



Message formulation and structural assembly: Describing “easy” and “hard” events with preferred and dispreferred syntactic structures

Maartje van de Velde^{a,b}, Antje S. Meyer^{a,c}, Agnieszka E. Konopka^{a,c,*}

^a Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

^b International Max Planck Research School for Language Sciences, P.O. Box 310, 6500 AH Nijmegen, The Netherlands

^c Donders Institute for Brain, Cognition and Behavior, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 4 June 2012

revision received 30 October 2013

Keywords:

Timecourse of sentence formulation
Incrementality
Structural processing

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.

© 2013 Elsevier Inc. All rights reserved.

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 (Levitt, 1989). Such choices can be influenced by long-term biases, such as a general preference to use 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 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 charac-

* Corresponding author. Address: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.

E-mail address: agnieszka.konopka@mpi.nl (A.E. Konopka).

ters 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 to begin their sentences with this character rather than the uncued character (Gleitman et al., 2007; Kuchinsky & Bock, 2010; Myachykov, Tomlin, & Posner, 2005; Tomlin, 1997). 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



Fig. 1. Example of a target item. The modal active description of this event is “The cat is catching the mouse.”

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. First, 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. Second, 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.

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. In addition, 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.

Experiment 1

Eye-tracked native speakers of Dutch 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.

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 450×450 to 600×600 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° 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) \times 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).

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 $\alpha < .05$.

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

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
<i>Experiment 1</i>				
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)
<i>Experiment 2</i>				
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)

condition: speakers looked at cued agents and cued patients on 87% and 62% of all scored trials respectively (Fig. 2).

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 ($\beta = -.58$, $z = 1.46$, 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.

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 ($\beta = -.19$, $z = 1.03$) and did not differ across blocks ($\beta = .18$, $z = 1.03$). There were also no effects of Event, Agent or Patient codability (all z s < .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 ($\beta = .54$, $z = 2.96$, 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 ($z = -.58$).

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 ($\beta = 2.31$, $z = 4.08$, 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 ($\beta = -2.73$, $z = -4.98$, 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

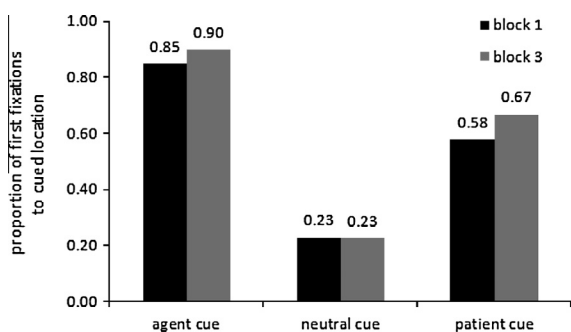


Fig. 2. Proportion of first fixations directed to the cued location (agent, patient, and neutral object) for all scored trials in Experiment 1.

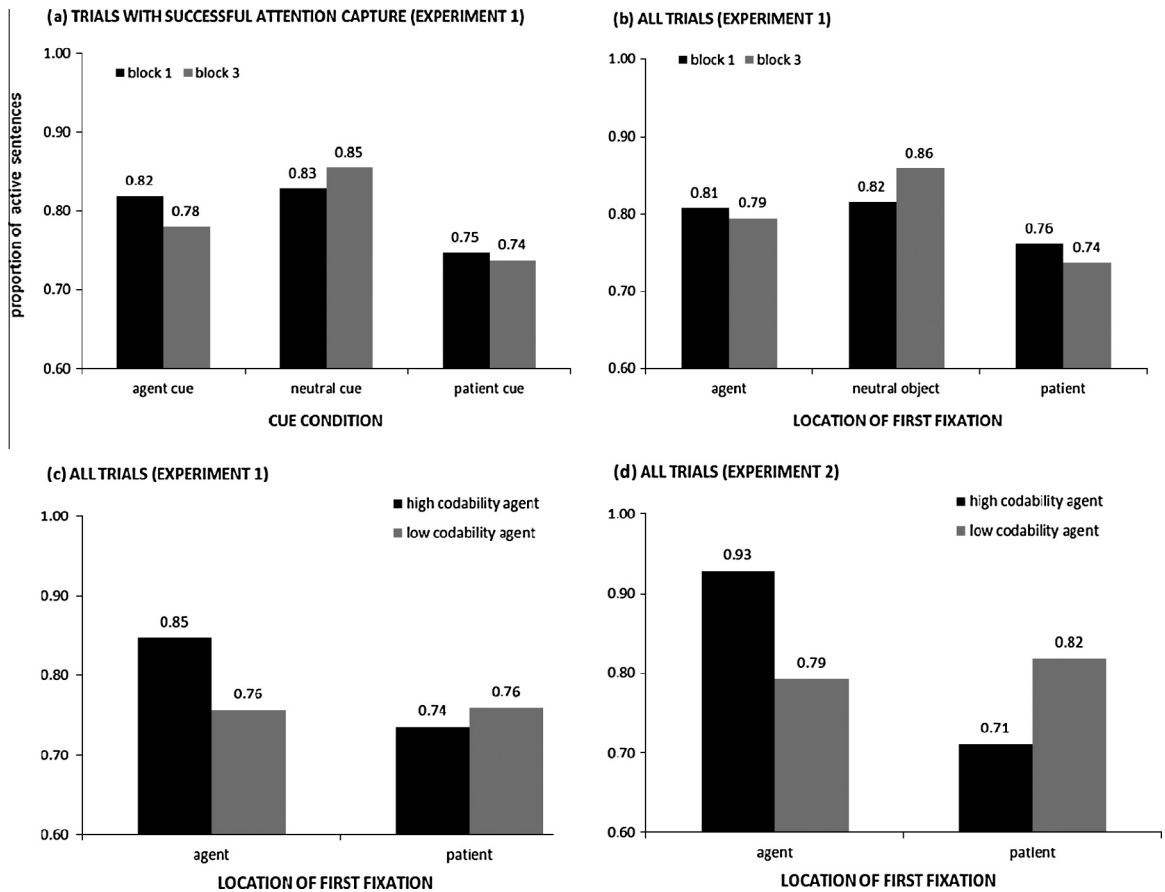


Fig. 3. Proportion of active sentences produced in (a)–(b) Experiment 1 when first fixations were directed to agents and patients. Figure (a) shows results for trials with successful attention capture; figure (b) shows results for all trials, irrespective of attention capture. Fixations to the neutral 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.

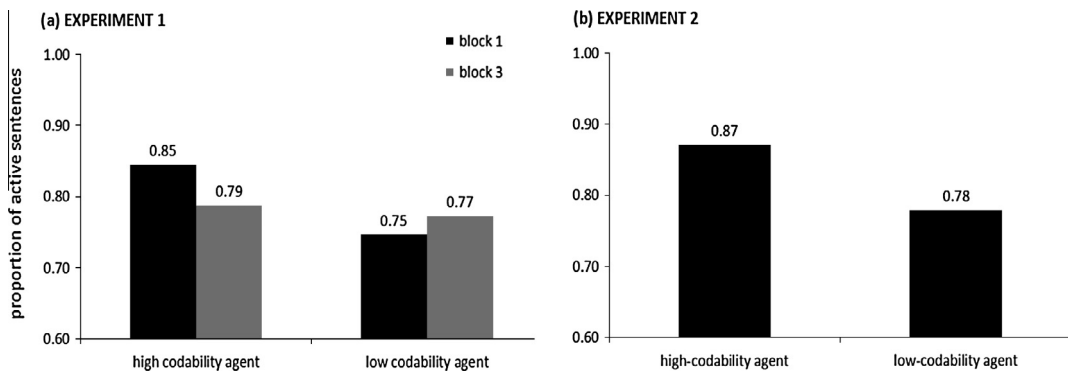


Fig. 4. Proportion of active sentences produced with respect to Agent codability in (a) Experiment 1 (across blocks) and (b) Experiment 2.

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

Fig. 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.

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). 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.

Supporting these predictions, agent fixations between 600 ms and speech onset (approximately 1800 ms) were sensitive both to Agent and Event codability (Table 3b).

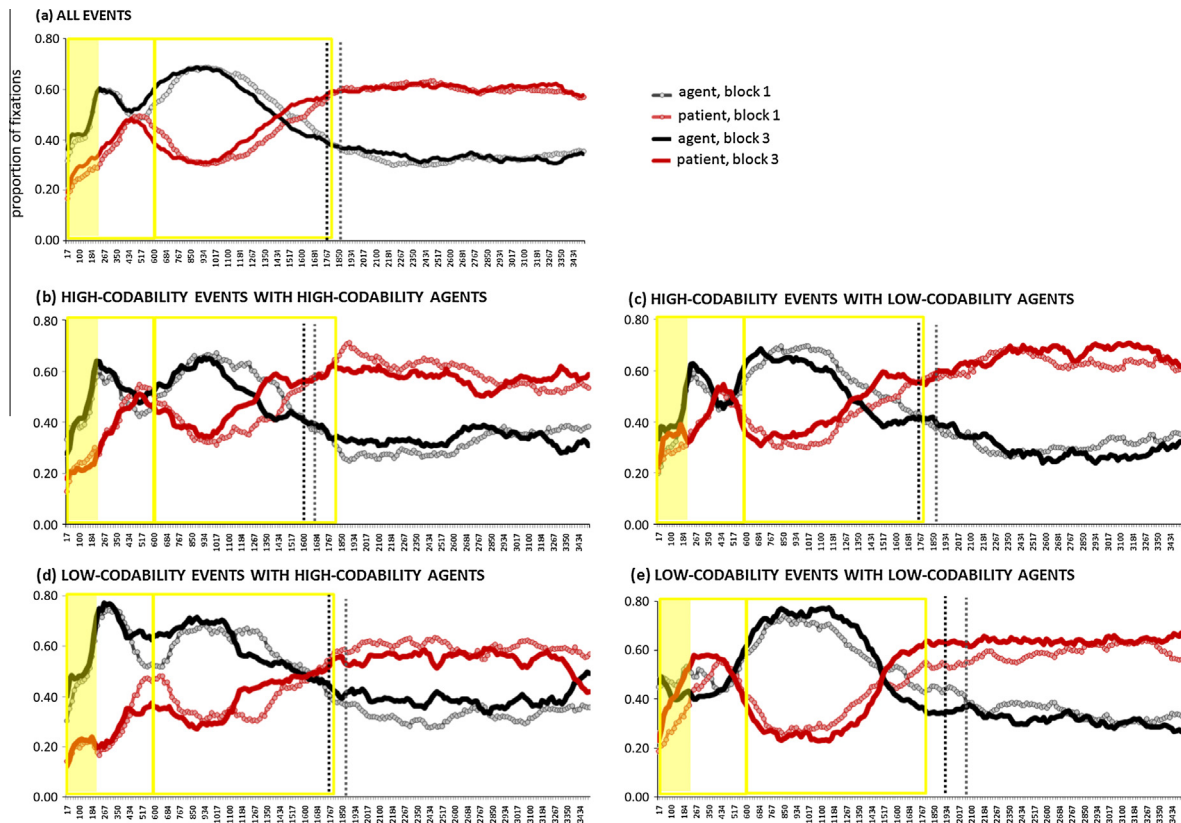


Fig. 5. Timecourse of formulation for active sentences in Blocks 1 and 3 of Experiment 1 for all events (a) and separately for higher- and lower-codability events with higher- and lower-codability agents (b)–(e). Yellow boxes show the time windows included in the analyses (0–200 ms, 0–600 ms, 600–1800 ms). The vertical dashed lines indicate speech onsets in Block 1 and Block 3. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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.

Fixed effects	By-participants			By-items		
	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.

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

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.

Fixed effects	By-participants			By-items		
	Estimate	SE	t-Value	Estimate	SE	t-Value
<i>(a) 0–600 ms</i>						
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	-3.45*	-.07	.02	-4.71* ^a
Event Cod. * Agent Cod. * Time bin	-1.09	.20	-5.59*	-.08	.02	-3.49* ^a
<i>(b) 600–1800 ms</i>						
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.

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 ($\beta = 169$, $t = 3.85$, 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 ($\beta = 118$, $t = 1.96$, for the interaction of Agent and Event codability). Onsets were also faster in Block 3 than Block 1 ($\beta = 107$, $t = 3.32$; 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, Fig. 6b and c and Fig. 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

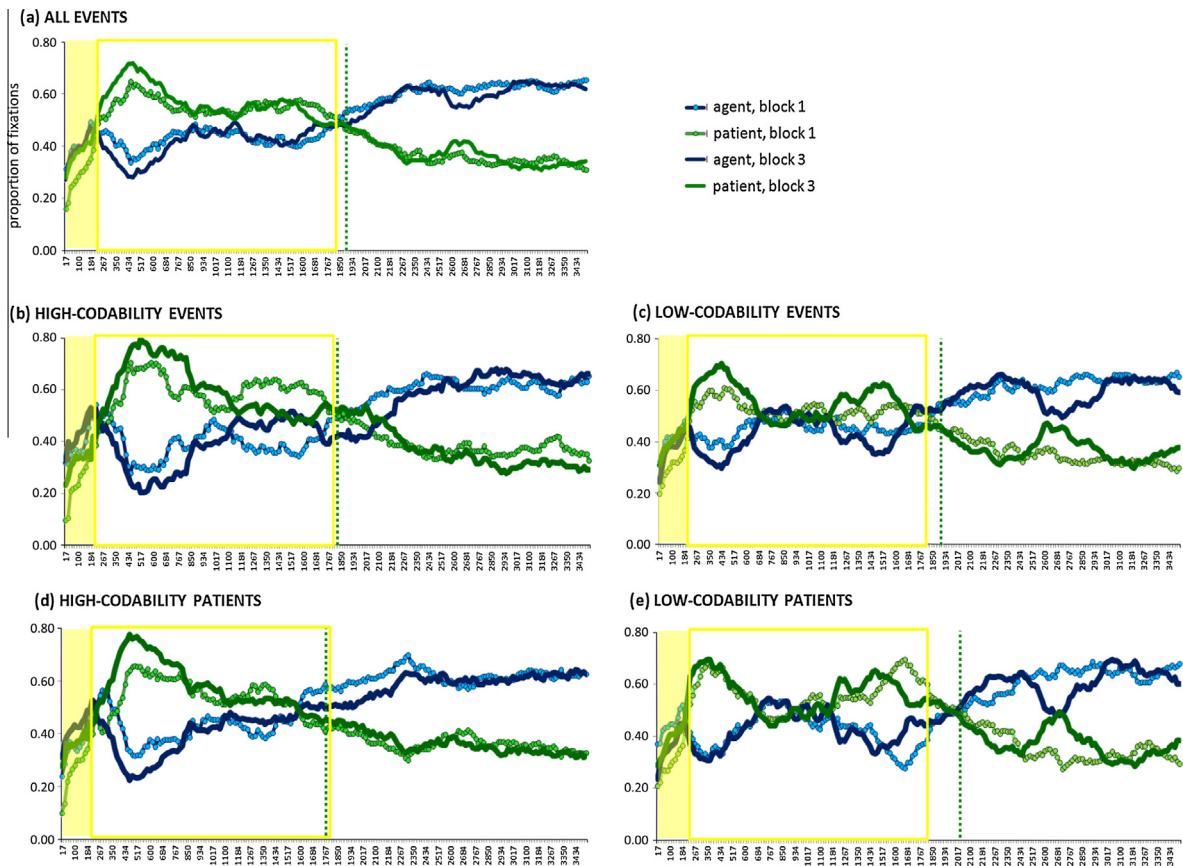


Fig. 6. Timecourse of formulation for passive sentences in Blocks 1 and 3 of Experiment 1 for all events (6a) and separately for higher- and lower-codability events (b and c) and higher- and lower-codability patients (d and e). The vertical dashed lines indicate speech onsets collapsed across Blocks.

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.

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

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.

Fixed effects	By-participants			By-items		
	Estimate	SE	t-Value	Estimate	SE	t-Value
<i>(a) 0–200 ms</i>						
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
<i>(b) 200–1800 ms</i>						
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.

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.

Speech onsets

Analyses of speech onsets show a joint influence of Patient codability, Event codability, and Block ($\beta = 264$, $t = 3.22$, 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.

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.

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.

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 re-starts), 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.

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 z 's > 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

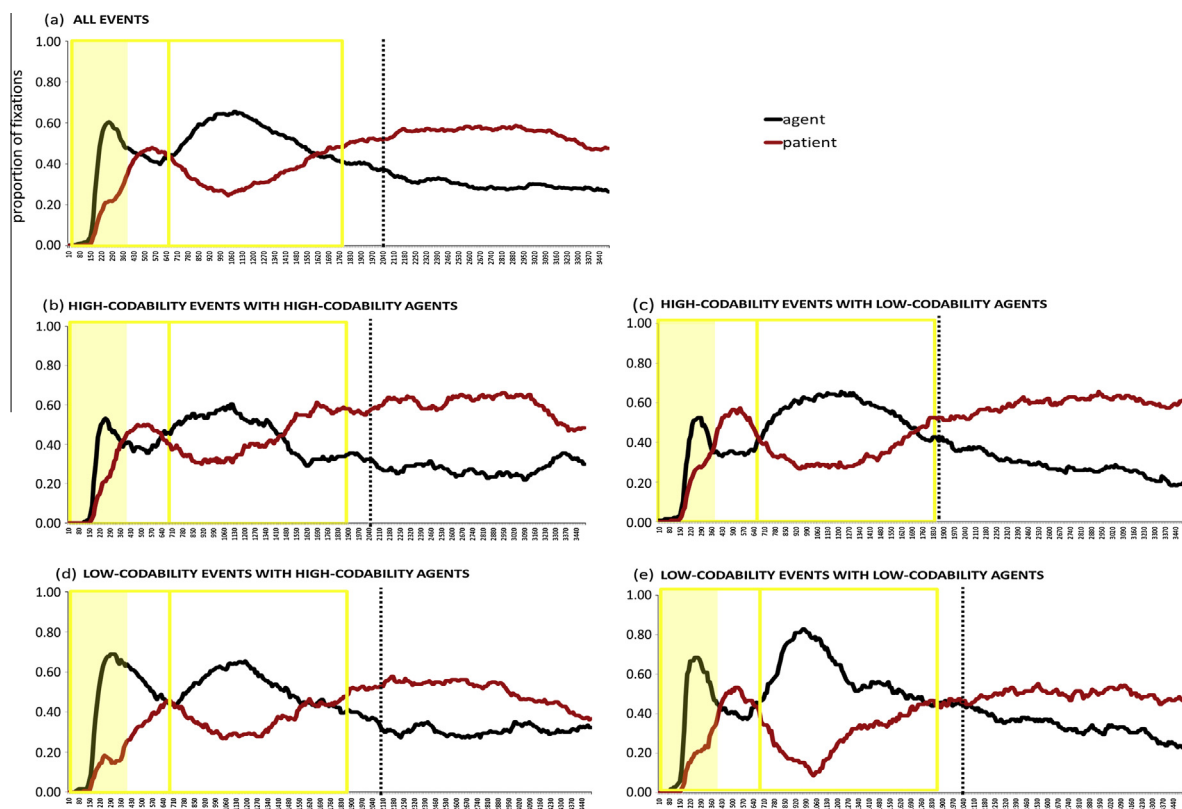


Fig. 7. Timecourse of formulation for active sentences for all events in Experiment 2 (7a) and separately for higher- and lower-codability events with higher- and lower-codability agents (b)–(e). Yellow boxes show the time windows included in the analyses (0–400 ms, 0–700 ms, 700–1900 ms). The vertical dashed lines indicate speech onsets. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the interaction between Event and Agent codability did not reach significance.

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.

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 ($\beta = 134$, $z > 2.08$; 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.

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

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.

Fixed effects	By-participants			By-items		
	Estimate	SE	t-Value	Estimate	SE	t-Value
<i>(a) 0–400 ms</i>						
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*
<i>(b) 0–700 ms</i>						
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*
Agent codability	–.09	.03	–2.79*	–.12	.04	–3.12*
<i>(c) 700–1900 ms</i>						
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*
Agent codability	.25	.02	13.51*	.33	.02	18.03*

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.

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).

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

<i>Items with high event codability, high agent codability</i>	<i>Items with low event codability, high agent codability</i>
Ambulance hitting car	Santa dragging Christmas tree
Lightning striking church	Wind blowing away papers
Ducklings following mother duck	Bomb hitting ship
Fireman saving child	Elephant lifting clown
Thief stealing painting	Shark attacking man
Ball hitting boy	Alarm waking up boy
Frog catching fly	Cat catching mouse
Stork bringing baby	Helicopter pulling diver out of the water
Horse kicking cow	Coach hugging tennis player
Boxer hitting cheerleader	Dog chasing mailman
Girl pushing boy on sled	Pirate digging up treasure
Girl tickling boy	Robot crushing computer
Snake frightening girl	Train crushing bus
Mother feeding baby	Bird pulling worm out of the ground
River breaking dam	
Avalanche approaching skiers	
<i>Items with high event codability, low agent codability</i>	<i>Items with low event codability, low agent codability</i>
Dog licking puppy's face	Truck towing car
Cop stopping car	Waiter kicking outman
Dog leading blind person	Astronaut catching an alien
Bishop crowning king	Cooks preparing cake
Mayor unveiling statue	Snow burying house
Man shooting woman	Bodyguard protecting the politician
Painter painting wall	Hunter roasting pig
Photographer filming model	Cop arresting man
Technician fixing tv	Bee stinging man
Football player scoring goal	Woman dressing boy
Journalist interviewing an actor	Sprinkler splashing old woman

Appendix A (continued)

<i>Items with high event codability, low agent codability</i>	<i>Items with low event codability, low agent codability</i>
Masseuse massaging man	Maid throwing out garbage
Professor congratulating student	Conductor directing orchestra
Firetruck hitting woman	Security searching bags
	Army attacking castle
	Bulldozer tearing down 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).

<i>Items with high event codability, high agent codability</i>	<i>Items with low event codability, high agent codability</i>
Girl pushing grandpa on sled	Cook preparing cake
Woman dressing boy	Crab cutting boy
Horse kicking cow	Frog eating fly
Superhero rescuing diver	Mayor unveiling statue
Girl taking cookie	Dog chasing mailman
Doctor examining patient	Bird pulling worm out
Cop arresting girl	Stork bringing baby
Ambulance ramming car	Cat catching mouse
Conductor conducting orchestra	Man shooting woman
Cowboy catching sheriff	Pirate burying treasure
Ballerina tickling boy	Santa dragging tree
	Shark attacking man

<i>Items with high event codability, low agent codability</i>	<i>Items with low event codability, low agent codability</i>
Bee stinging bear	Security searching bag
Professor congratulating student	Scout roasting pig
Technician fixing car	Maid throwing garbage out
Bulldozer destroys	Thief stealing painting

(continued on next page)

Appendix B (continued)

Items with high event codability, low agent codability	Items with low event codability, low agent codability
building	Bomb hitting ship
Bishop crowning king	Punk crushing computer
Woman interviewing an astronaut	Truck towing car
Painter painting wall	Windmill hitting farmer
Diver pushing paparazzi	Firetruck hitting traffic light
Dog licking puppy	Bodyguard protecting the queen
Woman massaging man	Lightning striking church
Assistant painting model	Army attacking castle
Wind blowing papers away	

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, 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, 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., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279.
- 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.
- Hartsuiker, R. J., & Kolk, H. J. (1998). Syntactic persistence in Dutch. *Language and Speech*, 41, 143–184.
- Jaeger, 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.