Encoding actions and verbs: Tracking the time-course of relational encoding during message and sentence formulation

Agnieszka E. Konopka

University of Aberdeen

Max Planck Institute for Psycholinguistics

Address correspondence to:

Agnieszka E. Konopka

School of Psychology

University of Aberdeen

agnieszka.konopka@abdn.ac.uk

Abstract

Two experiments tracked the encoding of relational information (*actions* at the level of the pre-linguistic message and *verbs* at the level of the sentence) during formulation of transitive event descriptions (e.g., *The tiger is scratching the photographer*). At what point during message and sentence formulation do speakers encode actions and verbs? Participants described pictures of transitive events in response to neutral questions (*What is happening?*), agent questions (*What is [the agent] doing?*), and patient questions (*What is happening with [the patient]?*). The agent and patient questions were intended to change the message-level focus of speakers' responses and to induce priority encoding of the event action and the sentence verb. The questions had a nearly categorical effect on speakers' choice of sentence form in their responses (characters mentioned in the questions were produced in subject position, as expected) and a strong effect on the time-course of sentence formulation: speakers rapidly directed their gaze to the part of the event needed to encode contextually *new*, *task-relevant* information – first the event action and the sentence verb, and then the sentence object. The distribution of fixations during the "verb-encoding" window showed that speakers encode relational information by fixating both event characters. Comparing formulation of sentences describing events with action-informative agents and action-informative patients showed a small preference for fixating the more informative character both immediately after picture onset (message-level encoding) and during the "verb-encoding" window (sentence-level encoding). The results identify action-specific and verb-specific eye movement signatures in message and sentence formulation.

Keywords: message and sentence formulation, incrementality, discourse, gaze-speech coordination, eye-tracking

In many of our interactions, we produce utterances that describe or recount events (*What happened at the party? What happened at the zoo?*). While there are multiple aspects to an event (the *who*, *what*, *when* and *where* of an event), a central component of event representations are the relationships and interactions between event protagonists (e.g., *The host* <u>*drank*</u> *all the wine; The tiger* <u>*scratched*</u> *the photographer*). Theories of event representation ascribe a pivotal role to relational information in the binding of individual characters into coherent propositions (e.g., Dobel, Glanemann, Kreysa, Zwitserlood, & Eisenbeiss, 2011; also see Biederman, Mezzanotte, & Rabinowitz, 1982).

This paper examines when and how speakers encode relational information when preparing to speak. Formulating (or "planning") a sentence like *The tiger scratched the photographer* involves a complex sequence of encoding operations. Formulation starts with generation of a conceptual, prelinguistic representation of the speakers' communicative intent, normally referred to as the *prelinguistic* or *preverbal message* (Levelt, 1989; see Konopka & Brown-Schmidt, 2014, and Bock & Ferreira, 2014, for reviews). In its simplest form, the message consists of information about *who did what to whom* in the event. In the current example, the message consists of information about the two event characters (the *who* and *whom* of the event: the concepts *tiger* and *photographer*) and the relationship between them (their event roles, or the difference between *who* and *whom*: agents and patients) and the *did what* aspect of the event (the action *scratching*)[1]. To map all of this information onto language, speakers must then retrieve suitable lexical items and assemble a suitable syntactic structure (*sentence-level* or *linguistic* encoding). Information about the *who*/*whom* and *did what* aspects of the event is mapped onto words from different classes: the identities of the event characters are expressed

[1] The focus of this paper is on encoding of actions (*who* <u>*did what*</u> *to whom*). See e.g., Hafri et al. (2013) for a discussion of role assignment to characters (agents and patients) in simple events.

with nouns ("*tiger"*, "*photographer"*) and the relationships between them can be expressed with verbs ("*scratched*")[2]. Here, encoding of *relational* information will be referred to as encoding of the event *action* at the message/conceptual level (*scratch*), and as encoding of the *verb* at the sentence/linguistic level ("*scratched*"). Encoding of event characters at the message level and sentence level will be referred to as *non-relational* encoding of the corresponding concepts (message level) and words (sentence level).

While speakers must eventually encode both relational and non-relational information in an event to produce a full sentence, the relative timing of the message-level and sentence-level encoding operations that must be performed to do so is still debated. This is because speakers rarely encode a full message and sentence before speech onset; instead, they tend to prepare their messages and sentences incrementally, i.e., in a series of small planning units (Levelt, 1989; Wheeldon, 2013). Thus, for example, speakers may encode only part of the message and sentence they will eventually produce before speech onset, and they continue to generate the rest of the message and sentence after speech onset. One of the implications of incrementality is that different message-level and sentence-level processes may receive priority over others at different points in time. Indeed, early accounts of message and sentence formulation ranged from the view that speakers can build up messages and sentences roughly one concept and one word at a time (Paul, 1880) to the view that formulation requires encoding of the *who did what to whom* content of an entire message before sentence-level encoding begins (Wundt, 1900; see Bock & Ferreira, 2014, Griffin & Bock, 2000, and Konopka & Brown-Schmidt, 2014, for reviews). The difference

[2] Relationships between message elements can, of course, also be conveyed by means of different referential terms (e.g., *host* vs. *guest*) and different sentence structures (e.g., intransitive vs. transitive constructions). The focus here is on verbs as one of the main carriers of relational information in simple SVO sentences (see Allum & Wheeldon 2007, 2009; Lee, Brown-Schmidt, & Watson, 2013; Konopka, 2012; Konopka & Meyer, 2014, for discussions of the involvement of structural processes in sentence planning).

between these accounts is articulated in current theories as *linearly incremental* and *hierarchically incremental* planning respectively, each making different predictions about the time-course of relational encoding.

Linear Incrementality posits that both message and sentence formulation can proceed in small increments, each one planned independently of the following increment, in the linear order of mention. For example, in SVO languages like English or Dutch, formulation of a message and sentence may begin with encoding of only one character at the message level (the concept *tiger*) and immediate encoding of this character at the sentence level (the phrase "*The tiger…*"; Gleitman, January, Nappa & Trueswell, 2007). The implication is that encoding of relational information (the action *scratch* and the verb "*scratched*") will occur in a subsequent, separate increment – according to the linear order of mention – and is thus not required during the earliest stages of formulation. In contrast, *Hierarchical Incrementality* posits that formulation begins with "a wholistic process of conceptualization" – i.e., an event apprehension phase during which speakers encode a larger message ("a coarse understanding of the event as a whole"; Griffin & Bock, 2000, p. 274, p.279). On this view, encoding of the action occurs early on, while encoding of the verb occurs around speech onset. This is because encoding of message-level relational information is inherently part of the early message formulation process: speakers must encode the event *action* (*scratch*) to encode the gist of an event (Bock, Irwin, & Davidson, 2004; Dobel et al., 2010; Konopka & Meyer, 2014). Sentence-level encoding starts only once a rough message-level representation is in place; thus, speakers should retrieve the words "*tiger*", "*scratched*" and "*photographer*" roughly in this order. Other proposals (e.g., Lindsley, 1975) argue that early formulation may involve selection and encoding of a subject as well as a verb (see Bock & Levelt, 1994; Ferreira, 2000). Together, these accounts illustrate different predictions about *how much* information and *what kind* of information speakers prepare before speech onset (also see Konopka, 2012, for a discussion of how prioritizing different types of information – relational vs. non-relational information – influences planning scope).

Experimentally, the time-course of message and sentence formulation can be tracked by examining speakers' eye movements as they prepare to describe a pictured event. Speakers systematically direct their attention to whatever part of the display they are processing with priority at a given point in time (e.g., Flecken, Gerwien, Carroll, M. & v. Stutterheim, 2015; Gleitman et al., 2007; Griffin & Bock, 2000; Hwang & Kaiser, 2014; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; see discussions in Henderson & Ferreira, 2004). To distinguish between *Linear* and *Hierarchical Incrementality*, it is important to identify the time windows during which speakers prioritize relational and non-relational encoding. Naturally, to do this, one must also identify fixation patterns that are specific to relational and non-relational encoding. Time windows associated with encoding of *non-relational* information (i.e., individual characters in an event) are easy to identify: event characters are discrete visual elements and are identified by distinct interest areas, so fixations to these areas index encoding of character-specific information. In fact, one of the most robust findings in this literature is that speakers fixate event characters in the order of mention (Griffin & Bock, 2000; Gleitman et al., 2007; Konopka & Meyer, 2014; Konopka et al., 2018; Kuchinsky & Bock, 2010; also see Bock et al., 2003) and maintains fixations on each character until they complete retrieval of its name (Griffin, 2001; Meyer, Sleiderink & Levelt, 1998). For pictures of transitive events (e.g., a tiger scratching a photographer), this pattern of gaze-speech coordination emerges approximately 200-400 ms after picture onset and is interpreted as showing sequential linguistic encoding of the two event characters. In contrast, hypotheses regarding the time-course of *relational* encoding (actions at the message level and verbs at the sentence level) are harder to test for at least two reasons.

First, tracking relational encoding is complicated by the lack of a clear measure indexing encoding of relational information (Dobel et al., 2010; see e.g., Gillette, Gleitman, Gleitman, & Lederer, 1999, for a discussion of verb learning from observation). Relational information is not encoded as easily as non-relational information by directing attention to a single region of an event. For example, in an event where one character kicks another character, the *kicking* action may be encoded by fixating the character performing this action (the agent) because lower-level perceptual information (such as body posture) can provide sufficient cues for action identification. Given the lack of ambiguity, this action may even be encoded parafoveally, prior to character fixation. In other events, encoding relational information requires fixating more than one character: e.g., in an event where one person is picking up a ball, the "action region" can include both the agent's hands as well as the ball. *Second*, the production system allows some degree of flexibility in the timing of various encoding operations (Ferreira & Swets, 2002; Konopka, 2012; Wagner, Jescheniak & Schriefers, 2010): speakers may begin relational encoding at different points during message and sentence formulation in different utterances, so there can be substantial overlap in the timing of various encoding operations before speech onset. Consequently, eye movements in any time window may reflect a combination of relational and non-relational processes, making it hard to identify consistent gaze patterns indexing message-level and sentence-level relational encoding alone.

So far, a number of studies tested for speakers' sensitivity to relational information immediately after picture onset with different methods, and yielded mixed evidence concerning the time-course of relational encoding. For example, Dobel et al. (2010) tracked fixations to "action regions" in simple events and showed that speakers did direct their gaze to action-relevant regions shortly after picture onset, consistent with *Hierarchical Incrementality*. Speakers were more likely to do this when producing event descriptions than when naming the event characters sequentially (also see Kreysa, Zwitzerlood, Bolte, Glanemann, & Dobel, 2010). These effects, however, varied across picture types (drawings vs. photographs) and event types (three-character dative events vs. two-character transitive events), which may have been due to the different locations of "action regions" in their events. In a different study, Hwang and Kaiser (2014) showed that English speakers do attend to action-relevant regions in two-character transitive events shortly after picture onset. However, speakers were familiarized with the pictured events and learned the target words (nouns and verbs) to use in their descriptions prior to the production task (see Myachykov and colleagues for a similar method); thus, while early fixations to action-relevant regions are again consistent with *Hierarchical Incrementality*, it is unclear whether the results would hold in more naturalistic production settings where speakers generate messages and sentences from scratch.

To address the problem of variability in the timing of relational encoding, Konopka and Meyer (2014) compared the time-course of message and sentence formulation for descriptions of events where relational information (actions and verbs) was either easier or more difficult to encode. Instead of tracking fixations to action-relevant areas, they compared the overall distribution of fixations to agents and patients in different time windows. At picture onset, speakers can direct their gaze to either of the two characters, and agent and patient fixations generally diverge by 400 ms. The timing of this divergence point indicates when speakers begin encoding the subject character with priority. *Linear Incrementality* predicts that fixations to the two characters should diverge well before 400 ms (Gleitman et al., 2007): speakers should begin fixating the subject character preferentially within 200 ms of picture onset, suggesting that they prioritize encoding of this character over encoding of any other information in the event. In contrast, *Hierarchical Incrementality* predicts that the divergence point in agent and patient fixations should occur later (as late as 400 ms): Griffin and Bock (2000) proposed that speakers should first distribute their attention between the two characters to encode information about the event as a whole, and only then direct their attention to the subject character to begin linguistic encoding. There is now ample evidence that speakers can extract basic gist information about an event within as little as 37 ms (Hafri, Papafragou, & Trueswell, 2013; Dobel, Gumnior, Bolte, & Zwitserlood, 2007) but that encoding of action information takes more time and is more attention-demanding than character recognition or role assignment (also see Biederman 1972; Biederman & Rabinowiz, 1974; Biederman et al., 1982). Thus, finding a convergence in fixations to agents and patients for as long as 400 ms is compatible with the hypothesis that speakers devote resources to encoding *relational* information early on during formulation.

The results from production of active SVO sentences in Dutch (Figure 3 in Konopka & Meyer, 2014) showed that early eye movements differ between events where relational information is easy to encode and difficult to encode (even after controlling for variables like the ease of naming the subject character). Speakers showed a strong preference for fixating the agent over the patient within 400 ms of picture onset (*linearly* incremental planning) when describing ambiguous events, i.e., events where the event actions could be interpreted in different ways and that were described with a range of verbs (low-codability events; Kuchinsky & Bock, 2010). This suggests that, when speakers are faced with multiple options as to how they can interpret and describe the event action, they may prefer to encode the subject character first and thus postpone encoding of relational information (both actions and verbs). In contrast, speakers showed a smaller preference for the agent over the patient before 400 ms (*hierarchically* incremental planning) when describing events that were easier to apprehend, i.e., events where the main action was easy to interpret and describe (high-codability events). In this case, speakers devoted more time to encoding relational information before beginning linguistic encoding of the subject character. Such differences in gaze patterns across events confirm that early formulation *is* sensitive to the ease of encoding the relational content of an event, but also that speakers can flexibly allocate resources to encoding different types of information. Importantly, given the timing of these effects (0-400 ms), the conclusions concern flexibility in the encoding of the event *action* (i.e., encoding of relational information at the message level); these experiments do not isolate time windows during which speakers encode the sentence *verb*[3].

In fact, incrementality and flexibility in message and sentence formulation complicate further assessment of the time-course of relational encoding. In the current example, flexibility implies that encoding of relational information at the message-level may be "distributed" over a rather broad window: relational encoding may begin in the first 0-400 ms of picture viewing, when speakers start encoding the event action (high-codability events), but may also continue after 400 ms (low-codability events). After 400 ms, however, speakers normally fixate event characters in the order of mention, so any action encoding that remains to be completed will largely overlap with encoding of the subject and object characters. A similar problem concerns retrieval of the sentence *verb*. Both *Linear* and *Hierarchical Incrementality* predict that verb retrieval in SVO sentences occurs after speakers finish retrieving the subject character name, i.e., approximately around speech onset. However, speakers must be encoding verbs during fixations

[3] Hwang and Kaiser (2014) interpret early fixations to action regions as indicating verb encoding; however, since participants in their study were familiarized with the target events and target words before beginning the experiment, it is possible that speakers began retrieving verbs from memory in the early time windows.

to the two characters: specifically, when fixating any given character, speakers may be encoding information about the character itself (its identity and name) as well as whatever relational information this character contributes to the event. Indeed, since verb retrieval generally takes longer than noun retrieval (Shao, Roelofs, & Meyer, 2012), it is possible that a large proportion of the time spent fixating a given character is dedicated to encoding the verb. But since eye movements after 400 ms reflect a combination of relational and non-relational encoding processes, this limits the possibility of observing unique verb-specific components in gaze-speech coordination.

Another approach to tracking the processing of relational information is to compare formulation of verb-initial, verb-medial and verb-final sentences across languages (e.g., Hwang & Kaiser, 2014; Konopka & Kuchinsky, 2015; Momma, Slevc, & Phillips, 2016, 2018; Norcliffe et al., 2015; Santesteban, Pickering, Laka, & Branigan, 2015; Sauppe, Norcliffe, Konopka, Brown, van Valin, & Levinson, 2013; Sauppe, 2017; Schriefers, Teruel, & Meinshausen, 1998). Although the consensus is that the timing of verb retrieval is determined roughly by linear word order, finer-grained comparisons of "verb-encoding" windows across studies are complicated by a combination of factors (such as the presence or absence of case marking, language-specific structural preferences for different types of events, differences in the selected structures across studies, as well as methodological differences; see Jaeger & Norcliffe, 2009).

This paper uses a new approach to examine when and how speakers encode relational information (*actions* at the message level and *verbs* at the sentence level). In two experiments, speakers described pictures of simple transitive events with active (SVO) and passive (OVS) sentences. The time-course of encoding actions and verbs in these sentences was assessed with two manipulations. *First*, the events included characters (agents and patients) that differed in the extent to which they were informative for encoding the event action. These differences were operationalized in terms of a continuous event-specific *action-informativeness* score (see the norming study below). *Second*, the event descriptions were produced in response to questions that either did or did not explicitly induce priority encoding of relational information. Questions were presented prior to picture onset, and either provided no context for the target event (*What is happening?*) or shifted the focus of speakers' responses to relational information (*What is the tiger doing? What is happening with the photographer?*). In Experiment 1, participants were asked to respond to the questions in any way they liked (the most frequent response for this event is *The tiger is scratching the photographer*). In Experiment 2, they were asked to begin their sentences with a pronoun if they heard a character-specific question (e.g., Q: *What is the tiger doing?* A: *It is scratching the photographer*). The questions provide a way of "zooming in" on encoding of the action and the verb; thus, instead of drawing inferences about relational encoding from the length of fixations or from gaze shifts between characters, this manipulation allows identification of a specific time window and a specific pattern of fixations that directly correspond to action and verb encoding.

Encoding relational information: "where" is the action?

Previous research assumed that speakers can encode relational information by fixating "action regions" in an event, usually defined as points of physical contact or proximity between event characters (see Dobel et al., 2011; Hwang & Kaiser, 2014). However, to assign characters to correct event roles and to describe an event accurately, speakers must also interpret and verify their interpretations of each character's actions in the context of the other character. This requires encoding action-relevant details of *both* characters. Thus, an alternative to analyzing fixations to a single "action region" is to use a derived measure like the likelihood of speakers fixating *both* characters (Griffin & Bock, 2000; Konopka & Meyer, 2014). Importantly, event characters can differ in the extent to which they provide cues associated with conceptual categories like agenthood (such as body movement and orientation; Hafri et al., 2013), and they also differ in the extent to which they provide information relevant for encoding a *specific* action and verb: in some events, encoding the agent can be sufficient to identify the action that this character is performing on the patient, while in other events, speakers may need to devote resources to the patient as well (Dobel et al., 2011; see Momma et al., 2016, for another characterization of the relationship between verbs and patients/objects).

In the current experiments, speakers described events that differed on two dimensions relevant for relational encoding. One dimension was the *action informativeness* of each event character, i.e., speakers' ability to encode information about the event action by fixating either the agent or the patient (i.e., encoding one character independently of other). The second dimension was the ease of encoding the action itself, i.e., *event codability* (Konopka & Meyer, 2014; Kuchinsky & Bock, 2010). The focus in this paper is on *action informativeness*, and event codability is included in the analyses for control purposes.

Differences in the action informativeness of agents and patients in each event were established in a norming study: participants saw incomplete pictures (pictures where one character was visible and one character was masked) and were asked to guess the event action by writing down as many verbs as they could think of. Guess accuracy was evaluated by comparing these responses against sentences produced in the main experiments (where speakers saw the "full" events). An *action informativeness* score for each event was calculated from the proportions of correct guesses for the agent-visible and patient-visible pictures: a higher proportion of correct guesses in the agent-visible condition indicates that the agent is more action-informative than the patient, and vice versa. The experiments then compared the distribution of fixations to agents and patients during production of descriptions for events with action-informative agents (*agent-action* events) and action-informative patients (*patient-action* events). The analyses used the action informativeness scores as a continuous predictor to gauge speakers' sensitivity to relational information over time.

On Griffin and Bock's (2000) and Konopka and Meyer's (2014) account, sensitivity to relational information at the *message* level is primarily indexed by the distribution of fixations in the 0-400 ms time window during formulation. Thus, if SVO syntax favors a *linearly incremental* planning strategy (i.e., encoding of the subject character before encoding of relational information), there should be little difference in fixations to agents in agent-action events and patient-action events across conditions. But if speakers encode some information about event actions early on, even when the first content word they produce is a character name (*hierarchically incremental* planning), there should be more fixations to the action-informative character than the less action-informative character in this time window (either the agent or the patient). This would provide strong evidence that speakers prioritize encoding of relational information (in addition to encoding any non-relational information about the fixated characters) during early event viewing.

To examine sensitivity to relational information at the *sentence* level, one can then compare the length of gazes on the subject character after 400 ms. For sentences produced out of context or in a neutral context (e.g., after questions like *What is happening?*), speakers should begin encoding the sentence verb around the time when they finish encoding the subject character (i.e., around speech onset): thus, they should fixate a subject character *longer* if this character is the main carrier of relational information in an event than when the object character is the main carrier of relational information.

Naturally, the length of fixations on individual characters also depends on the ease of character naming, so any differences in gaze patterns to agents in events with action-informative agents and events with action-informative patients are only an indirect index of verb encoding. Thus, to track verb encoding after 400 ms with greater temporal detail, the current experiments character-specific questions to identify a unique "verb-encoding" window.

Formulating sentences in response to questions

Production studies typically elicit sentences out of context, i.e., in situations in which all message elements are new or unfocused (Dobel et al., 2011; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Van de Velde et al., 2014). The task for speakers is thus to generate a preverbal message on the basis of their own interpretation of the event and to generate all the linguistic material needed to express this message. Lexical retrieval and structure choices may be primed across trials (e.g., Bock, 1986a, 1986b; Gleitman et al., 2007; Hwang & Kaiser, 2015; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Myachykov, Tomlin & Posner, 2005) or may be facilitated by familiarizing speakers with the stimuli before the experiment (e.g., Hwang & Kaiser, 2014; Momma et al., 2016; Konopka et al., 2018; Myachykov, 2007), but all of the information needed to produce an event description is nevertheless *contextually* new.

One way of tracking the encoding of individual message elements is to ask speakers to produce event descriptions in response to questions that bring those parts of the message into focus. Questions provide explicit cues that shape the conceptual focus of the prelinguistic message and the surface form of the sentence conveying this message (e.g., see discussions of givenness and perspective-taking: Bock, 1977; Givon, 1992; MacWhinney, 1977, 2005; MacWhinney & Bates, 1978; Thompson, Ling, Myachykov, Ferreira, & Scheepers, 2013). In the current experiments, speakers heard neutral questions (control condition: *What is happening?*) or character-specific questions (Agent-questions: *What is the tiger doing?*; Patient-questions: *What is happening with the photographer?*) before picture onset. The Agent- and Patient-questions were expected to coerce preferential encoding of relational information in these events: they establish the mentioned character as being contextually *given* or *grounded* and thus they shift the focus of speakers' responses to *new* information. In SVO/OVS sentences, the first piece of new conceptual information to be encoded is the event *action*, and the first piece of new linguistic information is the sentence *verb* (*The tiger is scratching the photographer; The photographer is being scratched by the tiger*). Thus, at the message level, these questions ask speakers to prioritize encoding of the *action* carried out by or on this character. At the sentence level, the questions reduce the costs of linguistic encoding the given character and they bias selection of a sentence structure where the given character is produced in subject position[4]; thus, they also facilitate early encoding of the sentence *verb*. In other words, questions influence the relative timing of encoding *given* and *new* message elements: speakers should quickly "zoom in" on relational information in the event, revealing how they encode the event action (before 400 ms) and the sentence verb (after 400 ms).

In early time windows (before 400 ms), the magnitude of any differences in the Agent-question and Patient-question conditions relative to the Neutral condition should depend on the

[4] The syntax of the agent- and patient-questions was deliberately different. The questions were chosen to unambiguously elicit the same *shift* in the message-level focus of speakers' responses in the two conditions with character-specific questions: agent- and patient-questions establish the mentioned character as the "doer" or the "undergoer" of the action respectively and they elicit sentences where the *old/given* character is mentioned first. Thus despite the difference in the syntax of active and passive sentences, the first *new* piece of information expressed linguistically in both sentences is the sentence verb followed by the *new* character.

degree to which speakers prioritize message-level relational encoding. If speakers normally prioritize encoding of the event action before 400 ms (i.e., if formulation is *hierarchically incremental* in the Neutral condition), then answering an Agent-question or Patient-question should show a similar pattern of eye movements, as speakers would be engaging in relational encoding to the same extent across conditions. There should also be a preference for fixating the more action-informative character in all conditions. If formulation is *linearly incremental* in the Neutral condition, then answering an Agent-question or Patient-question should produce a shift towards priority encoding of the event action, indexed by a shift away from fixating the subject character and towards the action-informative character (either the agent or the patient).

After 400 ms (i.e., during sentence-level encoding), the question manipulation should allow identification of a “verb-encoding” window. Agent-questions and patient-questions alter the *given-new* status of the subject character and thus reduce the costs of linguistic encoding of this character name, allowing speakers to quickly begin encoding the next content word –i.e., the verb. It is possible that repeating a character name may simply reduce the length of gazes on this character – i.e., speakers may simply spend less time fixating a character they already have a verbal label for – without changing the overall pattern of sequential looks to the two event characters in the order of mention. Support for this type of *linear deployment of attention* during linguistic encoding comes from studies showing that speakers systematically direct and *re-direct* their attention to each referent every time it is mentioned (e.g., Griffin, 2001; Meyer et al., 1998) Meyer, Wheeldon, van der Meulen, & Konopka, 2012; van der Meulen, Meyer & Levelt, 2001). Thus, in the current experiments, speakers may still fixate the old subject character preferentially (but perhaps more briefly) right before naming it. However, if changes in message focus are stronger predictors of attentional focus than linear order of mention, speakers may “skip” fixating the contextually *old* subject character (tiger) and immediately direct their gaze to whatever part of the display “contains” *new* information (the sentence verb). In this case, the tight coupling of gaze and speech observed during the time window corresponding to linguistic encoding of the subject character would be abandoned in favor of a less linear encoding strategy. This type of *discourse-driven deployment of attention* is compatible with theories proposing that eye movements during production reflect guidance from a higher-level message plan (e.g., Bock et al., 2004) and would demonstrate strong top-down effects of context on formulation (also see Chun, 2000, and Henderson, Malcolm, & Schandl, 2009). Experiment 1 tested how large discourse effects are in this task, and Experiment 2 tested how these effects change when the time spent producing the subject character is further reduced by asking participants to use pronouns (*he/she/it*) to refer to the subject character.

So, if speakers need little time to encode a contextually *old* character, what does verb-encoding look like? Applying the same logic to encoding of relational information at the message level (before 400 ms), selection and retrieval of the verb should also be indexed by a *convergence* of fixations to agents and patients after 400 ms. Thus, when speakers “zoom in” on encoding of the verb in the Agent-question and Patient-question conditions, they should distribute their gaze between agents and patients over a fairly broad window – beginning at 400 ms and lasting as long as it takes to retrieve the verb. Further, if the *new* information that speakers encode in the “verb-encoding” window is indeed verb-specific, then the distribution of fixations should be modulated by *action informativeness*: speakers should show a preference for fixating the more action-informative than the less action-informative character.

Summary of predictions and analyses

In two experiments, speakers described pictures of events in response to questions. In the Neutral-question condition, gaze patterns were expected to replicate earlier studies (Gleitman et al., 2007; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010): formulation should begin with priority fixations to agents before 400 ms (*Linear Incrementality*) or with an early period of converging fixations to agents and patients (*Hierarchical Incrementality*), and after 400 ms, it should proceed with sequential fixations to the two characters in the order of mention. In contrast, in the Agent-question and Patient-question conditions, speakers should quickly direct resources to encoding *new* information – in this case, relational information. The magnitude of the contrast between the Neutral-question condition and the character-question conditions is informative for assessing the strength of top-down context effects on formulation and for identifying a "verb-encoding" window.

Analyses were carried out for active sentences in each experiment individually, and formulation of passive sentences was examined in a joint analysis for both experiments due to sparse data (passives are dispreferred structures). For a fine-grained assessment of the time-course of relational encoding (actions and verbs), all analyses compared formulation in events with action-informative agents and action-informative patients across conditions. The prediction was that speakers would spend more time fixating the action-informative character whenever they devoted resources to encoding relational information. Selection of time windows for these analyses was both theory- and data-driven, as outlined below.

Finally, a complementary analysis was carried out to verify the validity of these conclusions using a different, already known index of relational encoding – *event codability* (the ease of encoding event gist; Kuchinsky & Bock, 2010). As outlined earlier, event codability influences the distribution of agent-directed and patient-directed fixations shortly after picture onset in sentences produced out of context (Konopka & Meyer, 2014). In the current experiments, effects of event codability should be observed in the "verb-encoding" window in the Agent- and Patient-question conditions as well. If the *new* information that speakers encode with priority after 400 ms is indeed verb-specific, then event codability should influence fixation patterns over and above any effects of action informativeness: the window during which speakers distribute their attention between agents and patients should be shorter in higher-codability events (where the verb is easy to select) and longer in lower-codability events (where the verb is harder to select). This provides independent evidence for interpreting fixations in the "verb-encoding" window as showing priority encoding of relational information.

Picture norming

The target items were 78 pictures of transitive events with unique agent-verb-patient combinations (see Appendix). 55 of the pictures had human agents and 23 had non-human agents; 40 had human patients and 38 had non-human patients[5].

The pictures were normed in a separate study with 36 participants (drawn from the same participant pool as Experiments 1 and 2; all studies received ethical approval from Rabdoud Universiteit Nijmegen). Participants received one of two versions of the norming booklets (18 participants per booklet). One version showed pictures in which the agent character was visible but in which the patient had been replaced by a grey oval, masking the identity of this character (*agent-visible* condition). The other version showed pictures in which the patient character was visible but in which the agent had been replaced by a grey oval (*patient-visible* condition). The

[5] Animacy had the expected effect on structure choice in the Neutral-question conditions of both experiments: speakers produced more active sentences to describe events with human than non-human agents (Experiment 1: .79 vs. .55; Experiment 2: .78 vs. .58), and more passive sentences to describe events with human than non-human patients (Experiment 1: .42 vs. .13; Experiment 2: .37 vs. .15). Animacy had no effect on structure choice in the remaining conditions and is not discussed further.

size of the oval corresponded to the size of that character in the original picture. A sentence preamble appeared below each picture: the preamble began with the name of the visible character and was followed by a dotted line (indicating the missing verb) and a grey box (representing the masked character) to indicate that all events were transitive: e.g., “The tiger is…… [grey box]” and “The photographer is being……by [grey box]”. The task for participants was to guess what action the visible character was performing on the masked character (participants were not asked to guess the identity of the masked character). This method provides a global assessment of how each character contributes to the event action, whichever cues participants rely on to make their judgments (world knowledge, character-specific perceptual cues, etc.; see Hafri et al., 2013).

Participants wrote down a total of 3434 transitive verbs (intransitive verbs were excluded; .05 of all responses), with an average of 1.29 verbs per picture (1.29 and 1.27 in the agent-visible and patient-visible conditions; range 1-5). To assess guess accuracy, responses were compared to verbs produced in Experiments 1 and 2. Responses were scored as *correct* if the verbs from the norming study had been used at least once in the experiments or if they were close synonyms of these verbs[6]. Novel verbs were scored as being *incorrect* guesses (e.g., *to photograph* for the picture of an army attacking a castle in the patient-visible condition).

The results showed high variability across items (*M*=.67 and .55 correct guesses in the agent-visible and patient-visible conditions respectively; *SD*=.29 in both conditions). For example, the picture of a boxer punching a cheerleader had 1.00 correct guesses in the agent-visible condition but only .23 correct guesses in the patient-visible condition, indicating that speakers could identify the event action (*punching* or *hitting*) by looking exclusively at the agent.

---

[6] For example, *vernielen* [*to destroy*] was accepted as a synonym of *verwoesten* [*to destroy*] for the picture of a bulldozer hitting a building. Including and excluding close synonyms of the used verbs produced similar results.

In contrast, the picture of a bird puncturing a hot-air balloon had .00 correct guesses in the agent-visible condition but .75 correct guesses in the patient-visible condition, indicating that the patient was critical for identifying the action (*puncturing*). Some items also elicited very few correct responses in both conditions, indicating that, in these items, both characters were crucial for action encoding. An *Action Informativeness* score was calculated for each item by taking the difference in guess accuracy in the agent-visible and patient-visible conditions, and dividing this difference by the overall accuracy guess rate for that item.

Based on these scores, 46 of the target events were categorized as including agents that were more action-informative than patients (i.e., *agent-action* events with action informativeness scores larger than 0), and 32 of the target events included patients that were more action-informative than agents (i.e., *patient-action* events with scores smaller than 0). This variable was then used as a continuous predictor in most analyses reported below, but for clarity, figures plot results for agent-action and patient-action items separately.

The Action Informativeness of agents and patients did not vary with other properties of the events: correlations of this variable with Event Codability, Agent Codability, and Patient Codability were negligible (all *r*s<.10 in both experiments; see Table 1 for means).

Table 1.
Mean character Action Informativeness scores and Codability ratings (with standard deviations) across items in Experiments 1 and 2.

Items	Action Informativeness	Event Codability	Agent Codability	Patient Codability
Experiment 1				
Agent-action events (*n*=46)	.79 (.65)	1.59 (.83)	.90 (.72)	.98 (.77)
Patient-action events (*n*=32)	-.67 (.64)	1.47 (.86)	.71 (.64)	.86 (.71)
Experiment 2				

Agent-action events (*n*=46)	.79 (.65)	1.42 (.76)	.98 (.75)	1.00 (.74)
Patient-action events (*n*=32)	-.67 (.64)	1.44 (.77)	.82 (.77)	.96 (.67)

Experiment 1

Participants described pictures in response to questions, using the full name of the subject character in their responses.

Method

Participants

Thirty native speakers of Dutch, mostly students at Radboud Universiteit Nijmegen, gave informed consent and participated for payment. Three had to be replaced due to track loss.

Materials and design

Target displays were preceded by three types of questions. Speakers heard either a neutral question (*What is happening?*), a question about the agent (e.g., *What is the tiger doing?*), or a question about the patient (e.g., *What is happening with the photographer?*). Question type was manipulated within-participants and within-items, so three lists were created to present each target picture in each condition to different participants. Within lists, each participant received 26 items in each condition. Two versions of each target picture were created by mirror-reversing the placement of the agents and patients (left vs. right hand-side of the event); agent placement was also counterbalanced across lists and conditions (resulting in a total of six lists), but all analyses collapsed across this factor.

Target pictures were separated by 2 or 3 filler trials. Filler trials showed pictures of one-character or multi-character events eliciting different types of descriptions (132 trials). These trials were preceded by different types of questions: neutral questions, agent questions, patient questions, and other questions (e.g., *Where is the snake?*). A small subset of filler trials (30 trials) showed single words instead of pictures (30 trials); these trials were included to disguise the purpose of the experiment (see below).

Procedure

The experiment was run with an Eye-link 1000 eye-tracker (500 Hz sampling rate). Instructions were printed on the screen and were paraphrased by the experimenter. Participants were not familiarized with the stimuli prior to starting the experiment.

Each trial began with a printed word (*LISTEN* or *READ*) at the top of the screen. On *LISTEN* trials, participants first heard a recorded question. Then they fixated on a fixation point at the top of the screen, pressed the spacebar to continue, and saw a picture in the center of the screen. They were instructed to describe the pictured event with one sentence mentioning all event characters, as quickly and fluently as possible, but received no further instructions regarding sentence content or form. The procedure was the same on *READ* trials, except that participants saw a printed word on the screen: they had to read the word out loud and decide whether they had already used it in the experiment by saying “yes” or “no.” The experiment began with six practice trials.

Sentence scoring

Sentences produced on target trials were scored as actives, full passives, or truncated passives. Other sentence types (intransitive sentences, get-passives, sentences beginning with a “there is/are” construction or with indefinite pronouns, as well as sentences with restarts) were rejected. In the remaining dataset, trials were excluded if the first fixation in that trial was not on the fixation point at the top of the screen (this resulted in the removal of 90 trials) and if onsets were longer than 5 seconds and 3 standard deviations from the grand mean (44 sentences; onsets were measured from the first word of the sentence, excluding disfluencies). The final dataset consisted of 1760 sentences (.58 actives, .30 passives, and .13 truncated passives). Analyses were conducted only on actives and full passives in both experiments.

Codability scoring

Codability scores were calculated as described by Kuchinsky and Bock (2010) and used as control variables in the analyses. Event Codability scores were based on the distribution of the different verbs used to describe target events across conditions (codability scores did not vary with condition). Character Codability scores were based on the distribution of the different agent and patient names in selected conditions. Because speakers repeated the agent and patient names given in the questions, repetition rates were 1.00 after Agent questions and .99 after Patient questions. Thus, Agent Codability scores were calculated based on responses in the Neutral-question and Patient-question conditions, and Patient Codability scores were calculated based on responses in the Neutral-question and Agent-question conditions. In all cases, lower codability scores (i.e., less heterogeneity in participants' responses) indicate *higher* codability and thus greater ease of encoding.

Event codability did not correlate with either Agent or Patient Codability (r=.02 and .05). Agent and Patient Codability were weakly but reliably correlated (r=.25, p<.05).

Analyses

Analyses were performed in three steps. For comparison with earlier studies, the first analysis tested whether properties of individual items influenced sentence form (actives vs. passives) in the Neutral-question condition (logit mixed models; Jaeger, 2008). The second analysis compared speech onsets across conditions (linear mixed effects models; Baayen et al., 2008). The third set of analyses compared the time-course of formulation for active sentences across conditions (by-participant and by-item Growth Curve Analyses, described in more detail below; Mirman, 2014). A joint analysis for passive sentences from Experiments 1 and 2 is also reported.

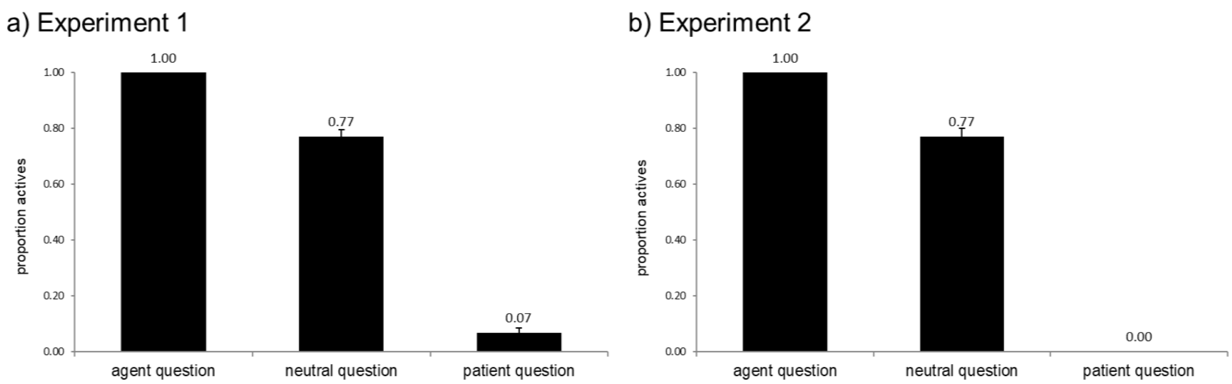
The categorical predictor (Question Condition) was coded with deviation coding. The continuous predictors (Action Informativeness, Event Codability, Agent Codability, and Patient Codability) were converted to z-scores. The models included participants and items as random effects, all three-way interactions between predictors of interest (four-way interactions were not reliable), as well as additive slopes for all factors (Barr, Levy, Scheepers, & Tily, 2013). In case of non-convergence, the random slope(s) for factors accounting for the least variance were removed first. Effects were considered to be reliable at $p<.05$, unless mentioned otherwise.

Results

Sentence form

Consistent with the distribution of responses in earlier studies (e.g., Konopka & Meyer, 2014), speakers produced .77 active sentences in the Neutral-question condition (Figure 1a). Structure choice in this condition was sensitive only to properties of the agent, with more active sentences produced to describe events with higher-codability (easy-to-name) than lower-codability (harder-to-name) agents (.79 vs. .73; β=.64, z=2.12, for the main effect of Agent Codability). Event Codability, Patient Codability and Action Informativeness did not influence sentence form (all *z*s <1.86 for main effects and interactions).

Figure 1. Proportions of active sentences across conditions in Experiments 1 and 2 (by-participant means).



In contrast, the two focus conditions had a radically different distribution of responses: Agent questions elicited exclusively active descriptions (1.00 actives) and Patient questions elicited almost exclusively passive descriptions (.07 actives; an analysis with Question condition as a three-level factor is not possible as one of the conditions had zero variance). Given the strength of these effects, none of the other predictors had any effect on sentence form in the Agent- and Patient-question conditions[7].

Speech onsets

Onsets for active and passive sentences were compared separately across two conditions each (actives: Neutral vs. Agent questions; passives: Neutral vs. Patient questions). Models predicting onsets of active and passive sentences included either Agent or Patient Codability respectively (in addition to Question Condition, Event Codability and Action Informativeness) to verify whether repeating character names given in the questions reduced encoding costs.

Onsets for active sentences (Figure 2a) were indeed shorter after Agent questions than Neutral questions (β=507, t=15.42). In the Neutral-question condition, sentences with higher-

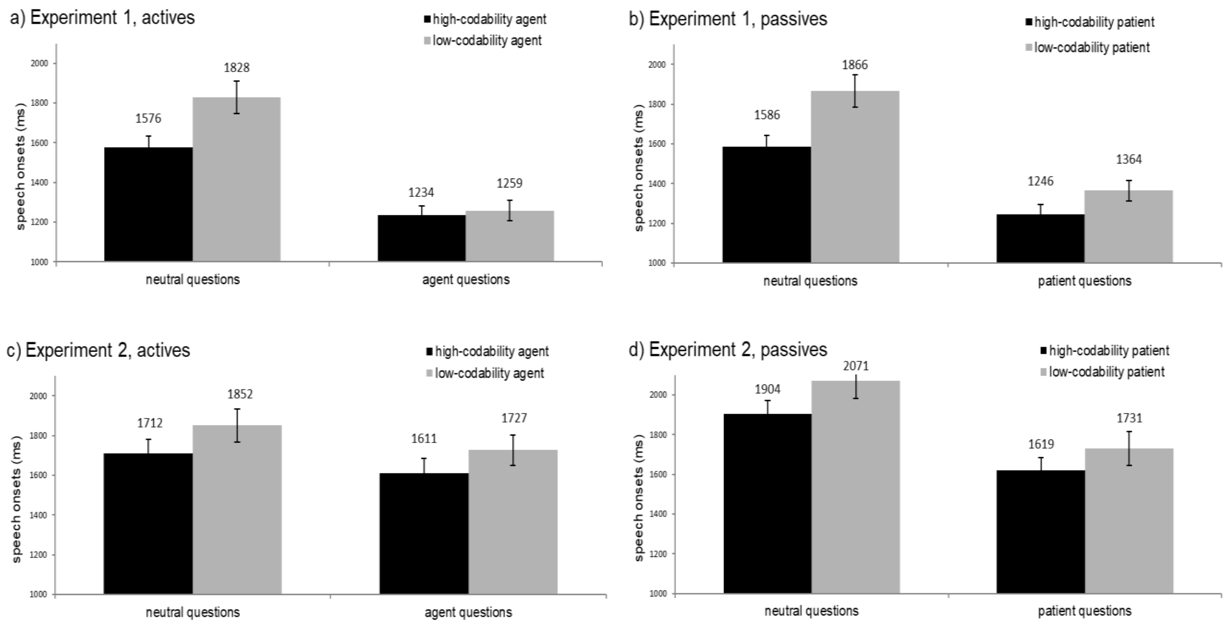
[7] Previous research also evaluated the influence of first character fixations on sentence form. Given the nearly categorical distribution of active and passive sentences across conditions in Experiments 1 and 2, this analysis was not conducted here. The question manipulation also did not influence the distribution of first character fixations (proportions of first fixations directed to the agent in the Neutral, Agent, and Patient conditions: Experiment 1 = .63, .59, and .60; Experiment 2 = .56, .56, and .59 respectively).

codability (easy-to-name) agents had shorter onsets than sentences with lower-codability (harder-to-name) agents, but this difference was nearly eliminated after Agent questions (resulting in an interaction between Agent Codability and Question Condition: β=94, z=3.50).

Similarly, onsets for passive sentences (Figure 2b) were shorter after Patient questions than Neutral questions (β=500, z=9.12). In the Neutral-question condition, sentences with higher-codability (easy-to-name) patients had shorter onsets than sentences with lower-codability (harder-to-name) patients, but this difference was again much smaller after Patient questions (resulting in an interaction between Patient Codability and Question Condition: β=174, z=3.19).

For both active and passive sentences, onsets were also shorter in sentences describing high-codability than lower-codability events (both ts>2), but did not vary with Action Informativeness. Finally, the analysis of passive sentences showed a theoretically uninteresting interaction between Event and Patient Codability (β=76, z=3.00; sentences with low-codability patients were initiated more slowly particularly for low-codability events).

Figure 2. Speech onsets for actives (left) and full passives (right) describing events with higher-codability and lower-codability agents and patients across conditions in Experiments 1 and 2 (participant means). Error bars are standard errors.



Time-course of formulation for active sentences

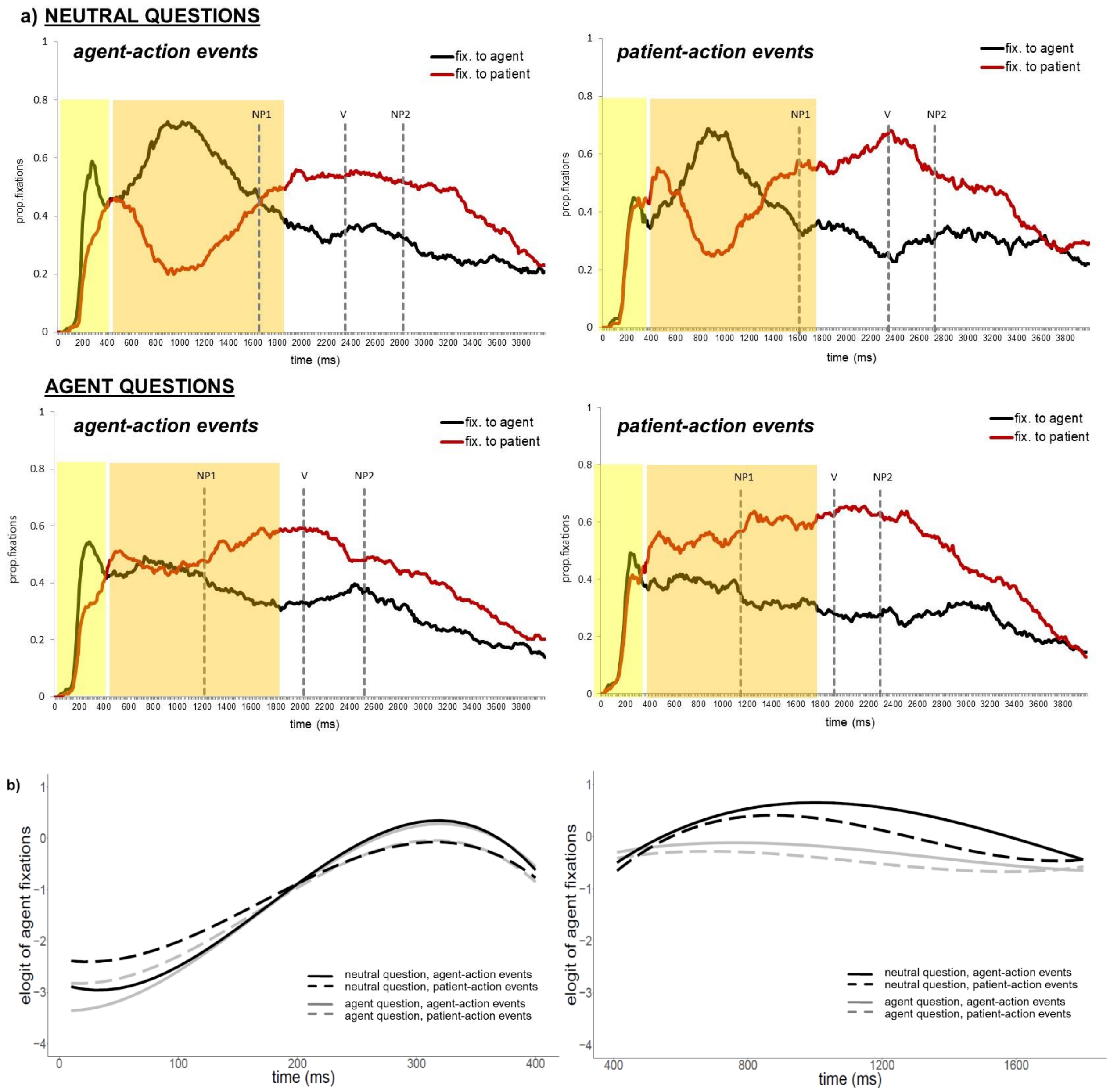
Growth Curve Analyses (GCA) were conducted over two time windows to compare the time-course of sentence formulation after Neutral and Agent questions, and with respect to Action Informativeness. The *first* time window spanned the first 400 ms of each trial (0-400 ms), where eye movements are used to assess the extent of message planning (Gleitman et al., 2007; Griffin & Bock, 2000; Konopka & Meyer, 2014; Konopka et al., 2018). The second time window included fixations between 400 and 1800 ms (speech onset), i.e., the rise and fall of fixations to the subject character in the Neutral-question condition during linguistic encoding of this character's name (as in earlier studies).

Fixations were binned into consecutive 10 ms time samples and an empirical logit was calculated for each time bin by-participants and by-items indexing the log odds of speakers fixating the agent out of the total number of fixations to the agent, patient, and to empty areas on the screen in that time bin. Models report interactions of the Question variable with the Linear time term (indexing the slope of fixations in a given time window), and the Quadratic and Cubic time terms (indexing non-linearities in fixation patterns – specifically, describing the shape of a primary and secondary inflection point in a given time window; Mirman, 2014). Action Informativeness scores were included as a binary predictor in by-participant analyses (*agent-action* vs. *patient-action* events; contrasts were .5 for agent-action events and -.5 for patient-action events) and as a continuous predictor in by-item analyses (agent-action events had positive scores, patient-action events had negative scores). All models included additive random slopes for all factors. The distributions of fixations across conditions and event types are shown in Figure 3a, and model fits are shown in Figure 3b.

0-400 ms (Table 2a). In the 0-400 ms time window, fixations to agents and patients were generally converging (i.e., partially overlapping) in all conditions, although there was a steeper rise and a sharper peak in the Neutral-question condition (interactions of Question Condition with the Linear and Quadratic time terms). This suggests that early viewing patterns were sensitive to discourse focus.

Agent-directed fixations also rose more quickly over time in agent-action events (by-participant interaction of Action Informativeness with the Linear time term), had a more shallow peak (interaction with the Quadratic term) but also a steeper drop right before 400 ms (interaction with the Cubic term) than patient-action events. The three-way interaction of Question Condition and Action Informativeness with the Cubic term showed that fixations in agent-action events did not differ by Question condition, but fixations in patient-action events were more sustained (i.e., less variable) in the Neutral-question condition. Thus, overall, speakers were more likely to quickly but briefly fixate agents when they were more action-informative than patients, demonstrating sensitivity to relational information from the earliest stages of formulation.

Figure 3. (a) Formulation of active sentences in Experiment 1: proportions of fixations to agents and patients after Neutral questions and Agent questions in events with action-informative agents and action-informative patients (*agent-action* vs. *patient-action events*). Dashed lines indicate speech onsets: agents (NP1), verbs (V) and patients (NP2). Analysis windows are highlighted in all figures. (b) By-participant model fits (empirical logits of agent fixations) across conditions in the two analysis windows.



400-1800 ms (Table 2b). Question Condition had a stronger effect on eye movements after 400 ms (Table 2b). In the Neutral-question condition, as in earlier studies, fixations to the agent increased between 400 and 1000 ms, and then began falling after 1000 ms. In contrast, there was no preference for fixating the agent after Agent questions; in fact, agent-directed fixations declined across the entire time window (interaction of Question Condition with the Linear term). Thus, shifts of gaze to the patient occurred much earlier after Agent questions (approximately 1000 ms after picture onset) than after Neutral questions (approximately1500 ms after picture onset). The distribution of agent-directed fixations also had more shallow inflection points in this condition than in the Neutral-question condition (interactions with the Quadratic and Cubic terms).

Eye movements were again sensitive to Action Informativeness. Speakers generally spent more time fixating agents in agent-action than patient-action events, indicating that speakers were encoding verb-related information while fixating the agent. This preference to fixate the action-informative character persisted across the entire time window, and as a result, shifts of gaze to the patient also occurred *later* when the agent was more action-informative than the patient.

Finally, agent-direction fixations dropped more slowly in agent-action events than in patient-action events after Neutral questions than Agent questions (three-way interaction of Question condition and Action Informativeness with the Linear term) and had a taller and sharper peak (by-item interaction with the Quadratic term). Overall, the sharpest rise and then drop in agent-directed fixations was observed in patient-action events in the Neutral-question condition (interaction with the Cubic term). In contrast, the distribution of agent and patient fixations in the Agent question condition was relatively “flat” across the entire time window. As the first *new* content word that speakers had to encode in this condition was the sentence verb, fixations to agents and patients briefly converged in agent-action events while speakers showed a preference for the patient in patient-action events. These effects in the Agent-question condition mark the existence of a “verb-encoding” window.

Discussion

Experiment 1 showed that sentence form and the time-course of sentence formulation can change dramatically with discourse context. As expected, sentence form was largely determined by question type: speakers produced a combination of active and passive sentences in the Neutral-question condition, but the Agent and Patient questions elicited almost exclusively descriptions where the *old* character appeared in subject position (actives after Agent questions, passives after Patient questions). Analyses of speech onsets further showed that prior mention of a character in the questions reduced encoding costs to the point that sentences beginning with easy-to-name and harder-to-name subject characters had comparable onsets. In short, the question manipulation served its intended purpose by influencing structure choice and facilitating encoding of the subject character.

Importantly, these effects were accompanied by large changes in the way speakers *planned* their responses to the questions, allowing an analysis of eye movements that index encoding of relational information. Comparing the time-course of formulation after Neutral questions and Agent questions showed two main results. First, in the 0-400 ms time window, speakers generally distributed their gaze between agents and patients (with a small advantage for the agent) in the Neutral question condition, suggesting some degree of early relational encoding (encoding of the event action) even when producing sentences out of context, consistent with *Hierarchical Incrementality*. Providing questions that shifted the focus of speakers’ responses to relational information in the event (Agent questions) further biased speakers towards hierarchically incremental planning (there was a smaller advantage for the agent after Agent questions, although this effect was numerically quite weak).

Second, effects of the question manipulation on eye movements emerged in time windows normally associated with linguistic encoding (after 400 ms), demonstrating top-down guidance of sentence formulation by a discourse-shaped message. Results from the Neutral-question condition showed *linear deployment of attention*: speakers looked at the two characters in the order of mention. In contrast, formulation of sentences with the same surface form in response to Agent questions showed a less tight coupling of gaze and speech. After 400 ms, there was no evidence of speakers fixating the subject character preferentially, even though subject characters were not visually familiar. Instead, speakers fixated both agents and patients consistently until speech onset (approximately 1200 ms) and then directed their gaze preferentially to the patient. Relative to the Neutral-question condition, this pattern of fixations shows *discourse-driven deployment of attention*: speakers quickly directed their attention to the part of the display they needed to encode *new* information. Crucially, the magnitude of the discourse context effect on eye movements allows identification of a clear “verb-encoding” window before speech onset (400-1200 ms). The convergence of fixations to agents and patients in this time window is consistent with the prediction that speakers need to fixate *both* event characters to encode the sentence verb.

Further evidence for tracking the time-course of relational encoding comes from the comparison of fixations in agent-action and patient-action events. In the Neutral-question condition, speakers were sensitive to relational information from picture onset until speech onset. Most strikingly, they fixated action-relevant characters in the early 0-400 ms time window, even though SVO syntax favors encoding of the subject character before encoding of any relational information. In addition, when speakers directed their gaze to the subject character after 400 ms, the length of gazes on this character still reflected its relevance for encoding relational information: speakers fixated action-informative agents until speech onset but shifted their attention more quickly to action-informative patients. Thus, fixations on the subject character after 400 ms do not index exclusively linguistic encoding of its name: speakers must encode the sentence verb at some point during formulation, and, although there is no clear “verb-encoding” window for these sentences, eye movements show modulation of attention by relational variables *during* linguistic encoding of individual characters.

Effects of Action Informativeness were also observed in the Agent-question condition in the “verb-encoding” window, where speakers showed a preference for fixating the more action-informative character over the less informative character. The timing of these effects confirms that speakers are indeed “zooming in” on the sentence verb in this time window, and it also validates the use of Action Informativeness as a measure of relational encoding in the Neutral-question condition. Overall, the similarity in results observed in early and late time windows with the Action Informativeness variable shows that speakers scan events in the same way when encoding relational information at the message level (*actions*) and at the sentence level (*verbs*).

Experiment 2

The Agent-question and Patient-question conditions in Experiment 1 permitted identification of a verb-specific pattern of eye movements after 400 ms (the “verb-encoding” window) by shifting the message-level focus of speakers’ responses to relational information and providing the subject character name. This reduced speech onsets, and interactions between Question Condition and Character Codability further confirmed that providing the character name in the question reduced encoding costs. In line with these results, the time-course analyses also showed that speakers did not fixate *old* (mentioned) characters preferentially but quickly proceeded to encode the sentence verb after 400 ms.

Experiment 2 aimed to replicate the main findings of Experiment 1 and to test, with a similar manipulation, whether encoding of relational information in the "verb-encoding" window *requires* that speakers distribute their gaze between agents and patients. Since speakers generally tend to look at the things they talk about, agent fixations in this window in active sentences may still partially reflect linear order of mention. For example, it is possible that fixations to the agent in the "verb-encoding" window in active sentences are observed at all only because speakers still fixate the character they are about to name first. This would result in a convergence of fixations to agents and patients that overestimates the extent to which fixating both characters is necessary for relational encoding. In fact, an additional analysis comparing formulation patterns in events with higher-codability and lower-codability agents in Experiment 1 showed that speakers were more likely to fixate harder-to-name agents than easier-to-name agents in the verb-encoding window, even though they simply repeated the agent name provided in the question. Thus, to further "zoom in" on encoding of the verb, Experiment 2 tested whether it is possible to reduce the time spent fixating the subject character after Agent and Patient questions by asking participants to refer to the "old" characters with a *pronoun*, instead of repeating their names. If speakers fixated the subject character during the "verb-encoding" window in Experiment 1 only because they were looking at a character that they were naming overtly, then they should show an earlier shift of gaze to the second character in the "verb encoding" window in Experiment 2.

Method

Participants, materials, design and procedure

Thirty native speakers of Dutch participated for payment. Two were replaced due to low numbers (<25%) of scorable responses.

The materials, design and procedure were identical to Experiment 1. The instructions differed from Experiment 1 in only one respect: participants were told that they should begin their responses to character-specific questions with a pronoun (e.g., *What is the tiger doing? <u>It</u> is scratching the photographer*).

Sentence scoring and analyses

Scoring and analyses were carried out as in Experiment 1. The same exclusion criteria were applied, and 226 additional sentences had to be excluded as speakers mistakenly repeated the name of the character provided in the question instead of using a pronoun. This left 1455 responses for analysis (.61 actives, .32 passives, and .07 truncated passives).

Codability scores were also calculated as in Experiment 1. Again, Event codability was not correlated with either Agent or Patient Codability (r=.04 and .16 respectively); Agent and Patient Codability were weakly but reliably correlated (r=.32).

Results

Sentence form

Speakers produced .77 active sentences on neutral trials (Figure 1b). A model including all three-way interactions between the factors of interest showed that structure choice was predicted by Agent codability more weakly than in Experiment 1 (.84 vs. .71 active descriptions of events with higher-codability and lower-codability agents; β=.84, z=1.59, p=.11, for the main effect of Agent Codability; the model did not include random slopes for Event and Agent codability). The analysis also showed a theoretically uninteresting interaction between Event and Patient Codability (β=1.60, z=2.36): Patient Codability predicted sentence form only in hard-to-describe events (there were more passive descriptions of low-codability events when the patient was easy to name than when it was harder to name). Simplifying the model by removing all (non-significant) three-way interactions showed the expected, statistically reliable effect of Agent Codability on structure choice (β=.88, z=2.19) and an interaction between Agent Codability and Action Informativeness (β=.96, z=2.13): high-codability, action-informative agents were most likely to become subjects.

Importantly, the character-specific questions again elicited descriptions where the contextually *old* character appeared in subject position: actives after Agent questions (1.00 actives) and passives after Patient questions (.00 actives).

Speech onsets

Speech onsets were longer than in Experiment 1 but showed the same pattern across conditions. Active sentences (Figure 2c) were initiated more quickly after Agent questions than Neutral questions (β=165, z=3.87). In the Neutral-question condition, sentences with higher-codability (easy-to-name) agents were initiated more quickly than sentences with lower-codability (harder-to-name) agents; this difference was smaller after Agent questions but the effect was weaker than in Experiment 1 (β=76, z=1.44). Simplifying the model to include only the interaction between Agent Codability and Question Condition showed the expected effect of Agent Codability β=77, z=2.30) and a marginal interaction (β=78, z=1.79, p=.079).

Similarly, passive sentences (Figure 2d) were initiated more quickly after Patient questions than Neutral questions (β=227, z=2.96, for the main effect of Question Condition). In the Neutral-question condition, sentences with higher-codability (easy-to-name) patients were again initiated more quickly than sentences with lower-codability (harder-to-name) patients; this difference was smaller after Patient questions, but the effect was also weaker than in Experiment

1 (β=138, z=1.32). Simplifying the model to include only the interactions between Patient Codability and Question Condition showed the expected effect of Patient Codability β=127, z=2.78) and a marginal interaction (β=142, z=1.89, p=.059).

For both active and passive sentences, speech onsets were also shorter in sentences describing higher-codability than lower-codability events (ts>2.7), but did not vary with Action Informativeness.

Time-course of formulation for active sentences

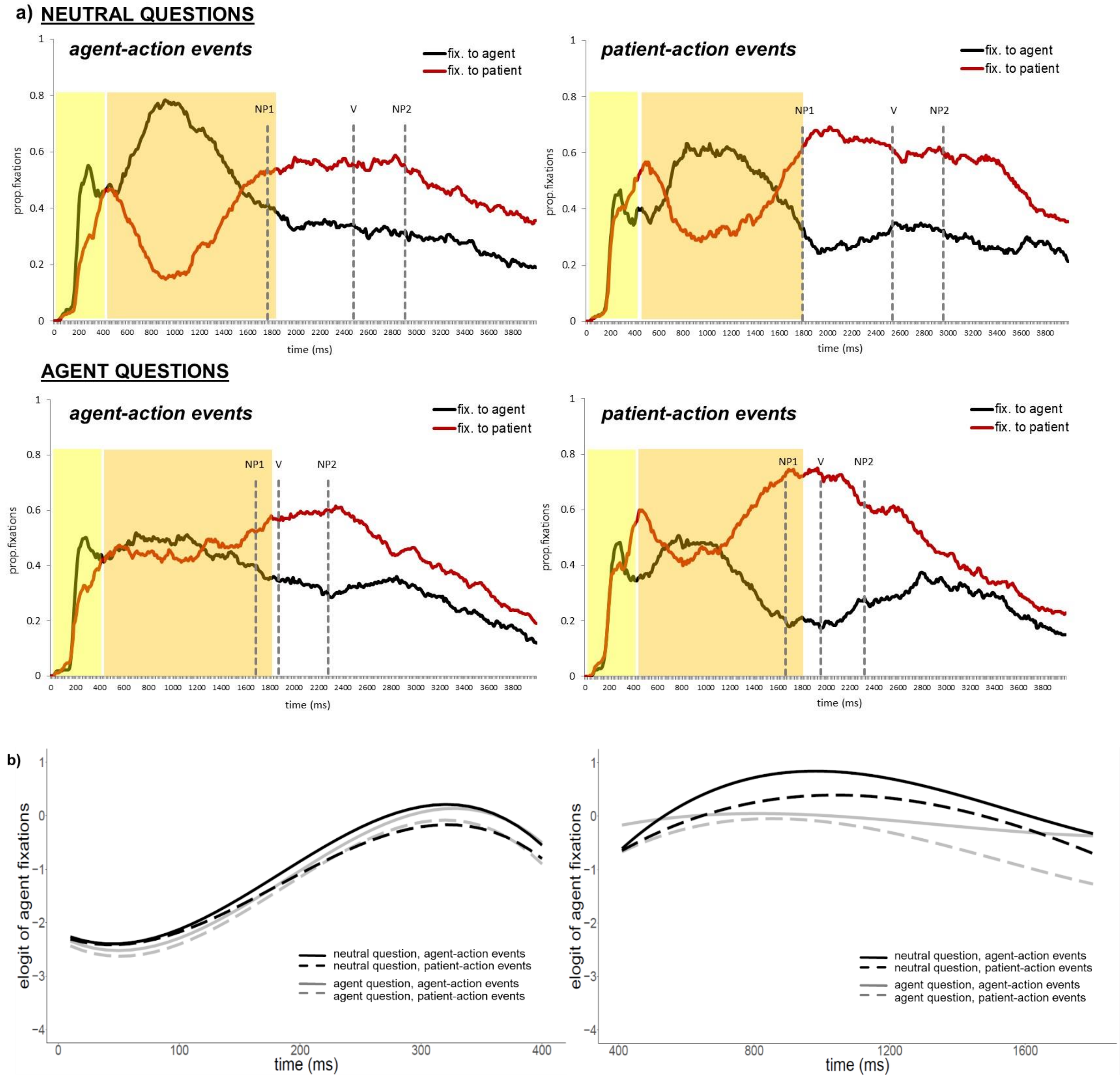
Analyses were carried out as in Experiment 1. Figure 4a shows agent-directed and patient-direction fixations across conditions; model fits are plotted in Figure 4b.

0-400 ms (Table 3a). In the 0-400 ms time window, fixations to agents and patients were generally converging (i.e., partially overlapping) in both the Neutral-question and Agent-question conditions, so differences between conditions were smaller than in Experiment 1.

The analysis replicated effects of Action Informativeness from Experiment 1, albeit the effects were less consistent across the by-participant and by-item analyses. Agent-directed fixations rose more quickly over time in agent-action events (interaction of Action Informativeness with the Linear time term) and had a more shallow peak (by-participant interaction with the Quadratic term), particularly after Agent questions (three-way interactions of Question Condition and Action Informativeness with the Quadratic term). In other words, speakers were again more likely to fixate agents when they were more action-informative than patients, particularly when prompted to “zoom in” on the verb by Agent questions, demonstrating sensitivity to relational information from the earliest stages of formulation.

Figure 4. (a) Formulation of active sentences in Experiment 2: proportions of fixations to agents and patients after Neutral questions and Agent questions in events with action-informative agents

and action-informative patients (*agent-action* vs. *patient-action events*). Dashed lines indicate speech onsets: agents (NP1), verbs (V) and patients (NP2). (b) By-participant model fits (empirical logits of agent fixations) across conditions in the two analysis windows.



400-1800 ms (Table 3b). As in Experiment 1, fixations to the agent increased before 1000 ms and began falling after 1000 ms in the Neutral-question condition, but there was no preference for fixating the agent after Agent questions. Thus, the distribution of agent-directed fixations showed a steeper and steadier decline after Agent questions than Neutral questions in this time window (interactions of Question condition with the Linear and Quadratic terms).

Effects of Action Informativeness were also similar to Experiment 1. Speakers were generally more likely to fixate agents in agent-action events and to look away from agents in patient-action events. This pattern varied with Question condition (resulting in three-way interactions of Question condition and Action Informativeness with the time terms). In the Neutral-question condition, agent-directed fixations had a sharper peak and thus also dropped more quickly in agent-action events than patient-action events. In contrast, in the Agent-question condition, the distribution of agent-directed fixations was generally more flat and had a faster drop towards the end of this window in patient-action events (interactions with the Linear and Quadratic terms). Overall, the largest fluctuations of agent-directed fixations was observed in agent-actions events after Neutral questions and in patient-action events after Agent questions (interaction with the Cubic term).

Discussion

Experiment 2 replicated Experiment 1 with respect to the influence of questions on sentence form, production speed and the time-course of formulation. Structure choice in the Agent-question and Patient-question conditions followed from discourse-level changes in message focus and the acquired *given-new* status of the agent and patient characters. Formulation of active sentences after Agent questions again showed a “verb-encoding” window during which speakers distributed their attention between agents and patients. Speakers were also more likely to fixate characters that were informative for encoding the event action in all conditions in early and late windows, demonstrating continuous sensitivity to relational information.

Importantly, the use of pronouns to refer to subject characters did not change the distribution of fixations to agents and patients in the "verb-encoding" window relative to Experiment 1 (despite longer onsets in Experiment 2). Namely, speakers did not shift their gaze to the patient more quickly in Experiment 2 than in Experiment 1. The similarity in formulation of active sentences after Agent questions in the two experiments suggests that speakers did not fixate the contextually *old* character after 400 ms to encode its name but to encode whatever relational information this character was contributing to the event. This confirms that fixation patterns in the "verb-encoding" window are reliable indicators of relational encoding proper.

Joint analyses for Experiments 1 and 2: active sentences

A joint analysis for Experiments 1 and 2 was carried out to compare the effects of Action Informativeness on formulation of active sentences against *Event Codability* – a variable previously used to track relational encoding in sentences produced out of context (Konopka & Meyer, 2014; Kuchinsky & Bock, 2010). Specifically, previous research showed that speakers devote more time to encoding relational information (event *actions*) in higher-codability (easy-to-describe) events at the outset of formulation (0-400 ms) but that they prefer to postpone encoding of this information in lower-codability (harder-to-describe) events. In the current experiments, the Agent questions induce priority encoding of relational information after 400 ms, so speakers do not have the option of postponing encoding of the verb in sentences describing hard-to-describe events as they may do in the Neutral-question condition. So, if speakers are indeed encoding the sentence verb in the "verb-encoding" window, then effects of event codability on formulation should be observed here as well: speakers should need less time to encode verbs in higher-codability than lower-codability events, resulting in shorter "verb-encoding" windows and faster shifts of gaze to the second character in higher-codability than lower-codability events. Further, effects of event codability should be present over and above any effects of action informativeness, as event codability concerns the ease of expressing the event action irrespective of it being an agent-action or a patient-action.

The joint time-course analysis included two predictors (Action Informativeness, and Event Codability, along with the three Time terms) and was performed over pooled data from the Agent-question condition in both experiments to increase statistical power (Figure 5)[8]. Analyses were performed over the 0-400 ms time window and then over a longer 400-2400 ms time window to assess theoretically interesting differences in fixations one second before and one second after speech onset (approximately 1400 ms). Figures 5a and 5b show the time-course of formulation and model fits respectively.

0-400 ms (Table 4a). As expected, in the 0-400 ms time window, agent-directed fixations in agent-action events had a sharper rise (by-participant interaction of Action Informativeness with the Linear term) and a more shallow peak (interaction with the Quadratic term), but also a sharper drop in fixations towards the end of this window (interaction with the Cubic term) than in patient-action events. This suggests that participants were sensitive to differences in the relational content of the two characters early in the formulation process.

The analysis also showed the expected effects of Event Codability. In events where the action was easy to encode, there was a steeper rise of fixations and then a more shallow peak of fixations in agent-action events than in patient-action events, showing a modulation of fixations by the action informativeness of individual characters; these effects were much weaker in events where the action was harder to encode (interactions of Event Codability and Action

[8] Fixation patterns in higher- and lower-codability events in the Neutral-question condition replicated Konopka and Meyer (2014) and are thus not reported: speakers were more likely to distribute their attention between agents and patients in the early 0-400 ms time window in higher- than lower-codability events.

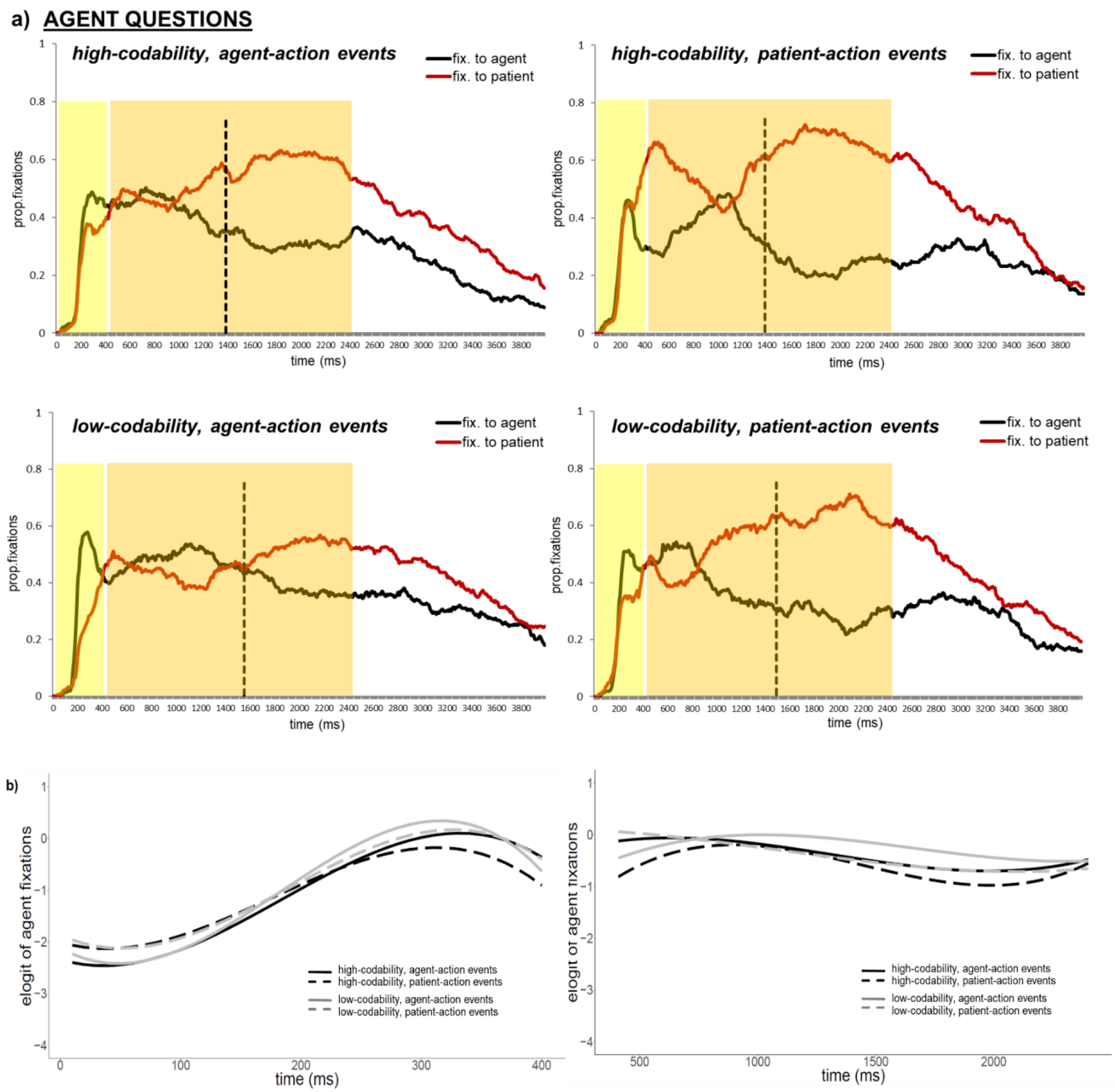
Informativeness with the Linear and Quadratic terms). Towards the end of this time window, there was also a sharper drop in fixations in hard-to-name events with action-informative agents (three-way interaction with the Cubic term). These results are globally consistent with Konopka and Meyer (2014), but as the interactions were reliable only in by-participant analyses, all effects are treated as suggestive.

400-2400 ms (Table 4b). Overall, there were more agent-directed fixations before speech onset (1400 ms) and more patient-directed fixations after speech onset. The timing of gaze shifts from the agent to the patient depended again on Agent Informativeness: speakers were more likely to fixate agents in agent-action events than patient-action events, so agent fixations dropped more slowly in agent-action events (interaction of Action Informativeness with the Linear term) and were more sustained (i.e., less variable; interaction with the Quadratic and Cubic terms) than fixations in patient-action events.

Speakers also began looking away from the agent and towards the patient earlier in higher- than lower-codability events. In higher-codability events, the linear decline in agent fixations was similar in agent-action than patient-action events (likely due to an unexpected rise and fall in agent fixations in patient-action events around 1000 ms). In low-codability events, speakers fixated agents in agent-action events much longer than in patient-action events, and the distribution of agent fixations had a more shallow peak than in patient-action events (three-way interactions of Action Informativeness and Event codability with the Linear and Quadratic terms). The three-way interaction with the Cubic term was due to unexpected fluctuation in high-codability patient-action events.

Figure 5. (a) Formulation of active sentences in Experiments 1 and 2 (pooled data): proportions of fixations to agents and patients after Agent questions in higher-codability and lower-codability events (easy-to-describe vs. harder-to-describe events) with action-informative agents and action-informative patients (*agent-action* vs. *patient-action events*). Dashed lines indicate speech

onsets for agents (NP1). (b) By-participant model fits (empirical logits of agent fixations) across conditions in the two analysis windows.



Discussion

Results from the 0-400 ms time window replicated the effects of Event Codability shown by Konopka and Meyer (2014) and suggest that speakers are sensitive to relational information from the earliest stages of formulation. Results from the later window (400-2400 ms) provide independent evidence that the convergence of fixations to agents and patients observed after 400 ms in the Agent-question condition corresponds to encoding of the sentence verb: speakers shifted their gaze to the second character earlier when the sentence verb was easy to select and later when the verb was harder to select (higher- vs. lower-codability events). This is analogous to results showing that the ease of encoding a *character* or *object name* (i.e., encoding of a *noun*) influences the length of gazes on that character or object (Griffin, 2001; Meyer et al., 1998). Importantly, the analyses showed effects of both Event Codability and Action Informativeness in the "verb-encoding" window, validating the use of both variables as indexes of relational encoding.

Joint analyses for Experiments 1 and 2: passive sentences

Passives were produced on nearly all trials after Patient questions but on less than a third of trials after Neutral questions. Formulation of passives across conditions was thus evaluated after pooling data from both experiments. Figure 6a shows results for passive sentences produced after Neutral questions and Patient questions for agent-action and patient-action events. Analyses were carried out over two time windows: 0-400 ms (as in active sentences) and 400-2000 ms (corresponding to the window during which speakers fixate the subject character in the Neutral-question condition before speech onset). Model fits are shown in Figure 6b.

0-400 ms (Table 5a). Agent-directed fixations had a steeper upward slope after Patient questions than Neutral questions (interaction of Question Condition with the Linear term). They also had a more shallow peak (interaction with the Quadratic term) and a more shallow drop towards the end of this window (resulting in an interaction with the Cubic term).

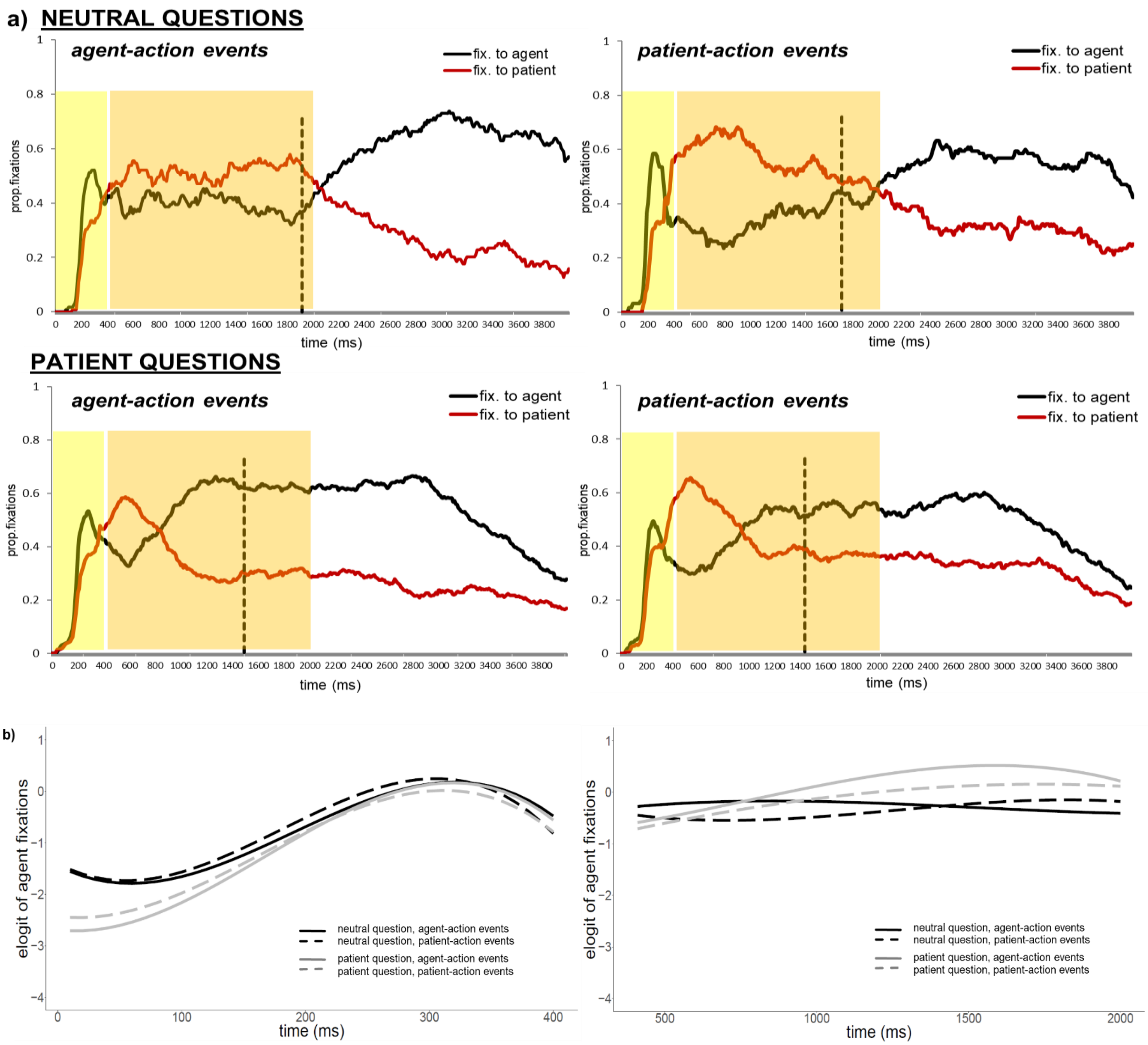
Effects of Action Informativeness were reliable only in by-participant analyses, so all effects are treated as suggestive. Agent-directed fixations had a steeper upward slope, particularly after Patient questions (interaction of Action Informativeness and Question

Condition with the Linear term). Agent-directed fixations also had a more shallow peak in agent-action than patient-action events in the Neutral Question condition (interaction with the Quadratic term).

400-2000 ms (Table 5b). Strong effects of Question Condition on sentence formulation emerged after 400 ms but showed a qualitatively different pattern from results observed with active sentences. In the Neutral-question condition, speakers directed their gaze to the patient (the subject character) and continued fixating this character approximately until speech onset, as expected in sentences produced out of context. In the Patient-question condition, however, speakers looked briefly at the patient, and then switched and maintained their attention on the agent. Thus, by speech onset, speakers were still fixating the patients in the Neutral-question condition but were fixating agents in the Patient-question condition. In other words, agent-directed fixations had a steeper upward slope (interaction of Question Condition with the Linear term), and an earlier and sharper peak in the Patient-question condition than in the Neutral-question condition (interactions with the Quadratic and Cubic term).

Across this entire time window, there were also more fixations to the agent in agent-action events than patient-action events. In the Neutral-question condition, agent-directed fixations in agent-action events remained stable over time; in patient-action events, speakers looked away from the agent after 400 ms and back to the agent by the end of this time window. In the Patient-question condition, agent-directed fixations rose more quickly over time in agent-action than patient-action events (three-way interaction of Question Condition and Action Informativeness with the Linear term). Overall, the largest gaze shifts over time were observed in the Neutral-question condition in patient-action events and in the Patient-question condition in agent-action events (three-way interaction with the Cubic term).

Figure 6. (a) Formulation of passive sentences in Experiments 1 and 2 (pooled data): proportions of fixations to agents and patients after Neutral questions and Patient questions in events with action-informative agents and patients (*agent-action* vs. *patient-action events*). Dashed lines indicate speech onsets for patients (NP1). (b) By-participant model fits (empirical logits of agent fixations) across conditions in the two analysis windows.



Discussion

Formulation of passive sentences in and out of context was only globally similar to formulation of active sentences. Relative to the Neutral-question condition, speakers rapidly shifted their gaze to *new* information in the event after Patient questions, and they directed more fixations to the action-informative character throughout formulation. However, unlike Experiment 1, the results showed *linear deployment of attention* rather than *discourse-driven deployment of attention* during linguistic encoding after Patient questions: there was no clear "verb-encoding" time window in the Patient-question condition because speakers fixated patients and agents sequentially. Preferential fixations to patients after 400 ms in this condition were very brief, but the overall pattern is similar to results from the Neutral-question condition (albeit on a different time scale).

A plausible explanation for differences in formulation of active and passive sentences in a discourse context concerns the large difference in the status of agents and patients. Agents are better "event builders" (i.e., better reference points in event representations) than patients (Cohn & Paczynski, 2009) and thus, arguably, they are also better sentence subjects (in fact, speakers consistently fixate agents before patients at the outset of formulation when producing passive sentences; Van de Velde, et al., 2014; also see Sauppe, 2017). This may influence the likelihood of speakers fixating agents and patients even in sentences produced in strongly-biasing discourse contexts. First, if patients are poorer sentence subjects than agents, speakers may need to devote more processing resources to patients than agents, resulting in more fixations to patients even when they are contextually *old* event characters. This explains the brief but preferential increase in fixations to patients after 400 ms in passive sentences in the current experiments and the absence of an analogous pattern for agents in active sentences. Second, if agents are better "event builders" than patients, a discourse context may produce a stronger shift of attention to the contextually *new* character when this character is an agent than when it is a patient. This may explain why speakers shift their attention to agents immediately after the short-lived period of fixations to the patient instead of fixating both characters in a “verb-encoding” window.

General Discussion

Summary

Two experiments compared the time-course of message and sentence formulation for sentences produced out of context and in a simple discourse context, i.e., in response to neutral questions and to character-specific questions, across a large number of items. The character-specific questions mentioned one of the characters and explicitly asked about the action in the event. The agents and patients in target events differed in their relative contribution of relational information in the event.

The question manipulation influenced both what speakers said (sentence form) and how they prepared to say it (the time-course of formulation). Relative to the Neutral-question condition, Agent questions increased production of active sentences and Patient questions increased production of passive sentences. Thus, the questions categorically shifted the focus of speakers’ responses to the action performed by the character mentioned in the question, leading speakers to begin sentences with the *old* character and to select sentence structures compatible with these message types. More importantly, questions also changed the way speakers assembled sentences online: when responding to Agent questions, formulation of active sentences after 400 ms included a broad window in which speakers clearly distributed their attention between agents and patients (with a small advantage for the more action-informative character) to encode contextually *new* information – i.e., the sentence verb.

The results support two conclusions. The *first* conclusion concerns the time-course of encoding of relational information. Using the fine-grained measure of action informativeness, the results showed that speakers were sensitive to relational information when producing SVO sentences in all conditions – both when they generated sentences out of context (and thus had to produce a new character in subject position) and when they generated sentences in context (and thus could immediately "zoom in" on the sentence verb). Early sensitivity to relational information is consistent with *Hierarchical Incrementality*. Further, the question manipulation allowed identification of a unique "verb-encoding" window after 400 ms in active sentences in the Agent-question conditions. Time-course analyses comparing fixations in sentences with action-informative agents and patients as well as sentences describing higher-codability and lower-codability events confirmed that the distribution of fixations and the length of this window depends on the ease of encoding the sentence verb.

The *second* conclusion concerns effects of context on message and sentence formulation. The question manipulation was an experimental tool for bringing relational information into focus (see Do, Kaiser, & Zhao, 2018, for a different use of questions to track linear encoding of lexical items), but the experiments also show theoretically interesting changes in the time-course of formulation in and out of context. Notably, speakers need not look at all referents in an event in the order of mention (*linear deployment of attention*). In active sentences, there was no evidence of preferential fixations to the first-mentioned character (the agent) after 400 ms in the Agent-question conditions, both when speakers repeated the agent name (Experiment 1) and when they used a pronoun (Experiment 2). Instead, eye movements quickly showed priority encoding of *task-relevant* information (*discourse-relevant deployment of attention*): shifts of gaze were tightly time-locked to encoding of the *verb* and then encoding of the object character. Fixations to agents and patients followed the order of mention more closely only in sentences with the dispreferred passive structure.

Implications for tracking encoding of relational information

When and how do speakers encode relational information? As noted, eye movements across conditions in these experiments show continuous sensitivity to relational information. In the Neutral-question condition, there was evidence of relational encoding in the early 0-400 ms time window, primarily associated with encoding of a pre-linguistic message, and then again between 400 ms and speech onset, i.e., in a time window associated with linguistic encoding of the subject character. The preference to fixate the action-informative character over a less informative character between 0 and 400 ms confirms other demonstrations of early relational encoding in similar sentences (e.g., Konopka & Meyer, 2014; van de Velde et al., 2014) with a new variable. Speakers also fixated the subject character longer between 400 ms and speech onset when this character was more action-informative than less action-informative. Thus, while there may be no unique “verb-encoding” window after 400 ms in SVO sentences produced out of context, fixations to individual characters during “linguistic” encoding (Griffin & Bock, 2000) reflect more than just lexical retrieval of the name of the fixated character.

The Agent-question conditions provide a more explicit demonstration of how relational encoding unfolds at the sentence level, as they highlight a window during which processing appeared to be dedicated primarily to verb encoding. This is the first experimental evidence of the time-course of verb encoding in SVO sentences and it identifies a specific pattern of eye movements indicating priority encoding of relational information at the sentence level.

The results of this condition are consistent with those of studies tracking formulation in *verb-initial* languages. For example, Norcliffe, Konopka, Brown and Levinson (2015) compared formulation of verb-initial (VOS/VSO) sentences in Tzeltal (a Mayan language spoken in Mexico) to verb-medial sentences (SVO) in Tzeltal and Dutch. Their study showed that early placement of the verb in verb-initial sentences results in speakers prioritizing relational encoding from the outset of formulation: as in the current experiments with SVO sentences, encoding of relational information in verb-initial Tzeltal sentences unfolded with speakers distributing their gaze between agents and patients over an extended time window before speech onset. In this sense, the question manipulation in the current experiments “simulated” formulation of *spontaneously produced* verb-initial sentences in Tzeltal using a different paradigm and a different language.

More generally, evidence of sensitivity to relational variables in nearly all time windows and all conditions in these Experiments suggests that speakers are unlikely to plan SVO sentences in a strictly linearly incremental fashion (e.g., Gleitman et al., 2007). Instead, early formulation appears to require some degree of encoding of information that binds the various message elements together into a larger proposition (Bock et al., 2004; Bock & Ferreira, 2014). While it is possible that the questions generally increased the likelihood of speakers attending to this type of information, the results are also consistent with a number of other studies showing an influence of relational variables on early and late eye movements for sentences produced out of context (e.g., Konopka & Meyer, 2014; Van de Velde et al., 2014; Sauppe, 2017). An interesting exception is provided by Hwang and Kaiser (2014) who showed that speakers may postpone encoding of relational information when verbs are produced in *sentence-final* position (e.g., in Korean).

An important goal for future research is to further specify processes involved in relational encoding. The current results show that encoding of relational information requires encoding the relationship *between* characters (leading to a convergence of fixations to agents and patients) as well as encoding of some *character-specific* information (leading to more fixations to action-informative characters). Interestingly, the question manipulation resulted in a clear change in fixations to agents and patients before speech onset, while the effect of Action Informativeness on eye movements was numerically smaller. This suggests that encoding of relational information may primarily require directing attention to multiple parts of an event because relational information is *distributed* across different parts of an event. Since "global" information naturally becomes available earlier than more detailed "local" information (e.g., Greene & Oliva, 2009a, 2009b), global-level relational information encoded by fixating *two* characters may indeed play a stronger role in organizing event representations. However, the specific balance between encoding relational information by fixating two characters and encoding relational information by processing action-relevant properties of each character individually remains to be determined (the two mechanisms may, for instance, have a different time-course in events described with more "general" or more "specific" verbs, e.g., *hit* vs. *thump*).

Implications for gaze-speech coordination

The results also have larger implications for interpreting eye movements and their relationship to cognitive processes during production. In general, production studies draw inferences about what information speakers process with priority by relying on the fact that eye movements closely anticipate linguistic mention. In fact, numerous studies have argued for a high degree of linearity in sentence formulation (albeit at different levels) based on the stability of the gaze-speech link. For example, Gleitman et al. (2007) proposed that linearity applies to the formulation of the preverbal message itself: speakers may generate messages one concept at a time because they tend to fixate each element of a message preferentially shortly before mentioning it in an utterance. Griffin and Bock (2000), on the other hand, proposed that linearity pertains to *linguistic* encoding: speakers first engage in a non-linear process of gist apprehension

(0-400 ms), and then proceed to retrieve the names of individual characters sequentially, in the order of mention (post-400 ms).

The current experiments show a departure from simple linear planning at both levels. First, there was evidence that speakers fixated characters that were informative for encoding the event action both during early formulation (0-400 ms) and during fixations to the subject character (post-400 ms). This argues against the hypothesis that speakers begin formulation by encoding as little as a single character and thus rules against linearity at the *message* level (also see Konopka & Meyer, 2014; Van de Velde et al., 2014). Second, more strikingly, speakers did not fixate the subject character after 400 ms when producing sentences in response to character-specific questions in either experiment: this suggests that formulation does not require tight gaze-speech coordination during linguistic encoding (at least for preferred sentence structures like actives) and thus challenges the assumption of linearity at the *sentence* level.

The lack of a consistent coupling of visual attention and linguistic mention at the sentence level appears to cast some doubt on the reliability of the gaze-speech link. However, both experiments also showed that speakers rapidly directed their attention to encoding *task-relevant* information. Task relevance was defined by linear word order: speakers prioritized linguistic encoding of the sentence verb before linguistic encoding of the sentence object. In addition, the convergence of fixations to agents and patients in the “verb-encoding” window, as well as sensitivity of these fixations to variables like Action Informativeness and Event Codability, shows fairly tight gaze-speech coordination during linguistic encoding of *new*, *task-relevant* information. In fact, the results show that the same principles apply to linguistic encoding of individual characters and to linguistic encoding of relational information: speakers look at *characters* as long as they need to in order to retrieve their names (Griffin, 2000;

Konopka & Meyer, 2014; Meyer et al., 1998) and they also distribute their gaze *between* two characters as long as they need to in order to retrieve a suitable verb. Thus overall, speakers' focus of visual attention is still isomorphic with linear word order, although it is a better indicator of speakers' larger production goals (*discourse-relevant deployment of attention*) than of what they are about to articulate.

Of course, discourse-driven production goals and the preparation of a message element immediately before its articulation may coincide: when speakers direct their attention to a character, they tend to maintain fixations on that character until they start producing its name. At the same time, when a specific communicative goal directs speakers' attention to one aspect of an event, speakers may direct their attention to the task-relevant part of the event long before they will need to mention this information in an utterance. This suggests that the *lag* between linguistic preparation and articulation may vary from context to context.

Further evidence of discourse context influencing gaze-speech coordination comes from Ganushchak, Konopka and Chen (2014) who examined formulation of sentences produced in response to questions with a narrower discourse focus. For example, speakers heard questions like *Who is scratching the photographer?* (agent focus) and *Who is the tiger scratching?* (patient focus). These questions named one of the event characters *as well as* the event action, and thus were expected to shift the focus of speakers' responses only to the new character. As predicted, speakers rapidly directed their gaze to the new character after 400 ms (i.e., during the time window associated with linguistic encoding) and there was little evidence of a "verb-encoding" window as in the current experiments. These results show a stronger departure from linear planning than in this paper and confirm that a close relationship between *looking* and *speaking* is observed when speakers encode *new* information and not when they simply repeat *old* information.

Conclusions

These experiments demonstrate sensitivity to relational information throughout message and sentence formulation, suggesting that relational information may receive priority over non-relational information in speakers' preverbal representations of events and during the message-to-language mapping (i.e., during both message and sentence formulation). This is consistent with accounts emphasizing *hierarchically* incremental planning in production.

The manipulation used to probe sensitivity to relational information in these experiments involved changes in message focus brought about by strong discourse cues, and thus also provides some of the first evidence about how formulation can change with discourse context. In short, when speakers encode events with a specific message focus, visual scan paths show anticipatory fixations to the part of the event containing contextually new, task-relevant information. Thus, to the extent that eye movements index shifts of attention and allocation of processing resources, speakers appear to quickly prioritize encoding of information that meets their communicative goals.

Naturally, question-answer paradigms may reveal only a small part of the flexibility in allocation of resources to task-relevant information that speakers might need during every-day language use. One example of context-sensitive changes in production regards changes in the *content* of speakers' utterances. A visual and/or linguistic context can influence the specificity of object descriptions: e.g., speakers use longer expressions to refer to new referents but produce reduced expressions for contextually old or unambiguous referents (*the blue cup* vs. *the cup*; see Konopka & Brown-Schmidt, 2014, for a review). Effects of conversational history on formulation of sentences with richer or poorer relational content are an important question for future research (e.g., Levinson & Torreira, 2015, and Meyer, Alday, Decuyper, & Knudsen, 2018).

Author note

Many thanks to Annelies van Wijngaarden and student assistants from the Psychology of Language Department (in particular Esther Kroese, Marloes Graauwmans, and Ilse Wagemakers) for help with data collection and processing, and Tess Forest and Antje Meyer for helpful discussions.

References

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

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

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.

Biederman, I. (1972). Perceiving real world scenes. *Science*, 177, 77-80.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143-177.

Biederman, I., Rabinowiz, J., et al. (1974). Information extracted from a glance at a scene. *Journal of Experimental Psychology*, 103, 597-600.

Bock, J. K. (1977). The effect of a pragmatic presupposition on syntactic structure in question answering. *Journal of Verbal Learning and Verbal Behavior, 16*, 723-734.

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

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

Bock, K., & Ferreira, V. S. (2014). *Syntactically Speaking*. In M. Goldrick, V. Ferreira, & M. Miozzo (Ed.), The Oxford Handbook of Language Production (pp. 21-46). Oxford University Press.

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., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. Gernsbacher (Ed.), Handbook of Psycholinguistics (pp. 945-984). San Diego, CA: Academic Press.

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

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*, 170-178.

Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of agents: The processing of semantic roles in visual narrative. *Cognitive Psychology, 67*, 73-97.

Do, M., Kaiser, E., & Zhao, P. (2018). What can wh-questions tell us about real-time language production: Evidence from English and Mandarin. *Proceedings of the Linguistic Society of America, 3, 46*, 1-12.

Dobel, D., Glanemann, R., Kreysa, H., Zwitserlood, P., & Eisenbeis, S. (2010). Visual encoding of coherent and non-coherent scenes. In E. Pedersen & J. Bohnemeyer (Eds.), *Event representation in language: Encoding events at the language cognition interface* (pp. 189-215). Cambridge, England: Cambridge University Press.

Dobel, C., Gumnior, H., Bolte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*, 129-143.

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

Flecken, M., Gerwien, J., Carroll, M. & v. Stutterheim, C. (2015). Analyzing gaze allocation during language planning: a cross-linguistic study on dynamic events. *Language and Cognition, 7,* 138-166.

Ganushchak, L., Konopka, A. E., & Chen, Y. (2014, May). *Focus planning during sentence production.* Poster presented at the 10th International Seminar on Speech Production [ISSP], Cologne, Germany.

Gilette, J., Gleitman, H., Gleitman, L., & Lederer, A. (1999). Human simulations of vocabulary learning. *Cognition, 73*, 135-176.

Givon, T. (1992). The grammar of referential coherence as mental processing instructions. *Linguistics, 30,* 5-55.

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.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science*, *20*, 464-472.

Greene, M. R., & Oliva, A. (2009). Recognition of natural scenes from global properties: Seeing the forest without representing the trees. *Cognitive Psychology, 58*, 137-176.

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

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

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

Hafri, A., Papafragou, A., & Trueswell, J. C. (2012). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, advance online publication. doi: 10.1037/a0030045.

Henderson, F. & Ferreira, M. (2004). *The integration of language, vision, and action: Eye movements and the visual world*. New York: Psychology Press.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review, 16*, 850-856.

Hwang, H., & Kaiser, E. (2014). The role of the verb in grammatical function assignment in English and Korean. *Journal of Experimental Psychology: Learning, Memory and Cognition.* Advance online publication: http://dx.doe.org/10.1037/a0036797.

Hwang, H., & Kaiser, E. (2015). Accessibility effects on production vary cross-linguistically: Evidence from English and Korean. *Journal of Memory and Language*, *84*, 190-204.

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.

Jaeger, T. F., & Norcliffe, E. J. (2009). The cross-linguistic study of sentence production. *Language and Linguistics Compass*, *3*, 866-887.

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

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

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

Konopka, A. E., Meyer, A. S., & Forest, T. A. (2018). Planning to speak in L1 and L2. *Cognitive Psychology*, *102*, 72-104.

Kreysa, H., Zwitzerlood, P., Bolte, J., Glanemann, R., & Dobel, C. (2010). *Task-specificity in allocating attention for sentence production: Naming agents and their actions*. Poster presented at the Language Production Workshop, University of Edinburgh, UK.

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

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

Levinson, S. C., & Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Frontiers in Psychology*, *6*, 731.

MacWhinney, B. (1977). Starting points. *Language*, *53*, 152-168.

MacWhinnney, B. (2005) The emergence of linguistic form in time. *Connection Science*, *17*, 191-211.

MacWhinney, B., & Bates, E. (1978). Sentential devices for conveying givenness and newness: A cross-cultural developmental study. *Journal of Verbal Learning and Verbal Behavior, 17*, 539-558.

Meyer, A. S., Alday, P. M., Decuyper, C., & Knudsen, B. (2018). Working Together: Contributions of Corpus Analyses and Experimental Psycholinguistics to Understanding Conversation. *Frontiers in Psychology*, *9*, 525.

Meyer, A. S., & Dobel, C. (2003). Application of eye tracking in speech production research. In J. Hyona, R. Radach., & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp 253-272). Amsterdam: Elsevier.

Meyer, A. S., Sleiderink, A., & Levelt, W. J. M. (1998). Viewing and naming objects: Eye movements during noun phrase production. *Cognition*, *66*, B25–B33.

Meyer, A. S., Wheeldon, L., van der Meulen, F., & Konopka, A. E. (2012). Effects of speech rate and practice on the allocation of visual attention in multiple object naming. *Frontiers in Psychology, 3*, 1-13.

Momma, S., Slevc, L. R., & Phillips, C. (2016). The timing of verb selection in Japanese sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*, 813-824.

Momma, S., Slevc, L. R., & Phillips, C. (2018). Unaccusativity in sentence production. *Linguistic Inquiry*, *49*, 181-194.

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

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. (2015). Word order affects the time-course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*.

Santesteban, M., Pickering, M. J., Laka, I., & Branigan, H. P. (2015). Effects of case-marking and head position on language production? Evidence from an ergative OV language. *Language, Cognition and Neuroscience*, *30*, 1175-1186.

Sauppe, S. (2017). Word order and voice influence the timing of verb planning in German sentence production. *Frontiers in Psychology*, *8*, 1648.

Sauppe, S., Norcliffe, E., Konopka, A. E., van Valin, R. D., & Levinson, S. (2013). Dependencies first: Eye-tracking evidence from sentence production in Tagalog. *Proceedings of the Annual Meeting of the Cognitive Science Society, 35*.

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

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

Thompson, D., Ling, S. P., Myachykov, A., Ferreira, F. & Scheepers, C. (2013). Patient-related constraints on get- and be-passive uses in English: evidence from el

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

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

Van der Meulen, F. F., Meyer, A. S., & Levelt, W. J. M. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition, 29*, 512-521.

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.

Wheeldon, L. (2013). Producing spoken sentences: The scope of incremental planning. In P. Perrier and P. Lang Verlag (Eds.), *Cognitive and physical models of speech production, speech perception, and production-perception integration*.

Table 2.
GCA results for active sentences in Experiment 1 (* $p<.05$ in all tables).

	By-participants			By-items		
Effect	Est.	SE	*t*	Est.	SE	*t*
a) 0-400 ms						
Intercept	-1.17	.04	-27.66*	-1.15	.09	-13.39*
Linear	6.42	.15	41.54*	5.96	.37	16.23*
Quadratic	-1.87	.27	-6.82*	-1.30	.27	-4.82*
Cubic	-1.75	.16	-10.85*	-1.92	.19	-9.89*
Question Condition	-.11	.07	-1.57	-.01	.06	-.16
Action Informativeness	.02	.08	.28	.11	.09	1.34
Question C. * Action Inf.	-.05	.04	-1.22	-.03	.06	-.47
Linear * Question C.	.70	.13	5.41*	1.07	.10	10.49*
Quadratic * Question C.	-.36	.12	-2.95*	-.31	.10	-3.21*
Cubic * Question C.	.09	.12	.76	.01	.10	.14
Linear * Action Inf.	-2.02	.13	15.54*	.64	.37	1.74
Quadratic * Action Inf.	.37	.12	-3.09*	-.19	.27	-.72
Cubic * Action Inf.	.47	.12	-4.09*	-.23	.20	-1.17
Linear * Question * Action Inf.	.38	.26	-1.46	-.46	.10	-4.55*
Quadr. * Question * Action Inf.	-.03	.24	.14	.25	.10	2.65*
Cubic * Question * Action Inf.	-.53	.23	2.27*	.43	.10	4.56*
b) 400-1800 ms						
Intercept	-.14	.03	-4.55*	-.13	.07	-1.93
Linear	-1.52	.40	-3.85*	-1.53	.56	-2.71*
Quadratic	-1.94	.32	-6.06*	-2.07	.50	-4.14*
Cubic	.89	.26	3.38*	.79	.36	2.22*
Question Condition	-.51	.06	-8.91*	-.53	.05	-9.81*
Action Informativeness	-.22	.05	-4.64*	.08	.07	1.22
Question C. * Action Inf.	.12	.02	6.37*	-.05	.05	-1.01
Linear * Question C.	-.38	.11	-3.54*	-.62	.10	-5.95*
Quadratic * Question C.	2.89	.11	27.19*	4.57	.10	34.58*
Cubic * Question C.	-.66	.11	-6.25*	-.79	.10	-7.77*
Linear * Action Inf.	-.38	.11	3.58*	.20	.57	.36
Quadratic * Action Inf.	.90	.11	-8.48*	.30	.50	.79
Cubic * Action Inf.	.65	.11	-6.11*	-.23	.36	-.64
Linear * Question * Action Inf.	1.27	.21	-5.91*	-.39	.10	-3.82*
Quadr. * Question * Action Inf.	-.32	.21	1.49	.55	.10	5.47*
Cubic * Question * Action Inf.	-.79	.21	3.71*	.49	.10	4.94*

Table 3.
GCA results for active sentences in Experiment 2.

	By-participants			By-items		
Effect	Est.	SE	*t*	Est.	SE	*t*
a) 0-400 ms						
Intercept	-1.21	.04	-28.77*	-1.14	.08	-14.87*
Linear	6.26	.22	28.35*	5.65	.32	17.51*
Quadratic	-.144	.27	-5.36*	-1.07	.24	-4.55*
Cubic	-1.84	.20	-9.35*	-1.85	.18	-10.02*
Question Condition	-.10	.07	-1.39	-.10	.07	-1.43
Action Informativeness	.02	.11	.22	.17	.08	2.17*
Question C. * Action Inf.	-.06	.04	-1.44	-.01	.07	-.10
Linear * Question C.	.13	.13	1.01	.33	.10	3.37*
Quadratic * Question C.	-.05	.13	-.38	.09	.09	1.02
Cubic * Question C.	.05	.12	.41	-.16	.09	-1.76
Linear * Action Inf.	1.18	.13	8.92*	.30	.33	.93
Quadratic * Action Inf.	-.28	.13	-2.21*	-.14	.24	-.57
Cubic * Action Inf.	.04	.12	.31	-.10	.19	-.54
Linear * Question * Action Inf.	-.24	.27	-.90	-.30	.10	-3.03*
Quadr. * Question * Action Inf.	.52	.25	2.09*	.25	.09	2.65*
Cubic * Question * Action Inf.	-.03	.24	-.10	.27	.09	2.95*
b) 400-1800 ms						
Intercept	-.06	.04	-1.79	-.01	.06	-.23
Linear	-1.50	.36	-4.21*	-1.35	.51	-2.66*
Quadratic	-2.76	.38	-7.26*	-2.87	.48	-5.97*
Cubic	.61	.24	2.56*	.65	.30	2.21*
Question Condition	-.45	.05	-8.35*	-.48	.06	-7.74*
Action Informativeness	.28	.05	5.91*	.12	.06	2.06*
Question C. * Action Inf.	-.06	.02	-3.05*	-.09	.06	-1.49
Linear * Question C.	-1.35	.11	-11.92*	-1.48	.10	-14.28*
Quadratic * Question C.	1.96	.11	17.17*	.245	.10	24.57*
Cubic * Question C.	.10	.11	.92	-.17	.10	-1.76
Linear * Action Inf.	.18	.11	1.62	.48	.51	.95
Quadratic * Action Inf.	-.01	.11	-.11	.34	.48	.71
Cubic * Action Inf.	.17	.11	1.49	-.13	.30	-.44
Linear * Question * Action Inf.	3.40	.23	14.86*	1.23	.10	12.20*
Quadr. * Question * Action Inf.	2.38	.23	10.34*	1.63	.10	16.14*
Cubic * Question * Action Inf.	-1.26	.23	-5.56*	-.73	.10	-7.29*

Table 4.
GCA results for active sentences produced after Agent questions in Experiments 1 and 2 (effects of event codability)

	By-participants			By-items		
Effect	Est.	SE	*t*	Est.	SE	*t*
a) 0-400 ms						
Intercept	-1.00	.03	-37.42*	-1.42	.09	-15.36*
Linear	5.26	.14	37.54*	7.59	.35	21.71*
Quadratic	-1.11	.16	-7.01*	-1.70	.31	-5.46*
Cubic	-1.69	.12	-14.19*	-2.08	.21	-10.00*
Event Codability	.13	.06	2.28*	.07	.09	.70
Action Informativeness	-.03	.06	-.45	.15	.10	1.63
Event Cod. * Action Inf.	-.04	.03	-1.41	.02	.09	.25
Linear * Event Cod.	.59	.09	6.56*	.57	.35	1.63
Quadratic * Event Cod.	-.03	.09	-.33	.05	.31	.16
Cubic * Event Cod.	-.46	.09	-5.32*	-.18	.21	-.88
Linear * Action Inf.	1.26	.09	13.82*	.28	.36	.78
Quadratic * Action Inf.	-.17	.09	-2.04*	-.19	.32	-.6
Cubic * Action Inf.	-.30	.09	-3.53*	-.001	.21	-.001
Linear * Event Cod. * Action Inf.	-.98	.18	-5.42*	.02	.09	.02
Quadr. * Event Cod. * Action Inf.	-1.01	.17	-5.83*	.01	.35	-1.01
Cubic * Event Cod. * Action Inf.	-.45	.17	-2.61*	-.32	.31	-.34
b) 400-2400 ms						
Intercept	-.41	.02	-16.47*	.50	.07	-7.31*
Linear	-2.83	.32	-8.92*	-3.57	.58	-6.12*
Quadratic	-.18	.30	-.61	-.10	.52	-.19
Cubic	1.29	.27	4.82*	1.51	.40	3.80*
Event Codability	.15	.04	3.30*	.14	.07	2.09*
Action Informativeness	.19	.04	4.55*	.08	.07	1.11
Event Cod. * Action Inf.	.03	.01	2.39*	.01	.07	.13
Linear * Event Cod.	.32	.09	3.39*	.42	.59	.71
Quadratic * Event Cod.	-.40	.09	-4.26*	-.21	.53	-.39
Cubic * Event Cod.	-1.14	.09	-12.27*	-.87	.40	-2.17*
Linear * Action Inf.	.78	.09	8.33*	.56	.60	.94
Quadratic * Action Inf.	-.56	.09	-5.94*	-.08	.54	-.15
Cubic * Action Inf.	-.42	.09	-4.47*	-.84	.41	-2.05*
Linear * Event Cod. * Action Inf.	2.13	.19	11.42*	.75	.58	1.28
Quadr. * Event Cod. * Action Inf.	-3.56	.19	-19.13*	-.89	.52	-1.70
Cubic * Event Cod. * Action Inf.	2.11	.19	11.32*	.54	.40	1.35

Table 5.
GCA results for passive sentences in Experiments 1 and 2.

	By-participants			By-items		
Effect	Est.	SE	*t*	Est.	SE	*t*
a) 0-400 ms						
Intercept	-.88	.04	-24.57*	-1.10	.09	-12.22*
Linear	4.87	.16	31.35*	5.79	.09	66.25*
Quadratic	-1.39	.17	-8.41*	-1.64	.08	-19.85*
Cubic	-1.65	.13	-12.94*	-2.07	.08	-25.28*
Question Condition	-.31	.08	-3.97*	-.32	.03	-11.58*
Action Informativeness	-.14	.07	-1.92	.11	.09	1.17
Question C. * Action Inf.	.21	.03	6.25*	-.02	.03	-.73
Linear * Question C.	1.49	.10	14.81*	1.60	.17	9.18*
Quadratic * Question C.	-.73	.10	-7.53*	-.53	.17	-3.20*
Cubic * Question C.	.27	.10	2.82*	.39	.16	2.41*
Linear * Action Inf.	.70	.10	7.02*	.04	.09	.44
Quadratic * Action Inf.	.33	.10	3.48*	.03	.08	.37
Cubic * Action Inf.	.10	.09	1.02	.02	.08	.29
Linear * Question * Action Inf.	.67	.20	3.38*	-.18	.17	-1.04
Quadr. * Question * Action Inf.	-.67	.19	-3.47*	-.21	.17	-1.25
Cubic * Question * Action Inf.	-.36	.19	-1.93	-.13	.16	-.79
b) 400-2000 ms						
Intercept	-.12	.03	-3.58*	-.16	.05	-3.01*
Linear	1.88	.35	5.40*	1.96	.53	3.71*
Quadratic	-.82	.25	-3.32*	-1.32	.45	-2.91*
Cubic	-.20	.26	-.80	-.21	.35	-.59
Question Condition	.30	.05	5.62*	.46	.07	6.72*
Action Informativeness	.21	.04	4.85*	.02	.05	.39
Question C. * Action Inf.	.09	.02	5.13*	.07	.07	.95
Linear * Question C.	3.23	.10	31.58*	3.79	.13	29.44*
Quadratic * Question C.	-1.53	.10	-15.01*	-1.99	.13	-15.74*
Cubic * Question C.	-.26	.10	-2.55*	-.25	.13	-1.97
Linear * Action Inf.	-.81	.10	-8.12*	-.36	.53	-.67
Quadratic * Action Inf.	-.74	.10	-7.47*	.64	.46	1.41
Cubic * Action Inf.	.06	.10	.59	.11	.36	.30
Linear * Question * Action Inf.	3.07	.20	15.29*	.86	.13	6.52*
Quadr. * Question * Action Inf.	-.10	.20	-.51	.92	.13	7.00*
Cubic * Question * Action Inf.	-1.21	.20	-6.09*	-.40	.13	-3.09*

Appendix

Events with action-informative agents and action-informative patients (*agent-action* vs. *patient-action events*) in both experiments.

Agent-action events

Man shooting woman
Nurse holding baby
Pirate burying treasure
Assistant painting model
Bee stinging man
Boxer punching cheerleader
Girl hanging sheet
Windmill hitting farmer
Mother dressing boy
Truck towing car
Journalist interviewing actor
Photographer filming model
Girl tickling boy
Shark attacking man
Waiter kicking man out
Mechanic fixing car
Horse kicking farmer
Woman sweeping stairs
Alien pulling astronaut
Eskimo attacking bear
Worker moving piano
Barber shaving prisoner
Cat catching mouse
Ballerina slapped pianist
Paparazzi photographing queen
Man brushing showdog
Cook decorating cake
Lifeguard watching child
Burglar smashing window
Lightning striking church
Punk spraying fence
Baby eating toy bear
Bomb hitting ship
Detective burning drugs
Army attaching castle
Man chopping log
Thief stealing painting
Tailor sketching bride
Santa Claus dragging Christmas tree
Nurse bandaging patient's head
Bodyguard protecting president
Garden hose splashing old woman
Man carving statue
Frog catching fly
Gardener watering sunflower
Monkey holding crab
Cowboy catching sheriff

Patient-action events

Cop arresting boy
Elephant lifting clown
Cop stopping truck
Dog chasing mailman
Train crushing bus
Ambulance hitting woman
Robot smashing computer
Girl pushing boy on sled
Diver pushing paparazzi into pool
Bulldozer destroying building
Bishop crowning king
Fireman saving boy
Crab cutting boy's foot
Girl taking cookie
Mayor unveiling statue
Maid eating chocolate
Tiger scratching photographer
Doctor washing baby
Farmer pulling donkey
Grandfather kissing toddler
Snake biting leopard
Girl throwing away present
Dog guiding blind man
Hippie tripping waiter
Professor congratulating student
Helicopter pulling diver out of sea
Boy letting birds out of a cage
Bird puncturing balloon
Boy scout roasting pig
Owl taking briefcase
Woman massaging man