers deal with syntactic choice available to them?

To answer these questions, this chapter first summarizes the main findings from the preceding chapters in chronological order. Subsequently, theoretical and methodological issues that require further attention are discussed, and questions for future research are highlighted.

6.2 Summary

Chapter 2 investigated how lower level perceptual and linguistic factors (salience and character codability) on the one hand and higher level linguistic factors (event codability and ease of structure assembly) on the other hand influenced structure choices and the time course of sentence planning and formulation in two eye-tracking experiments. Eye-gazes have been found to be tightly linked to message and sentence formulation processes (Griffin & Bock, 2000). For example, when asked to describe pictured events, speakers normally direct their gaze to the characters in the event in the order of mention and the timing of a gaze shift has been found to coincide with the completion of lexical encoding of the fixated referent (Griffin & Bock, 2000; Meyer, Sleiderink, & Levelt, 1998).

In two experiments, participants described a long series of pictures, including transitive target events eliciting descriptions with the preferred active construction or dispreferred passive construction. The depicted target events exhibited a wide range of codability (i.e., ease of conceptual encoding of the action) and featured agent and patient characters that were easy or difficult to name (character codability). Experiment 1 additionally employed a subliminal cuing manipulation to increase character salience and a reading task that included a high number of passive sentences mid-way through the experiment (i.e., passive syntax training) to facilitate passive structure assembly. Speakers were expected to shift from a *linearly incremental* planning strategy towards one that was more *hierarchically incremental* as the message-to-structure mapping was facilitated (i.e., when describing "easy" events and/or using a structure that is easy to assemble). Results indicated influences of factors modulating encoding of structural information (event structure and linguistic structure) on starting point selection and on the timecourse of formulation for active and passive sentences. In actives and passives, the early preference for fixating the agent was attenuated in highly codable events and after exposure to passive syntax, respectively. Only in low codability events, did speakers attempt to encode information about the most accessible character first (consistent with *linear incrementality*). The absence of a perceptual cuing effect on structure choices excluded an extreme form of linear incrementality in which speakers immediately begin linguistic encoding of a first-fixated character, irrespective of its status in the event. In sum, Chapter 2 confirms earlier findings showing that the influence of lower-level linguistic factors in causing a shift towards linearly incremental planning is constrained by higher level-linguistic factors concerned with the ease of the message-to-structure mapping (Kuchinsky & Bock, 2010). It extends earlier findings by showing the consequences of these effects on starting point selection and event apprehension for the entire timecourse of formulation.

Chapter 3 investigated how speakers make syntactic choices during grammatical encoding and whether executive control, specifically selective inhibition (SI), is involved in this process. The main question was whether syntactic alternatives compete for selection during sentence production. In an experiment making use of the RSVP paradigm, speakers produced dative verb phrases varying in syntactic flexibility. Syntactic flexibility was manipulated by eliciting sentences featuring verbs with a strong bias for one syntactic alternative (i.e., low syntactic flexibility) and sentences featuring verbs with a weak bias for one alternative (i.e., high syntactic flexibility). A competition-based view would predict that syntactic flexibility delays sentence production, due to competition between syntactic frames. In line with this view–and in contrast to earlier studies of syntactic flexibility with dative verbs (Ferreira, 1996)– participants were faster to initiate sentences featuring strong-bias verbs than weak-bias verbs. Additionally, participants with good SI (as assessed with the flanker task) showed a smaller verb bias effect than participants with poor SI, specifically in sentences with incongruent verb-structure pairings. This finding suggests that SI can support structure selection in sentences with high syntactic flexibility when grammatical encoding is difficult. The results of Chapter 3 speak to the second key question of this thesis by indicating that equipotent syntactic alternatives can compete during sentence production and that speakers make use of SI skills to resolve this competition efficiently.

Chapter 4 combined the themes of incrementality and syntactic flexibility by investigating whether the availability of multiple equipotent syntactic alternatives can influence the size of the advance planning scope during sentence recall. Again, an RSVP paradigm was used to elicit dative verb phrases with high vs. low syntactic flexibility. Additionally, to assess lexical planning scope, the frequency of the first (Experiment 1) or the second (Experiment 2) post-verbal noun (N1 or N2) was varied. Effects of N1 frequency on sentence onset latency would indicate that the advance planning scope included N1, while effects of N2 frequency would imply that the planning scope included N2 as well. Results showed no effect of N1 frequency on onset latencies in Experiment 1. An additional analysis revealed that any facilitatory effect of a high frequency N1 on onset latencies must have been cancelled out by the time needed to prepare the lower frequency N2s in Experiment 1. After we then manipulated the frequency of the second post-verbal noun in Experiment 2, an effect of N2 frequency was found only in syntactically flexible sentences and especially in slow speakers. These findings suggest that lexical planning scope can be extended up to and including the second noun phrase in sentences with high syntactic flexibility. Syntactic flexibility can thus be added to the range of linguistic factors influencing speakers' variations in planning scope.

Chapter 5 used a corpus-based approach to investigate whether items within the advance planning scope are linearly or hierarchically organized. To address this question, verb bias effects on structure choices in (S)VO and (S)OV dative clauses within a corpus of spoken Dutch (CGN) were compared. This comparison allowed inferences to be made about the size and structure of the advance planning scope during spontaneous speaking: if the verb is an obligatory component of clausal planning scope (in line with the hypothesis of verb-centered *hierarchical incrementality*), then biases should exert an influence on structure choices regardless of the early (VO) or late (OV) position of the verb in the clause. Conversely, if planning proceeds in a piecemeal fashion, strictly guided by lexical availability (in line with *linear incrementality*), then the verb and its associated bias can only influence structure choices in VO sentences. Results showed that verb preferences were a significant predictor of structure choices in OV sentences, but that the effect was weaker than in VO sentences. These findings were taken as evidence for advance planning that is mostly, but not entirely, in line with verb-centered hierarchical incrementality. Importantly, findings exclude strictly linearly incremental planning as the default planning strategy in spontaneous speech.

6.3 Discussion

Onset latencies as an index of both planning and competition processes

In most chapters, response latencies were used as an index of the sentence initial processing load associated with either planning (Chapter 2 and Chapter 4) or competition processes (Chapter 3). In other words, long(er) onset latencies were used interchangeably as a reflection of competition between structural frames and as proof of a shift towards parallel/ hierarchically incremental planning. But how is it possible to claim the existence of competition between frames if the same result (longer latencies) can also be explained with a change in planning strategies?

As the outcome of Chapter 4 suggests, competition and planning processes may be interrelated. Competition between abstract syntactic frames (allowing for different placement options of lexical material) when generating a syntactically flexible sentence, may increase the size of planning scope. Vice versa, syntactic structures can only engage in competition when planning is not piecemeal and solely guided by lexical accessibility, but when it occurs in a more wholistic fashion (Myachykov, Scheepers, Garrod, Thompson, & Fedorova, 2013). Chapter 4 has shown that the sentential planning scope is broadened in syntactically flexible sentences. This finding has been tentatively attributed to the fact that syntactic flexibility promotes the early activation of upcoming lexical material. Another possibility is that the time that it takes to resolve competition between frames in syntactically flexible sentences is used to retrieve in parallel words that are further ahead in the developing clause. This would explain why the frequency of N2 could have an effect on sentence onset latencies in syntactically flexible sentences.

One way of gaining more insight into sentence planning and formulation is to use eye-tracking measures. In Chapter 2, first fixations and eye-gaze patterns spanning the entire sentence planning and formulation process (i.e., before and after speech onset) were examined. In this study, the timecourse of eye-fixations was used as a marker differentiating linearly incremental or hierarchically incremental planning processes. Combining onset latencies with eye-gazes over time provides a way to infer the unique contributions of competition and planning processes in sentential onset latencies (see section on Syntactic flexibility in transitive sentences).

In Chapter 3 we sought to validate the competition-hypothesis by comparing speakers' susceptibility to syntactic flexibility with their performance on a flanker task measuring selective inhibitory control (SI) skills. Although results were not strong–there was a marginally significant trend suggesting that speakers with poor SI were delayed at initiating incongruent, syntactically flexible sentences compared to speakers with good SI–, the involvement of selective inhibition is suggestive of interference from competing representations during structure selection. The absence of a significant interaction between Flanker score and effects of syntactic flexibility may be due to insufficient power. Future research on the role of executive control in sentence production may test a larger sample of participants and add a more comprehensive assessment of SI skills.

A more direct (and efficient) method for identifying competition processes is to make use of event related potentials (ERPs). More specifically the N200 component of the ERP may be well-suited to index the selection of a sentence frame under syntactic flexibility (see Folstein & Van Petten, 2008, for a review). The N200 component, a negative wave peaking between 200 and 350 ms after stimulus onset, is known as a marker for inhibition processes when selecting a target response amongst competing representations in non-linguistic and linguistic tasks. In psycholinguistic research, the N200 component has been found to indicate inhibition processes during monolingual lexical selection (Shao, Roelofs, Acheson, & Meyer, 2014) and during bilingual language control and language switching (Christoffels, Firk, & Schiller, 2007; Hoshino & Thierry, 2011; Moreno, Rodríguez-Fornells, & Laine, 2008, for a review; Verhoef, Roelofs, & Chwilla, 2009), but it has never been used to examine syntactic selection processes within one language.

Another measure that could be added to exemplify the cognitive load associated with the resolution of competition of syntactic frames is pupillometry. Task evoked pupillary responses (TEPR) have been used in various domains as an indicator of cognitive effort (see Beatty, 1982, for a review). While differences in speech onsets may be dependent on shifts in sentence planning strategies, pupillary responses occur automatically and are therefore relatively independent of task strategies. Recently, pupillometry has been applied to tap cognitive load associated with syntactic complexity during sentence production (Sevilla, Maldonado, & Shalóm, 2014). In the study of Sevilla et al., participants had to describe pictures displaying transitive events, using sentence frames with varying syntactic complexity (i.e., actives, passives and clitic left dislocated [CLLD] sentences in Spanish) as signaled by a color cue indicating which frame to use. Results showed a higher increase in pupil size for syntactically complex sentences following a non-canonical thematic role order (i.e., passives and CLLD sentences) than for canonical counterparts (i.e., actives). Most interestingly, the same effect of syntactic complexity was found in pupil dilations when the color cue was presented prior to picture onset. This means that participants started structure building before reference could be made to lexical content. Hence, pupil size measures can be used to capture the generation of abstract structural frames of varying syntactic complexity. Similarly, TEPR may be sensitive to increased syntactic processing effort caused by the competition between syntactic frames. In fact, Sevilla et al. already introduce a competition view as an alternative account for their results: the pupil dilation during the selection of non-canonical sentence structures may be associated with the suppression of the more preferred canonical active structure. In a future study, the proposed competition account could be put to test by comparing TEPR (along with sentence onsets and gaze-patterns) during the production of syntactically flexible and inflexible sentences.

Syntactic flexibility in transitive sentences

In Chapter 3, the syntactic flexibility of to-be-produced sentences was varied by eliciting sentence frames with verbs that either had a strong or a weak bias towards one structural alternative. Another way of manipulating syntactic flexibility is to directly influence the activation level associated with one of the alternative syntactic frames. In Chapter 2, where speakers described pictures of transitive events, a passive syntax training was implemented half-way through Experiment 1. This training served to increase the ease of assembling a passive sentence structure, enabling a direct comparison between the production of a 'difficult sentence structure' (passives in part 1) and an 'easier to assemble structure' (passives in part 3, after the syntax training). At the same time, by cumulatively priming the passive sentence structure, passives presumably received a higher activation level (cf. Kaschak, 2007; Segaert, Menenti, Weber, & Hagoort, 2011). By implication, the activation level of the passive sentence structure became more equipotent to the well-practiced active structure, leading to increased syntactic flexibility. Critically, this change in activation levels should have led to an increase in the production of passives from part 1 to part 3 of Experiment 1. However, the effect of cumulative passive priming was only reflected in response tendencies for events with easy-to-name agents. Closer data inspection revealed that there was no overall increase in the production of passives, because participants showed large individual differences in susceptibility (as reflected in response tendencies) to the passive syntax training. Roughly half of the participants (n = 27) showed an increase in the proportion of passives produced from part 1 to part 3 of Experiment 1 (the 'trained' group), while the other half (n =21) did not show any change in their rate of producing passives or showed a decrease (the 'untrained' group). Hence, cumulatively priming the passive may have only led to an increase in the activation level of this sentence type for the 'trained' group of participants.

Thus taking a different perspective, the data from Experiment 1 can also be used to test predictions from a competition based model of grammatical encoding, as was done in Chapter 3. Following such a competition based account, selection times should increase when two structural alternatives share roughly the same activation level, delaying sentence production specifically in 'trained' participants. This hypothesis can be put to the test by examining more closely the onset latencies of passives in part 1 vs. part 3 of the experiment in 'trained' vs. 'untrained' participants. Since evaluating the influence of syntactic flexibility on production speed was not the original purpose of the research reported in Chapter 2, onset latencies were hitherto only analyzed as a complementary measure to the eye-gaze measures for inferring speakers' planning strategies.

A new linear mixed effects model predicting onset latencies in passives, yielded a significant effect of Experimental part in the 'untrained' group (β = -218.49, SE =86.10, t =-2.54) but not in the 'trained' group (t = -0.35). Figure 1 shows that in the 'untrained' group, speakers were significantly faster to initiate passives in part 3 than in part 1 of the experiment, whereas onset latencies of 'trained' speakers did not differ between experimental parts.[1]

[1] Interestingly, onset latencies followed a different pattern for active sentences: the 'trained' group initiated actives significantly faster in part 3 than in part 1 of the experiment (β = -149.44, SE =33.43, t =-4.47, whereas onset latencies for actives only decreased slightly (t = -1.43) in the 'untrained' group. Although there was no explicit priming manipulation promoting the active, this sentence structure was used so often throughout the experiment (on 79% of all scorable trials, collapsing across groups and experimental parts) that we can assume that practice effects would occur for this sentence type by the time participants reach part 3 of the experiment, leading to faster onset latencies. The fact that 'untrained' speakers did not show a significant practice effect in onset latencies for actives, may be caused by the fact that they were already very fast at producing this sentence type in part 1 of the experiment (onset latency: M = 1781, SD = 600 vs. M = 1933, SD = 602 in the 'trained' group). This 'ceiling effect' may have obscured facilitatory effects on active sentence formulation.

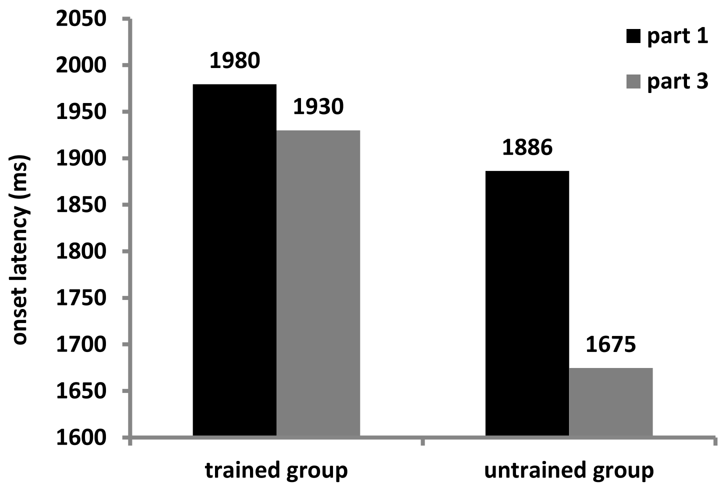

Figure 1. Onset latencies (ms) for passive sentences produced in part 1 vs. part 3 of Experiment 1 by 'trained' (primed) and 'untrained' participants.

This pattern of results is in line with findings from Segaert et al. (2011) who showed priming in response tendencies, but not in response latencies for passives. Segaert et al. explained these findings with a speed-selection trade-off. According to this account, there is a trade-off between the likelihood of selecting a particular structural alternative and the speed at which that selection is made. Priming a dispreferred structure (i.e., the passive) will facilitate execution speed but increases competition between the active and passive structures, which can result in slower selection of the dispreferred alternative. The inhibition effect may cancel out the facilitation in execution speed (due to a practice effect), leading to similar response latencies for passives in primed and unprimed conditions. Conversely, priming an active sentence structure only further increases the difference with the activation levels of the dispreferred passive structure, which may slightly shorten selection speed. Adding this up with the facilitatory practice effect on execution speed will lead to faster onset latencies for practiced (or primed) actives.

Similarly, in Experiment 1 of this thesis, 'trained' (i.e., primed) speakers may have been well practiced in executing structure assembly for the passive sentence structure, but the extra time needed to resolve competition during the selection stage may have cancelled out the facilitatory effect of practice on sentence initiation latency. Conversely, the 'untrained' group, although not showing an effect of cumulative priming of the passive in response tendencies, did get faster at initiating passives from part 1 to part 3 of the experiment. This finding suggests that this group may not have experienced as much syntactic competition when producing the passive in part 3 of the experiment as the 'trained' group. In the 'untrained' group, cumulative priming may have led to a slight increase in the activation level of the passive, but not to the extent that actives and passives became equipotent candidates. Instead, the cumulative priming manipulation only led to facilitation of execution speed for the passive.

Interestingly, the putative competition between sentence frames did not lead to a delay in the production of passives, but rather to an absence of a facilitatory practice effect on onset latencies. In Chapter 3, competition between frames led to a delay in the production of syntactically flexible ditransitives. The discrepancy between these findings may be due to the difference in imbalance between the sentence types under investigation. As for transitives, active and passive sentences structures exhibit a large difference in their frequency of occurrence (cf. Hartsuiker & Kolk, 1998). The dative alternation is more balanced; with double object datives (DO) being only slightly preferred over prepositional object (PO) datives in spoken language.

As in Chapter 3, it would be interesting to examine the role of selective inhibition (SI) in supporting competition resolution between syntactic frames. Based on the findings in Chapter 3, we would expect that speakers with good SI skills are faster at resolving competition between competing frames. Therefore, within the 'trained' group, differences in sentence onsets between part 1 and part 3 of Experiment 1 may be related to differences in SI skills: speakers with good SI skills may be fast at resolving competition during the selection stage of grammatical encoding, showing only a facilitatory practice effect on onset latencies.

Additionally, it would be worthwhile investigating why (nearly) half of the speakers in Experiment 1 in Chapter 1, did not show an effect of cumulative passive priming on response tendencies. In general, findings from priming studies indicate that maximum (100%) priming is hardly ever reached, regardless of whether preferred or dispreferred structures are primed (Pickering & Ferreira, 2008). This means that speakers continue to use both primed and unprimed structures. The extent to which speakers produce primed or unprimed structures (i.e., their susceptibility to priming) on target trials in priming experiments might be determined by individual-level variables. A possible candidate for this role is implicit learning. In various models, implicit learning of abstract structural representations has been appointed as the driving force behind structural priming (Bock & Griffin, 2000; Chang, Dell, & Bock, 2006; Chang, Dell, Bock, & Griffin, 2000). Until recently, implicit learning was not regarded as an ability with meaningful individual differences. However, Kaufman et al. (2010) used a probabilistic version of the serial reaction time task (SRT) to measure implicit learning and found that differences in this skill are related to foreign language acquisition, verbal analogical reasoning and processing speed. Other factors that might determine susceptibility to priming are working memory capacity, cognitive control and personality traits such as Extraversion and Neuroticism (Baving, Wagner, Cohen, & Rockstroh, 2001; Friederici, Schriefers, & Lindenberger, 1998; Kiefer, Ahlegian, & Spitzer, 2005). The latter have been hypothesized to determine the likelihood of priming in dialogue (i.e., alignment) specifically (Gill, Harrison, & Oberlander, 2004). Lastly, there may be other additional processing differences underlying the discrepancy in the results of the 'trained' and the 'untrained' group. Not only was the 'untrained' group less susceptible to the passive syntax training, speakers in this group were also much faster to initiate sentences in the picture description task (M = 1767, SD = 592) than the 'trained' group (M = 1883, SD = 604). Future research may examine the underlying factor linking these differences and evaluate to what extent this factor is related to task strategies specific to the picture description paradigm used here, or to sentence planning in general.

Methodological considerations: choosing good paradigms for sentence production research

A challenge in sentence production research is to find a paradigm for eliciting target-type (depending on the research question) sentences that is naturalistic on the one hand, but constrains options for the participant on the other hand. In this thesis, I have made use of a picture description paradigm (Chapter 2), a rapid-serial visual presentation (RSVP) paradigm (Chapters 3 and 4) and a written sentence completion task (Chapter 5). Whereas picture description and sentence completion tasks are used regularly in sentence production research, the use of the RSVP paradigm is rather novel (see also Chang, Bock, & Goldberg, 2003; Konopka & Bock, 2009; Tooley & Bock, 2014). In this paradigm, participants are presented with a sentence in a word-by-word fashion at a very high speed (100 ms per word); subsequently they perform a short distractor task, and then see a sentence preamble (e.g., the subject noun phrase) which they have to complete to form the presented sentence. It is assumed that the fast presentation of the sentence, with the distractor task intervening, leads to the formation of a conceptual representation of the sentence, which has to be 'rebuilt' during later recall (Potter & Lombardi, 1990). In this sense it resembles a natural production process in which a conceptual message needs to be translated to a linguistic sequence with linearly organized lexical units (e.g., Bock & Levelt, 1994).

There are also obvious differences between sentence recall via the RSVP paradigm and everyday sentence production. Most importantly, the participants do not generate the message based on the thoughts they wish to express, but instead read a sentence and store its content in working memory. Although it has been suggested that RSVP taps sentence reconstruction processes rather than episodic memory retrieval, in Chapters 3 and 4 sentence recall using this paradigm was often near verbatim. Speakers mostly produced sentences with the exact same structure and wording as the sentences presented to them. Especially in Chapter 4, where only congruent structure-verb bias pairings (e.g., PO preferring verbs in PO frames) and structure-noun frequency pairings (e.g., PO frames with a higher frequency direct than indirect object) were presented, speakers hardly 'flipped' the sentence's structure. Potter and Lombardi (1998) explain the verbatim recall with the fact that speakers are likely to reuse the recently activated (unordered) lexical entries and syntactic structures. Hence, during reconstruction in the recall phase, regular sentence production mechanisms are used to linearize lexical items retrieved from memory (Flores D'Arcais, 1974). Studies of the relationship between verbal working memory and language production are consistent with this view (Acheson & MacDonald, 2009a, 2009b; Jefferies, Lambon Ralph, & Baddeley, 2004).

Alternatively, one could argue that (some) speakers simply recall the sentence's surface structure from episodic memory: These speakers may have adopted a strategy of covertly repeating the presented sentence during the distractor task, such that they could retrieve the verbatim representation during later recall. In that sense, the recall task would resemble a delayed naming task only measuring the time needed to initiate the articulation of a production-ready string of words rather than the planning process itself (e.g., Balota & Chumbley, 1985).

There are several findings in Chapters 3 and 4 that counteract this alternative hypothesis. Firstly, patterns of errors suggest that participants did have difficulty in retaining a verbatim representation of the sentence material and instead stored conceptual representations. In both studies, the majority of the errors were substitutions of the target verb and/or a noun with often conceptually similar words, e.g., the use of 'entrepreneur' instead of 'buyer', such that the underlying sentence meaning remained unchanged.

Secondly, the effect of verb bias in both studies suggests that sentence initiation times in the RSVP paradigm are sensitive to 'normal' language processes responsible for the integration of verbs with their arguments. Namely, if sentence initiation times depended only on the time needed to initiate the articulation of the memorized verb (see e.g., Roelofs, 2002; Santiago, MacKay, & Palma, 2002 , for a discussion on the lower limit of articulation initiation), then no systematic effect of verb bias would be predicted, as verbs in weak and strong bias conditions were carefully matched on characteristics influencing their ease of accessibility (e.g., frequency, plausibility) and pronounceability (e.g., length and syllable count). Moreover, even in delayed naming (or pronunciation tasks), sentence initiation times were found to depend not only on characteristics of the initial word in a sentence, but on the sentence's number of phonological words and syntactic complexity–operationalized as the number of syntactic nodes (Ferreira, 1991)–, and on sentence structure (Savin & Perchonock, 1965). Together, these findings suggest that sentence memory is chunked according to an abstract syntactic hierarchy instead of a linearized list of words.

Thirdly, consistent processing differences between fast and slow speakers (as indexed by average onset latencies) were found across studies. Both in Chapter 3 and Chapter 4, slow speakers showed a larger advance planning scope than fast speakers. Interestingly, in both studies, speakers who initiated their utterance more slowly also showed slower overall speech rates. Memory research has revealed that speech rate is highly correlated to the amount of information that can be immediately recalled from memory (e.g., Baddeley, Thomson, & Buchanan, 1975; Smyth & Scholey, 1996). A memory account would thus predict that fast speakers have a longer recall list and hence may be able to produce the entire sentence from immediate memory. However, as verbal working memory is affected by lexical frequency (e.g., Roodenrys, Hulme, Lethbridge, Hinton, & Nimmo, 2002), the same memory account would also predict that fast speakers show an effect of N2 frequency in their onset latencies. Instead the N2 frequency effect in syntactically flexible sentences was carried mostly by the slow speakers in the tested sample (Experiment 2 in Chapter 4). Thus, the processing differences between fast and slow speakers do not seem to be attributable to working memory capacity. This conclusion is supported by findings from Wagner et al. (2010), who found that the planning scope of sentences with phrasal structures remained unchanged with varying memory load. Further examination is needed to investigate which processing differences underlie the difference in planning scope between slow and fast speakers. One possibility introduced in Chapter 4 is that slow speakers have better cognitive control skills than fast speakers, enabling them to engage in extensive advance planning which, in turn, is more cognitively demanding than piecemeal planning (Wagner, et al., 2010). Future research could evaluate this possibility by correlating slow and fast speakers' onset latencies in a sentence planning task to their performance on (a range of) tasks tapping cognitive control skills.

Taken together, findings from Chapters 3 and 4 constrain the role of episodic memory processes in sentence recall; although memorized content is retrieved upon recall, the ordering and unification of this content is grounded in everyday syntactic processes. Therefore, RSVP offers a viable way of studying the generation of sentences that are otherwise not easy to elicit. Whereas earlier studies used RSVP to study (constraints on) syntactic priming, the experiments in this thesis have shown that it can also be used to study sentence production in the absence of structural repetition.

Methodological considerations: caveats and improvements

Next to the challenge of finding a suitable language production paradigm for eliciting target-type sentences, there were more hurdles to be overcome when designing the experiments reported in this thesis. Here I will describe some possible shortcomings of the experimental designs and suggest improvements for future studies.

The first caveat concerns the use of post-hoc measures rather than experimental manipulations/orthogonal variations of factors in some of the reported experiments. In Experiment 1 of Chapter 2, the factors that were found most important in determining planning strategies, event codability and agent codability, were measured from post-hoc ratings in the experiment itself rather than varied a priori. This problem was addressed by running a second experiment in which items were allocated to four codability categories based on ratings obtained in four earlier experiments (based on the distribution of responses from different groups of subjects). Using pre-selected items varying on two codability dimensions and a simpler design (no Block and Cuing manipulation) results from Experiment 1 were fully replicated. However, in Chapters 3 and 4 we used a post hoc measure of production speed without validating it with an extra test. Although average production speed was based on onset latencies on neutral filler trials, it may still reflect paradigm-specific production speed rather than a general measure. It therefore prohibits further generalizations. Moreover, the experiments in both chapters were not set up to investigate individual differences in advanced planning: participants were all University students. Ideally, future experiments aiming to clarify the role of production speed in affecting advance planning scope, should test (a heterogeneous group of) participants on a range of tasks tapping their sentence production speed and cognitive performance in general.

Secondly, the use of a different sentence production paradigm (RSVP) in Chapter 3 prevents a direct comparison to the results of Ferreira (1996). Although there are several reasons to believe speakers' planning strategies or experimental materials in this study obscured possible competition effects, these reasons remain speculative. Ideally, two additional experiments, using the materials of Chapter 3, could be conducted in which the role of the paradigm in showing/obscuring effects of competition between syntactic frames during grammatical encoding would be clarified. Crucially, one experiment should be an exact replication of Ferreira's paradigm with Dutch dative materials, while the second experiment should use the same paradigm but without the timing bar manipulation (i.e., a quickly expiring timing bar that replaced the two-word preamble as soon as the participant started speaking). The latter has been argued to impose a linearly incremental planning strategy, which may make it impossible to observe effects of syntactic competition. In addition to these extra experiments, it would have been useful –for comparison to Ferreira's study and as an extreme case of syntactic inflexibility– if the materials in Chapter 3 had contained a non-alternating condition. In fact, we have tried to include this category of items, but unfortunately, in contrast to English, Dutch only has a few non-alternating ditransitive verbs and most of these are used in idiomatic expressions with fixed arguments (e.g., *zorgen baren* 'worry').

Finally, in some of the reported experiments sample size was not optimal. As noted before, the relatively small sample size in Chapter 3 may have caused the absence of interactions between experimental factors, such as Flanker score and Verb bias, due to limited power. Moreover, to take an individual differences approach, many more participants from a more heterogeneous group are actually needed. To further clarify the role of (other) specific subcomponents of executive control (EC), future experiments could also include an EC test battery instead of two measures of EC constructs (i.e., selective and nonselective inhibition).

In Chapter 5, the limited corpus size of both Alpino (written Dutch) and CGN (spoken Dutch) imposed certain restrictions on the way analyses were conducted. For example, it was impossible to directly compare OV and VO-based verb biases within one corpus of spoken language. Instead, VO-based biases were obtained from a written sentence completion task. The Alpino and CGN corpora were chosen because they are syntactically annotated, while other–sometimes larger–corpora are automatically parsed. Since automatic parsing is error-prone and time was too limited to check annotations manually, the two fully annotated corpora were chosen for analysis. It would be interesting to rerun the analyses in Chapter 5 on a larger corpus of spoken Dutch that allows for a within-corpus analysis of verb bias effects in OV and VO datives.

Contributions to language production models

Throughout this thesis the term *sentence structure* has been used to refer to the mental representation that symbolizes how constituents are related and ordered with respect to each other. Key focus in this thesis was whether this structure is linearly or hierarchically organized (Chapter 5), whether it engages in competition with alternative structures (Chapter 3) and whether the mapping to a *linear* surface structure is conceptually or lexically guided (Chapter

2). However, several questions remain about the nature of the sentence structure. For instance, is this structure abstract or dependent on lexical items? Does a syntactic structure have its own inherent meaning? And does the priming of frames reflect implicit learning or residual activation? Although the reported studies were not designed to answer these questions, we can infer some general features applying to the sentence production process from them.

Firstly, the verb bias effects in Chapter 3, 4 and 5 suggest that syntactic processing is lexicalist in nature. In other words, syntactic information is tied to individual lexical items and retrieved from the mental lexicon (Hagoort, 2005; Jackendoff, 2002; Vosse & Kempen, 2000). It is hard to reconcile verb bias effects with traditional accounts viewing the lexicon and the grammar of a language as qualitatively completely different, as these would not predict any dependencies between particular lexical items and particular syntactic structures. By measuring verb onset times instead of onset times at the syntactic choice point of a sentence (i.e., at the first object noun in datives), we implicitly assume that verb lemmas are directly bound to syntactic information. More specifically, we assume that upon retrieval of the verb lemma, subcategorization frames specifying the verb's arguments (or syntactic rules, see Pickering & Branigan, 1998) are activated to the degree specified by the verb's bias. This assumption is based on work by Melinger and Dobel (2005), who showed that the presentation of a non-alternator verb in isolation is sufficient to prime the use of its subcategorization frame during a subsequent production situation involving a syntactic choice. A replication of this experiment using verbs with different degrees of bias (instead of non-alternating verbs) would be necessary to directly test this assumption.

Relatedly, verb bias effects have been used as evidence for constructional meaning, the fact that sentence frames by themselves represent meaning, which is suggested by construction-based approaches (Stefanowitsch & Gries, 2003). According to construction-based approaches, syntactic alternates are not derived from the same underlying structure, but represent separate constructions. Sometimes, a syntactic alternate can be used specifically to convey a certain meaning. For example, according to a longstanding hypothesis, double object datives tend to be used to express 'caused possession', while prepositional object datives are associated with 'caused motion' (Goldberg, 1992). Verb preferences are seen as a reflection of the verbs' semantic compatibility with each of the two alternative frames: If a verb is better understood as caused possession than caused motion, it will have a DO preference (e.g., *refuse*). If a verb's semantics is compatible with both meanings, then it may not have a significant preference for either frame (leading to syntactic flexibility, see Chapters 3 and 4). Various studies have examined the distribution of verb preferences in corpora of spoken and written language to infer the inherent semantics of both dative alternates, i.e., the PO versus the DO dative (Colleman, 2009; Gries & Stefanowitsch, 2004). By grouping together verbs with strong preferences for one frame and deriving their joint semantics, constructional meaning could be inferred. In Chapter 5, we compared the outcome of the verb bias analysis of a spoken corpus of Dutch (CGN) with an analysis of a Dutch written corpus by Colleman (2009) to test the assumption that biases in written and spoken language are highly similar. The correlation analysis yielded a strong correspondence between verb biases based on spoken and written language. At the same time, this correlation provides converging evidence for the semantic factors driving the dative alternation that Colleman identified based on verb preferences. In short, the consistency in verb preferences across studies lends support for grammar formalisms that view sentence structures as meaningful units (e.g., Goldberg, 1995).

Finally, results from Chapters 3 and 4 can be interpreted in the light of theories specifying the mechanism underlying syntactic priming effects. These theories roughly fall into two main categories; one group of theories assumes that syntactic priming relies on error-based implicit learning (Chang, 2002; Chang, et al., 2006; Chang, Janciauskas, & Fitz, 2012), while the other group attributes priming effects to residual activation in recently processed representations (Malhotra, Pickering, Branigan, & Bednar, 2008; Pickering & Branigan, 1998)[2]. Only the former group of theories can explain *inverse preference effects* in priming: the effect that repetition of a *dispreferred* structure causes larger priming effects than the repetition of a *preferred* structural alternative (Ferreira & Bock, 2006). In error-based learning theories, encountering a less expected structure leads to a greater prediction error and thus to greater weight changes in the system, resulting in larger priming effects. Inverse preference effects have also been found in relation to verb preferences. If a prime sentence features a verb with an opposite structural preference, it is less expected and therefore priming effects will be larger than for prime sentences featuring a verb with congruent verb bias (Bernolet & Hartsuiker, 2010).

In the experiment reported in Chapter 3, participants had to reproduce sentences featuring verbs with preferences that were either congruent or incongruent with the structural configuration of the presented sentence. Results showed that speakers mostly reproduced the sentence structure as it was presented to them, even when there was a mismatch with the verb's preference. This finding was explained with the *inverse preference effect*: incongruent verb-structure pairings lead to a greater effect of surprisal than congruent pairings and

[2] For a hybrid account see Reitter, Keller, & Moore (2011).

therefore re-use of the presented structure is promoted. Typically, verb-structure congruence did not affect onset latencies in the same study.

In a recent study by Segaert et al. (2014) positive preference effects of verb-structure priming were found on response latencies: target sentences were initiated faster after primes with a congruent verb-structure pairing than after primes with an incongruent verb-structure pairing. Segaert et al. explained these findings with a speed-selection trade-off (see above). In short, this account predicts a trade-off between the likelihood of selecting a particular structural alternative and the speed at which that selection is made. When priming a preferred sentence structure–given the preference of the verb–priming further increases the difference in activation levels between the target structure and syntactic alternatives. As a result, the time needed to select a syntactic alternative decreases. At the same time, practice facilitates planning of the structure and the sentence can be initiated more quickly. When priming a dispreferred structure–given the verb's preference–sentence planning is facilitated but competition between the target structure and its syntactic alternative increases. Hence, the time needed to select a syntactic alternative increases and overrides the facilitatory practice effect on response latencies for dispreferred structures.

Although there was no effect of verb-structure congruence on onset latencies, the obtained verb bias effects in Chapters 3 and 4 could also be interpreted on the basis of the competition model put forward by Segaert et al (2014). Namely, a strong congruent verb bias may effectuate a larger difference in activation levels between the presented syntactic structure and its competitor than a weak congruent verb bias. Consequently, selection times for producing a sentence with a weak bias verb are longer than selection times for a sentence with a strong bias verb. In both situations, priming facilitates the execution of the speech plan, but because of the additive effect of verb preference on selection times, only in sentences with strong bias verbs will this lead to faster onset latencies.

6.4 Conclusions and implications

Within language research, the topic 'sentence production' is relatively understudied, yielding 13.300 hits on Google Scholar against, for example, 51.000 hits for 'lexical access'. Yet it is the most naturalistic level of language production: nearly all of our daily conversational speech consists of connected words. A lot more research is needed to unravel exactly how speakers go from a conceptual non-linear notion to a linearly structured sequence of words. This thesis only focused on two aspects of the sentence production process: incrementality and syntactic flexibility.

One aim was to further specify the linguistic conditions under which speakers make use of linearly and hierarchically incremental planning strategies. The reported experiments showed a joint influence of factors modulating the encoding of structural information (event structure and linguistic structure) on starting point selection and on the timecourse of formulation for active and passive sentences. Importantly, the influence of lower-level linguistic factors in promoting piecemeal planning was constrained to situations in which the concept-to-structure mapping was complicated. Together with the finding that the availability of multiple syntactic frames (i.e., syntactic flexibility) broadens planning scope during sentence recall, these results place important restrictions on how the message-to-language mapping proceeds. That is, higher-level structural factors seem to determine to what extent lexical items can enter speakers' planning scope and subsequently influence grammatical encoding processes. The finding that hierarchically proximate–while linearly distant–factors (i.e., a verb in final sentential position) can drive structure choices is in line with this conclusion.

Another goal was to investigate the influence of syntactic flexibility on grammatical encoding. The results of two of the reported studies show that syntactic flexibility delays the sentence production process and thereby provides support for competition-based models of grammatical encoding. Onsets for syntactically flexible sentences were specifically delayed in speakers with poor selective inhibitory control, suggesting that inhibitory control may be a skill needed for efficiently selecting a structural alternative among competitors.

Additionally, effects of syntactic flexibility were especially apparent in slow speakers, who have been shown–in previous studies–to engage more in extensive advance sentence planning than fast speakers. This latter finding implies that a broad scope of planning may be a prerequisite for syntactic flexibility to exert an effect: a narrow planning scope may not allow for the global weighing of syntactic choices. Vice versa, in this thesis, syntactic flexibility was shown to directly broaden the advance planning scope in dative verb phrases, possibly because the availability of multiple equipotent syntactic frames leads to the activation of all lexical material that can fill the post-verbal sentence slot.

Whereas this thesis was concerned with flexibility at the structural level, it is important to also address the influence of flexibility at different stages of the message-to-language mapping on language production. Numerous studies have investigated how speakers deal with flexibility at the lexical level and have found that executive control plays an important role in selecting a target representation among competitors (e.g., Shao, et al., 2014). However, how speakers deal with flexibility higher up in the sentence production hierarchy, e.g., with conceptual flexibility, remains unclear. Patient studies have revealed that unconstrained production situations generally lead to large production problems for speakers with impaired cognitive control skills (see Novick, Trueswell, & Thompson-Schill, 2010 for a review). Robinson et al. (2010; 2005) found that patients with left inferior frontal gyrus (LIFG; an area crucial for cognitive control functions) damage exhibited, among other things, an acute failure to complete sentences when multiple continuations (e.g., *the man entered the house and*...) were possible as compared to more constrained conditions (e.g.*, the man entered the cinema and*..). They concluded that the LIFG is crucial for selection from competing conceptual representations during sentence generation. Furthermore, they indicate that the reduced spontaneous speech characteristic of dynamic aphasia may be underpinned by the inability to select a verbal response when multiple competing alternatives are activated and no prepotent response is available. Future studies examining the role of cognitive control in dealing with production flexibility (at any level of the production process) in healthy participants could thus make a significant contribution to understanding the language - cognitive control connection and could support the development of therapies for situations where the connection is disrupted.

References

Acheson, D. J., & MacDonald, M. C. (2009a). Twisting tongues and memories: Explorations of the relationship between language production and verbal working memory. *Journal of Memory and Language, 60*(3), 329-350.

Acheson, D. J., & MacDonald, M. C. (2009b). Verbal working memory and language production: Common approaches to the serial ordering of verbal information. *Psychological Bulletin, 135*(1), 50.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Balota, D. A., & Chumbley, J. I. (1985). The locus of word-frequency effects in the pronunciation task: Lexical access and/or production? *Journal of Memory and Language, 24*(1), 89-106.

Baving, L., Wagner, M., Cohen, R., & Rockstroh, B. (2001). Increased semantic and repetition priming in schizophrenic patients. *Journal of Abnormal Psychology, 110*(1), 67-75.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91*(2), 276.

Bernolet, S., & Hartsuiker, R. J. (2010). Does verb bias modulate syntactic priming? [Research Support, Non-U.S. Gov't]. *Cognition, 114*(3), 455-461.

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

Bock, K., & Griffin, Z. M. (2000). The persistence of structural priming: Transient activation or implicit learning? [Article]. *Journal of Experimental Psychology-General, 129*(2), 177-192.

Chang, F. (2002). Symbolically speaking: a connectionist model of sentence production. [Article]. *Cognitive Science, 26*(5), 609-651.

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

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

Chang, F., Dell, G. S., Bock, K., & Griffin, Z. M. (2000). Structural priming as implicit learning: A comparison of models of sentence production. [Article; Proceedings Paper]. *Journal of Psycholinguistic Research, 29*(2), 217-229.

Chang, F., Janciauskas, M., & Fitz, H. (2012). Language adaptation and learning: Getting explicit about implicit learning. *Language and Linguistics Compass, 6*(5), 259-278.

Christoffels, I. K., Firk, C., & Schiller, N. O. (2007). Bilingual language control: An event-related brain potential study. *Brain research, 1147*, 192-208.

Colleman, T. (2009). Verb disposition in argument structure alternations: a corpus study of the dative alternation in Dutch. *Language Sciences, 31*(5), 593-611.

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, V. S. (1996). Is it better to give than to donate? Syntactic flexibility in language production. *Journal of Memory and Language, 35*(5), 724-755.

Ferreira, V. S., & Bock, K. (2006). The functions of structural priming. [Article]. *Language and Cognitive Processes, 21*(7-8), 1011-1029.

Flores D'Arcais, G. B. (1974). Is there a memory for sentences? *Acta Psychologica, 38*(1), 33-58.

Folstein, J. R., & Van Petten, C. (2008). Influence of cognitive control and mismatch on the N2 component of the ERP: a review. *Psychophysiology, 45*(1), 152-170.

Friederici, A. D., Schriefers, H., & Lindenberger, U. (1998). Differential age effects on semantic and syntactic priming. [Article]. *International Journal of Behavioral Development, 22*(4), 813-845.
Gill, A., Harrison, A., & Oberlander, J. (2004). *Interpersonality: Individual differences and interpersonal priming.* Paper presented at the Proceedings of the 26th annual conference of the cognitive science society.
Goldberg, A. E. (1992). The inherent semantics of argument structure: The case of the English ditransitive construction. *Cognitive Linguistics (includes Cognitive Linguistic Bibliography), 3*(1), 37-74.
Goldberg, A. E. (1995). *Construction grammar*: Wiley Online Library.
Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspective onalternations'. *International journal of corpus linguistics, 9*(1), 97-129.
Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science, 11*(4), 274-279.
Hagoort, P. (2005). On Broca, brain, and binding: a new framework. *Trends in Cognitive Sciences, 9*(9), 416-423.
Hartsuiker, R. J., & Kolk, H. H. J. (1998). Syntactic persistence in Dutch. [Article]. *Language and Speech, 41*, 143-184.
Hoshino, N., & Thierry, G. (2011). Language selection in bilingual word production: electrophysiological evidence for cross-language competition. *Brain research, 1371*, 100-109.
Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution*: Oxford University Press.
Jefferies, E., Lambon Ralph, M. A., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language, 51*(4), 623-643.
Kaschak, M. P. (2007). Long-term structural priming affects subsequent patterns of language production. [Article]. *Memory & Cognition, 35*(5), 925-937.
Kaufman, S. B., Deyoung, C. G., Gray, J. R., Jimenez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition, 116*(3), 321-340.
Kiefer, M., Ahlegian, M., & Spitzer, M. (2005). Working memory capacity, indirect semantic priming, and stroop interference: Pattern of interindividual prefrontal performance differences in healthy volunteers. *Neuropsychology, 19*(3), 332-344.
Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology, 58*(1), 68-101.
Kuchinsky, S. E., & Bock, J. K. (2010). *From seeing to saying: Perceiving, planning, producing.* Paper presented at the 23rd meeting of the CUNY human sentence processing conference.
Malhotra, G., Pickering, M., Branigan, H., & Bednar, J. A. (2008). *On the persistence of structural priming: Mechanisms of decay and influence of word-forms.* Paper presented at the Proceedings of the 30th annual conference of the cognitive science society.
Melinger, A., & Dobel, C. (2005). Lexically-driven syntactic priming. [Research Support, Non-U.S. Gov't]. *Cognition, 98*(1), B11-20.
Meyer, A. S., Sleiderink, A. M., & Levelt, W. J. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition, 66*(2), B25-B33.
Moreno, E. M., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics, 21*(6), 477-508.

Myachykov, A., Scheepers, C., Garrod, S., Thompson, D., & Fedorova, O. (2013). Syntactic flexibility and competition in sentence production: the case of English and Russian. [Research Support, Non-U.S. Gov't]. *Quarterly Journal of Experimental Psychology, 66*(8), 1601-1619.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2010). Broca's area and language processing: Evidence for the cognitive control connection. *Language and Linguistics Compass, 4*(10), 906-924.

Pickering, M. J., & Branigan, H. P. (1998). The representation of verbs: Evidence from syntactic priming in language production. *Journal of Memory and Language, 39*(4), 633-651.

Potter, M. C., & Lombardi, L. (1990). Regeneration in the short-term recall of sentences. *Journal of Memory and Language, 29*(6), 633-654.

Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science, 35*(4), 587-637.

Robinson, G., Shallice, T., Bozzali, M., & Cipolotti, L. (2010). Conceptual proposition selection and the LIFG: Neuropsychological evidence from a focal frontal group. *Neuropsychologia, 48*(6), 1652-1663.

Robinson, G., Shallice, T., & Cipolotti, L. (2005). A failure of high level verbal response selection in progressive dynamic aphasia. *Cognitive neuropsychology, 22*(6), 661-694.

Roelofs, A. (2002). Spoken language planning and the initiation of articulation. *The Quarterly Journal of Experimental Psychology: Section A, 55*(2), 465-483.

Roodenrys, S., Hulme, C., Lethbridge, A., Hinton, M., & Nimmo, L. M. (2002). Word-frequency and phonological-neighborhood effects on verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(6), 1019.

Santiago, J., MacKay, D. G., & Palma, A. (2002). Length effects turn out to be syllable structure effects: Response to Roelofs (2002). *Language and Cognitive Processes, 17*(1), 15-29.

Savin, H. B., & Perchonock, E. (1965). Grammatical structure and the immediate recall of English sentences. *Journal of Verbal Learning and Verbal Behavior, 4*(5), 348-353.

Segaert, K., Menenti, L., Weber, K., & Hagoort, P. (2011). A paradox of syntactic priming: why response tendencies show priming for passives, and response latencies show priming for actives. [Research Support, Non-U.S. Gov't]. *Plos One, 6*(10), e24209.

Segaert, K., Weber, K., Cladder-Micus, M., & Hagoort, P. (2014). The influence of verb-bound syntactic preferences on the processing of syntactic structures. *Journal of experimental psychology. Learning, memory, and cognition*.

Sevilla, Y., Maldonado, M., & Shalóm, D. E. (2014). Pupillary dynamics reveal computational cost in sentence planning. *The Quarterly Journal of Experimental Psychology*(just-accepted), 1-24.

Shao, Z., Roelofs, A., Acheson, D. J., & Meyer, A. S. (2014). Electrophysiological Evidence that Inhibition Supports Lexical Selection in Picture Naming. *Submitted for publication*.

Smyth, M. M., & Scholey, K. A. (1996). The relationship between articulation time and memory performance in verbal and visuospatial tasks. *British Journal of Psychology, 87*(2), 179-191.

Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction of words and constructions. *International journal of corpus linguistics, 8*(2), 209-243.

Tooley, K. M., & Bock, K. (2014). On the parity of structural persistence in language production and comprehension. *Cognition, 132*(2), 101-136.

Verhoef, K., Roelofs, A., & Chwilla, D. J. (2009). Role of inhibition in language switching: Evidence from event-related brain potentials in overt picture naming. *Cognition, 110*(1), 84-99.

Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. *Cognition, 75*(2), 105-143.

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.

NEDERLANDSE SAMENVATTING

Als we communiceren dan doen we dat meestal niet met losse woorden, maar in volledige zinnen. Voor de productie van een zin coördineert de spreker verschillende processen. Zo dienen de woorden, die de intentie van de spreker omvatten, opgehaald te worden uit het lange termijn geheugen. Deze woorden worden gecombineerd tot een lineaire reeks die voldoet aan de eisen die de grammatica van de gesproken taal stelt. Vaak zijn er meerdere mogelijkheden om woorden te combineren tot een zin. In mijn proefschrift noem ik dit 'syntactische flexibiliteit'.

Toch – met zoveel beperkingen aan de ene en vrijheden aan de andere kant – is de gemiddelde spreeksnelheid bij het produceren van een zin erg hoog; Nederlanders spreken gemiddeld 4,23 lettergrepen uit per seconde. Daarom is het onvermijdelijk dat een deel van het spreekplan wordt voorbereid, voordat men begint met het uitspreken van een zin. In dit proefschrift heb ik onderzocht hoeveel van het spreekplan wordt voorbereid (d.w.z. de mate van incrementaliteit) onder verschillende omstandigheden en hoe sprekers omgaan met syntactische flexibiliteit. Hieronder worden enkele bevindingen kort aangestipt.

Eerder onderzoek heeft aangetoond dat sprekers hun zinnen op twee manieren kunnen plannen: lineair of hiërarchisch incrementeel. In het eerste geval begint de spreker met praten zodra het eerste woord van de zin opgehaald is uit het geheugen. De verdere zin wordt vervolgens stukje bij beetje gepland, waarbij de vorm bepaald wordt door de relatieve toegankelijkheid van de afzonderlijke woorden. Bijvoorbeeld, als de spreker een gebeurtenis met een postbode en een rennende hond beschrijft en het woord *postbode* als eerste wordt opgehaald, dan leidt dit automatisch tot de productie van een passieve zinsstructuur *de postbode wordt achternagezeten door een hond.* Bij hiërarchisch incrementeel plannen, construeert de spreker eerst een hiërarchisch, overkoepelend plan voor de zin en vult de woorden pas later in.

Hoofdstuk 2 onderzocht hoe perceptuele en linguïstische factoren van verschillende niveaus in het taalproductieproces, invloed uitoefenen op de manier waarop sprekers hun zinnen plannen. In twee experimenten beschreven proefpersonen een lange reeks plaatjes met actieve of passieve zinsstructuren, terwijl ondertussen hun oogbewegingen werden gemeten. Oogbewegingen geven namelijk een goed beeld van de processen die omgaan in de spreker, voorafgaand aan de productie van een zin. Zo richten sprekers, als ze een plaatje beschrijven, hun blik op de afgebeelde personages in de volgorde waarin ze deze later noemen en verleggen ze hun blik als ze het woord voor het gefixeerde personage hebben opgehaald.

De plaatjes in de twee experimenten verschilden in de mate van benoembaarheid (‘codability’) van de afgebeelde acties en personages. Daarnaast werden proefpersonen in Experiment 1 onbewust gecued om naar een bepaald personage te kijken en kregen ze een training, halverwege het experiment, om de productie van de passieve zinsstructuur te vergemakkelijken. Verwacht werd dat proefpersonen van een *lineair* naar een *hiërarchisch* incrementele planning strategie zouden switchen naarmate het makkelijker werd om de over-te-brengen boodschap te structureren (d.w.z. bij het beschrijven van makkelijk benoembare acties en bij het gebruik van een geoefende zinsstructuur). Uit de resultaten blijkt dat sprekers alleen overgingen tot lineair incrementeel plannen als de afgebeelde actie moeilijk benoembaar was. De bevindingen bevestigen eerder onderzoek dat laat zien dat factoren hoger in het taalproductieproces de invloed van factoren van lager niveau beperken.

Hoofdstuk 3 onderzocht hoe sprekers syntactische keuzes maken en of cognitieve controle (specifiek: selectieve inhibitie) hierbij een rol speelt. De hoofdvraag was of syntactische alternatieven de competitie aangaan, wanneer er een zinsstructuur geselecteerd wordt. Ik heb dit onderzocht door proefpersonen datieve zinnen (bijv. *De ober serveert de klant de maaltijd*) te laten produceren, die varieerden in syntactische flexibiliteit. Deze zinnen werden ontlokt met behulp van het Rapid Serial Visual Presentation (RSVP) paradigma. In dit paradigma krijgt de proefpersoon een zin te zien die woord voor woord, op hoge snelheid (100 ms per woord), wordt getoond. Daarna volgt er een afleidingstaak; de proefpersoon beoordeelt of een getal deel uitmaakte van een daarvoor getoonde reeks. Op vertoon van een preambule (*De student..*) dient de proefpersoon vervolgens de zin aan te vullen met hetgeen hij daarvoor heeft gelezen. Door de hoge presentatiesnelheid en de afleidingstaak, wordt verondersteld dat de proefpersoon op dat moment slechts een conceptuele representatie van de zin paraat heeft en dat de zin dus opnieuw (zoals in een natuurlijk spraakproces) opgebouwd moet worden.

Syntactische flexibiliteit werd gevarieerd door proefpersonen zinnen te laten produceren met werkwoorden die een sterke voorkeur hebben voor een syntactisch alternatief (lage syntactische flexibiliteit, bijv. *voorschotelen*) of een zwakke, niet significante voorkeur hebben voor een alternatief (hoge syntactische flexibiliteit, bijv. *serveren*). Theorieën die uitgaan van competitie tussen syntactische alternatieven voorspellen dat syntactische flexibiliteit zinsproductie vertraagt. In lijn met deze voorspelling, initieerden proefpersonen sneller de productie van zinnen met werkwoorden met een sterke syntactische voorkeur dan met een zwakke voorkeur. Daarnaast waren proefpersonen met een goede cognitieve controle sneller in het produceren van zinnen onder syntactische flexibiliteit dan proefpersonen met een slechtere cognitieve controle. De resultaten bevestigen dat gelijkwaardige syntactische alternatieven onderling concurreren om geselecteerd te worden en dat cognitieve controle kan helpen om deze competitie efficiënt op te lossen.

Hoofdstuk 4 combineerde de thema's incrementaliteit en syntactische flexibiliteit door te onderzoeken of de beschikbaarheid van meerdere gelijkwaardige syntactische alternatieven de mate van vooruitplannen tijdens zinsproductie beïnvloedt. Door opnieuw gebruik te maken van het RSVP paradigma, liet ik sprekers datieve zinnen met een hoge en lage syntactische flexibiliteit produceren. Daarnaast werd binnen deze zinnen de frequentie van het eerste (Experiment 1, zie 1a en 1b) of tweede (Experiment 2, zie 2a en 2b) zelfstandig naamwoord (ZNW) gemanipuleerd.

1a) *De ober serveert/schotelt* ***de monarch*** *het feestmaal voor.* (laagfrequent ZNW 1)

1b) *De ober serveert/schotelt* ***de koning*** *het feestmaal voor.* (hoogfrequent ZNW 1)

2a) *De ober serveert/schotelt de klant* ***het feestmaal*** *voor.* (laagfrequent ZNW 2)

2b) *De ober serveert/schotelt de klant* ***de maaltijd*** *voor.* (hoogfrequent ZNW 2)

Als de frequentie van het eerste ZNW invloed zou uitoefenen op de snelheid waarmee sprekers zinsproductie initiëren, dan geeft dit aan dat het eerste ZNW zich in de planning scope bevindt. Een effect van ZNW2 op de reactiesnelheid van proefpersonen zou aangeven dat ook dit ZNW uitmaakt van de planning scope. Resultaten lieten een effect zien van de frequentie van het tweede ZNW, alleen in zinnen met hoge syntactische flexibiliteit. Deze bevindingen suggereren dat syntactische flexibiliteit de mate van vooruitplannen kan beïnvloeden en onze lexicale planning scope kan verbreden tot en met het tweede ZNW in een datieve werkwoordzin.

Hoofdstuk 5 maakte gebruik van een corpusstudie om te onderzoeken of de onderdelen in de planning scope van een spreker lineair of hiërarchisch georganiseerd zijn. Ik heb hiervoor gekeken naar de invloed van werkwoordsvoorkeuren op structuurkeuzes in datieve zinnen met het werkwoord aan het begin (SVO volgorde) of aan het einde (SOV volgorde). Als de planning scope van een spreker hiërarchisch georganiseerd is en het werkwoord, bovenaan de hiërarchie, als eerst gepland wordt, dan zal deze altijd (in SVO en SOV zinnen) zijn invloed kunnen uitoefenen op structuurkeuzes. In andere woorden, een werkwoord met een sterke voorkeur voor een aan-datief zal, ook aan het eind van de bijzin, een significante invloed kunnen uitoefenen op de uiteindelijk geproduceerde zinsstructuur (bijvoorbeeld: *de student, [die het plan aan de professor voorlegt], zit in haar derde jaar*). Als de planning scope van een spreker daarentegen lineair georganiseerd is, dan kunnen werkwoordsvoorkeuren alleen hun invloed uitoefenen als het werkwoord vooraan staat in de zin (SVO volgorde). Een analyse van het Corpus Gesproken Nederlands (CGN) liet zien dat werkwoordsvoorkeuren een significante voorspeller zijn van structuurkeuzes in zowel SVO als SOV zinnen, maar met een zwakker effect in zinnen met laatstgenoemde woordvolgorde. De bevindingen suggereren dat het plannen van zinnen tijdens spontane spraak vooral op hiërarchisch incrementele wijze gebeurt.

ACKNOWLEDGEMENTS

The cover of this thesis not only presents an analog to one of the questions of my PhD project, it also reflects part of the PhD process; at times I couldn't see the wood for the trees. Fortunately, I had many great people by my side who helped to keep me on track.

First of all, I would like to thank my supervisors Antje Meyer and Agnieszka Konopka. Antje, thank you for your guidance throughout my project. I appreciate that you were always available for questions and that you provided feedback on my writings, mostly, within a day. You also helped me with not getting lost in the big forest of research questions and possible analyses. Agnieszka, thank you for putting so much effort and energy in the projects that we have carried out together. I learned a lot from your lessons on R and Excel.

Thank you Gerard Kempen for your always kind and patient advice, for your genuine interest in my research and personal life, and for your enthusiasm about our project. I very much enjoyed our collaboration, together with Karin who introduced me to corpus research.

I am grateful to Ardi Roelofs, Peter Hagoort, and Linda Wheeldon for reading and evaluating this thesis. Linda Wheeldon, also thank you for your feedback along the way during your stays at the MPI.

I would also like to thank Fenna Poletiek, with whom I now work on various research and teaching projects at Leiden University. Fenna, thank you for offering me the opportunity to work with your students and for the trust and freedom that you have always given me.

Svetlana and Annelies, I am very happy to have you by my side as my paranymphs. Svetlana, I've spent most time of my PhD project in one office with you and I am very happy about that. We have shared many ups and downs and you were the first to notice one major 'up': my pregnancy of Nora. Your calmness and warm personality were of great help in times of despair. I am very grateful to have you as a friend and to share so many of life's milestones. Annelies, you brought back the music in my PhD life! Thank you for your always positive energy, your groundedness and our chats about, primarily, non-research stuff. I really enjoyed our secret (or not so secret anymore..) singing sessions in the basement.

Thank you to my other colleagues at the MPI. A special thanks goes out to my former office mates. Florian and Jannie, my Beyoncés, I still have to laugh when I hear the song by Cee Lo Green (you know which one!). Francesca, I cannot sum up the funny and crazy moments we shared, because there are too many. Thank you for supporting me and teaching me many things about (PhD) life. Alma, we shared the Party Office and the last eventful phase of our PhD project. I thank you for making this last bit of my PhD project a good one. Joost, we have never shared an office, but it surely felt that way. Thank you for having geeky statistics chats and for your always thoughtful answers to my sometimes silly questions. Matthias, thanks for making the many trips to and from Nijmegen so nice! Katrien, I am very happy that we met, although quite late in my PhD project. We do not only share a research interest, but turned out to have many other common interests and children that like to do baby-yoga together. IMPRS workshop team, thank you for an inspiring and instructive experience; I really enjoyed organizing the Linguistic Relativity workshop with you.

For all administrative struggles, Evelyn was there to help. Evelyn, thank you for always being so kind and interested! I also want to thank the TG for their help with technical challenges. A huge thank you to IMPRS office; Els, Dirkje and Rachel, you were always there for a listening ear and helped me with many non-scientific issues. You have been of great support! The same holds for all the assistants of the PoL department who have helped me annotate endless speech material. Thank you!

Lieve vrienden, ik voel me zo gelukkig dat jullie er nog steeds zijn, ondanks mijn proefschrift en nu baby-isolement. Ik beloof dat ik jullie (voorlopig) niet meer lastig zal vallen met zinnetjes om te normeren (C&M, jullie zijn veilig!) en dat ik weer meer tijd ga vrijmaken. Hen, Rens en Pan, bedankt voor de vele bezoekjes, steun en de (stoofpeertaxi) lol die we altijd hebben. Vilma, ik ben zo blij dat je er weer bent! Anna, bedankt voor je immer positieve instelling en inspirerende creativiteit. Nelise, je bent een schat, bedankt dat je altijd aan me denkt. Caduten, jullie wil ik bedanken voor de altijd hilarische afleiding, vooral mijn jaargenoten, Inge, Dieneke en Lisa. Algon, bedankt voor gezellige schildersessies en het uitwisselen van babyperikelen en –kleding. Clofo’s, dankzij jullie ga ik zingend door ‘t leven!

En dan als laatste, de personen die het allermeest voor me betekenen. Lieve pap en mam, bedankt voor jullie steun nu en vroeger en voor een warm en liefdevol thuis. Mam, bedankt ook voor je hulp bij de cover van mijn dissertatie. Ook Stefan, Suzanne, Sharon en mijn lieve schoonfamilie, bedankt voor de interesse – ook al was het soms onduidelijk waar ik precies mee bezig was. Een speciaal dankjewel voor Koen, die mijn proefschrift al twee keer gered heeft uit grote computerellende. Ten slotte wil ik ook mijn gezin (ja, dat is het nu echt!) bedanken. Lieve Wouter, zonder jou zou dit boekje hier nu niet liggen. Naast jou voel ik me alsof ik de hele wereld aankan. Dank je voor je liefde en steun en voor onze twee scheetjes van kinderen; Nora en Abel. Jullie zijn mijn alles!

Curriculum Vitae

Maartje van de Velde was born in Vlagtwedde, the Netherlands. After obtaining her Gymnasium diploma in 2003, she studied Psychology at the University of Leiden with a minor in Design at the Royal Academy of Art in The Hague. In 2007 she obtained her BA degree and continued to do a research MA in Cognitive Neuroscience at the University of Leiden. She completed this research MA cum laude in 2009 with a thesis on lexical inhibition in bilinguals under the supervision of Maria Teresa Bajo (University of Granada) and Wido La Heij. In 2010, after some months of traveling, she started her PhD project in the Psychology of Language department at the Max Planck Institute for Psycholinguistics under the supervision of Antje Meyer and Agnieszka Konopka. The results of her PhD research are described in this thesis. She is currently working as education coordinator of the Wageningen School of Social Sciences and as lecturer in Cognitive Psychology at Leiden University.

PUBLICATIONS

Poletiek, F.H., Van de Velde, M., *Hearing true speech helps learning a language with hierarchical structure: The effect of reflecting statistical characteristics of the reference world in language, on learning its grammar.* Manuscript in preparation.

Van de Velde, M., Kempen, G., & Harbusch K. (under revision). Dative alternation and planning scope in spoken language: A corpus study on effects of verb bias in VO and OV clauses of Dutch.

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.

MPI Series in Psycholinguistics

1. The electrophysiology of speaking : investigations on the time course of semantic, syntactic, and phonological processing. *Miranda van Turennout*
2. The role of the syllable in speech production : evidence from lexical statistics, metalinguistics, masked priming, and electromagnetic midsagittal articulography. *Niels Schiller*
3. Lexical access in the production of ellipsis and pronouns. *Bernadette Schmitt*
4. The open-/closed class distinction in spoken-word recognition. *Alette Haveman*
5. The acquisition of phonetic categories in young infants : a self-organising artificial neural network approach. *Kay Behnke*
6. Gesture and speech production. *Jan-Peter de Ruiter*
7. Comparative intonational phonology : English and German. *Esther Grabe*
8. Finiteness in adult and child German. *Ingeborg Lasser*
9. Language input for word discovery. *Joost van de Weijer*
10. Inherent complement verbs revisited : towards an understanding of argument structure in Ewe. *James Essegbey*
11. Producing past and plural inflections. *Dirk Janssen*
12. Valence and transitivity in Saliba : an Oceanic language of Papua New Guinea. *Anna Margetts*
13. From speech to words. *Arie van der Lugt*
14. Simple and complex verbs in Jaminjung : a study of event categorisation in an Australian language. *Eva Schultze-Berndt*
15. Interpreting indefinites : an experimental study of children's language comprehension. *Irene Krämer*
16. Language-specific listening : the case of phonetic sequences. *Andrea Weber*
17. Moving eyes and naming objects. *Femke van der Meulen*
18. Analogy in morphology : the selection of linking elements in dutch compounds. *Andrea Krott*
19. Morphology in speech comprehension. *Kerstin Mauth*
20. Morphological families in the mental lexicon. *Nivja de Jong*
21. Fixed expressions and the production of idioms. *Simone Sprenger*
22. The grammatical coding of postural semantics in Goemai (a West Chadic language of Nigeria). *Birgit Hellwig*
23. Paradigmatic structures in morphological processing : computational and cross-linguistic experimental studies. *Fermín Moscoso del Prado Martín*
24. Contextual influences on spoken-word processing : an electrophysiological approach. *Daniëlle van den Brink*
25. Perceptual relevance of prevoicing in Dutch. *Petra van Alphen*
26. Syllables in speech production : effects of syllable preparation and syllable frequency. *Joana Cholin*
27. Producing complex spoken numerals for time and space. *Marjolein Meeuwissen*

28. Morphology in auditory lexical processing : sensitivity to fine phonetic detail and insensitivity to suffix reduction. *Rachèl Kemps*
29. At the same time.. : the expression of simultaneity in learner varieties. *Barbara Schmiedtová*
30. A grammar of Jalonke argument structure. *Friederike Lüpke*
31. Agrammatic comprehension : an electrophysiological approach. *Marlies Wassenaar*
32. The structure and use of shape-based noun classes in Miraña (North West Amazon). *Frank Seifart*
33. Prosodically-conditioned detail in the recognition of spoken words. *Anne Pier Salverda*
34. Phonetic and lexical processing in a second language. *Mirjam Broersma*
35. Retrieving semantic and syntactic word properties : ERP studies on the time course in language comprehension. *Oliver Müller*
36. Lexically-guided perceptual learning in speech processing. *Frank Eisner*
37. Sensitivity to detailed acoustic information in word recognition. *Keren Shatzman*
38. The relationship between spoken word production and comprehension. *Rebecca Özdemir*
39. Disfluency : interrupting speech and gesture. *Mandana Seyfeddinipur*
40. The acquisition of phonological structure : distinguishing contrastive from non-constrative variation. *Christiane Dietrich*
41. Cognitive cladistics and the relativity of spatial cognition. *Daniel Haun*
42. The acquisition of auditory categories. *Martijn Goudbeek*
43. Affix reduction in spoken Dutch : probabilistic effects in production and perception. *Mark Plumaekers*
44. Continuous-speech segmentation at the beginning of language acquisition : Electrophysiological evidence. *Valesca Kooijman*
45. Space and iconicity in German sign language (DGS). *Pamela Perniss*
46. On the production of morphologically complex words with special attention to effects of frequency. *Heidrun Bien*
47. Crosslinguistic influence in first and second languages : convergence in speech and gesture. *Amanda Brown*
48. The acquisition of verb compounding in Mandarin Chinese. *Jidong Chen*
49. Phoneme inventories and patterns of speech sound perception. *Anita Wagner*
50. Lexical processing of morphologically complex words : an information-theoretical perspective. *Victor Kuperman*
51. A grammar of Savosavo : a Papuan language of the Solomon Islands. *Claudia Wegener*
52. Prosodic structure in speech production and perception. *Claudia Kuzla*
53. The acquisition of finiteness by Turkish learners of German and Turkish learners of French: investigating knowledge of forms and functions in production and comprehension. *Sarah Schimke*
54. Studies on intonation and information structure in child and adult German. *Laura de Ruiter*
55. Processing the fine temporal structure of spoken words. *Eva Reinisch*
56. Semantics and (ir)regular inflection in morphological processing. *Wieke Tabak*
57. Processing strongly reduced forms in casual speech. *Susanne Brouwer*
58. Ambiguous pronoun resolution in L1 and L2 German and Dutch. *Miriam Ellert*

59. Lexical interactions in non-native speech comprehension : evidence from electro-encephalography, eye-tracking, and functional magnetic resonance imaging. *Ian FitzPatrick*
60. Processing casual speech in native and non-native language. *Annelie Tuinman*
61. Split intransitivity in Rotokas, a Papuan language of Bougainville. *Stuart Robinson*
62. Evidentiality and intersubjectivity in Yurakaré : an interactional account. *Sonja Gipper*
63. The influence of information structure on language comprehension : a neurocognitive perspective. *Lin Wang*
64. The meaning and use of ideophones in Siwu. *Mark Dingemanse*
65. The role of acoustic detail and context in the comprehension of reduced pronunciation variants. *Marco van de Ven*
66. Speech reduction in spontaneous French and Spanish. *Francisco Torreira*
67. The relevance of early word recognition : insights from the infant brain. *Caroline Junge*
68. Adjusting to different speakers : extrinsic normalization in vowel perception. *Matthias Sjerps*
69. Structuring language : contributions to the neurocognition of syntax. *Katrien Segaert*
70. Infants' appreciation of others' mental states in prelinguistic communication : a second person approach to mindreading. *Birgit Knudsen*
71. Gaze behavior in face-to-face interaction. *Federico Rossano*
72. Sign-spatiality in Kata Kolok: how a village sign language of Bali inscribes its signing space. *Connie de Vos*
73. Who is talking? Behavioural and neural evidence for norm-based coding in voice identity learning. *Attila Andics*
74. Lexical processing of foreign-accented speech: Rapid and flexible adaptation. *Marijt Witteman*
75. The use of deictic versus representational gestures in infancy. *Daniel Puccini*
76. Territories of knowledge in Japanese conversation. *Kaoru Hayano*
77. Family and neighbourhood relations in the mental lexicon: A cross-language perspective. *Kimberley Mulder*
78. Contributions of executive control to individual differences in word production. *Zeshu Shao*
79. Hearing speech and seeing speech: Perceptual adjustments in auditory-visual processing. *Patrick van der Zande*
80. High pitches and thick voices. The role of language in space-pitch associations. *Sarah Dolscheid*
81. Seeing what's next: Processing and anticipating language referring to objects. *Joost Rommers*
82. Mental representation and processing of reduced words in casual speech. *Iris Hanique*
83. The many ways listeners adapt to reductions in casual speech. *Katja Poellmann*
84. Contrasting opposite polarity in Germanic and Romance languages: Verum Focus and affirmative particles in native speakers and advanced L2 learners. *Giuseppina Turco*
85. Morphological processing in younger and older people: Evidence for flexible dual-route access. *Jana Reifegerste*

86. Semantic and syntactic constraints on the production of subject-verb agreement. *Alma Veenstra*
87. The acquisition of morphophonological alternations across languages. *Helen Buckler*
88. The evolutionary dynamics of motion event encoding. *Annemarie Verkerk*
89. Rediscovering a forgotten language. *Jiyoun Choi*
90. The road to native listening: Language-general perception, language-specific input. *Sho Tsuji*
91. Infants'understanding of communication as participants and observers. *Gudmundur Thorgrímsson*
92. Information structure in Avatime. *Saskia van Putten*
93. Switch reference in Whitesands. *Jeremy Hammond*
94. Machine learning for gesture recognition from videos. *Binyam Gebrekidan Gebre*
95. Acquisition of spatial language by signing and speaking children: a comparison of Turkish sign language (TID) and Turkish. *Beyza Sumer*
96. An ear for pitch: on the effects of experience and aptitude in processing pitch in language and music. *Salomi Asaridou*
97. Incrementality and flexibility in sentence production. *Maartje van de Velde*