

This article was downloaded by: [Max Planck Institut Fur Psycholinguistik]

On: 12 March 2015, At: 08:56

Publisher: Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Language, Cognition and Neuroscience

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/plcp21>

Word order affects the time course of sentence formulation in Tzeltal

Elisabeth Norcliffe^a, Agnieszka E. Konopka^{bc}, Penelope Brown^a & Stephen C. Levinson^{acd}

^a Language and Cognition Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

^b Psychology of Language Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands

^c Donders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands

^d Linguistics Department, Radboud University, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands

Published online: 17 Feb 2015.



[Click for updates](#)

To cite this article: Elisabeth Norcliffe, Agnieszka E. Konopka, Penelope Brown & Stephen C. Levinson (2015): Word order affects the time course of sentence formulation in Tzeltal, *Language, Cognition and Neuroscience*, DOI: [10.1080/23273798.2015.1006238](https://doi.org/10.1080/23273798.2015.1006238)

To link to this article: <http://dx.doi.org/10.1080/23273798.2015.1006238>

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the "Content") contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at <http://www.tandfonline.com/page/terms-and-conditions>

Word order affects the time course of sentence formulation in Tzeltal

Elisabeth Norcliffe^{a*}, Agnieszka E. Konopka^{b,c}, Penelope Brown^a and Stephen C. Levinson^{a,c,d}

^aLanguage and Cognition Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands;

^bPsychology of Language Department, Max Planck Institute for Psycholinguistics, Wundtlaan 1, 6525 XD Nijmegen, The Netherlands;

^cDonders Institute for Brain, Cognition and Behaviour, Radboud University, Kapittelweg 29, 6525 EN Nijmegen, The Netherlands;

^dLinguistics Department, Radboud University, Erasmusplein 1, 6525 HT Nijmegen, The Netherlands

The scope of planning during sentence formulation is known to be flexible, as it can be influenced by speakers' communicative goals and language production pressures (among other factors). Two eye-tracked picture description experiments tested whether the time course of formulation is also modulated by grammatical structure and thus whether differences in linear word order across languages affect the breadth and order of conceptual and linguistic encoding operations. Native speakers of Tzeltal [a primarily verb–object–subject (VOS) language] and Dutch [a subject–verb–object (SVO) language] described pictures of transitive events. Analyses compared speakers' choice of sentence structure across events with more accessible and less accessible characters as well as the time course of formulation for sentences with different word orders. Character accessibility influenced subject selection in both languages in subject-initial and subject-final sentences, ruling against a radically incremental formulation process. In Tzeltal, subject-initial word orders were preferred over verb-initial orders when event characters had matching animacy features, suggesting a possible role for similarity-based interference in influencing word order choice. Time course analyses revealed a strong effect of sentence structure on formulation: In subject-initial sentences, in both Tzeltal and Dutch, event characters were largely fixated sequentially, while in verb-initial sentences in Tzeltal, relational information received priority over encoding of either character during the earliest stages of formulation. The results show a tight parallelism between grammatical structure and the order of encoding operations carried out during sentence formulation.

Keywords: incrementality; message formulation; sentence formulation; cross-linguistic comparisons of sentence production; verb-initial languages

To produce an utterance, speakers must transform an abstract thought into a linearly ordered sequence of words that conforms to the grammatical constraints of the target language. According to most models of sentence production (e.g., Levelt, 1989), the first stage of this process involves formulating a *message*, a non-verbal representation of the information speakers want to express. This message must then undergo *linguistic encoding*: Speakers must select and retrieve suitable words to express the individual concepts of the message and must integrate them into a syntactic structure. Subsequently, speakers retrieve phonological information in preparation for articulation.

Language production thus involves a fundamental linearisation of complex hierarchical structures. Yet languages vary widely in their 'basic word order', the most frequent and unmarked order of subject, object and verb in a basic transitive clause. Amongst the rarer word orders, some 5% of languages put the verb first and the subject last [verb–object–subject (VOS) order]. In this paper we examine how the time course of the sentence production process is influenced by the word order and associated grammatical properties of the target language. Specifically, we investigate the processes involved in producing

sentences in Tzeltal, a Mayan language with VOS basic word order, and we compare the formulation process to Dutch, a language with subject–verb–object (SVO) basic word order. Our goal is to test whether and how differences in the linear ordering of constituents in a sentence affect the temporal order by which message-level and sentence-level increments are planned. In doing so, we present the first study of sentence formulation in a verb-initial language, and broach a critical, yet underexplored, theoretical question in production research: to what extent are the processing routines involved in sentence production affected by the grammatical properties of individual languages?

Incrementality and planning scope

Producing a sentence takes time. It is generally agreed that speakers do not wait until processing is completed at all levels of production prior to initiating speech. Instead, most production models assume that planning proceeds incrementally (Ferreira & Swets, 2002; Kempen & Hoenkamp, 1987; Levelt, 1989): As a unit (or increment) of information becomes available at one level of processing, it triggers processing at the next level in the system, potentially all the

*Corresponding author. Email: elisabeth.norcliffe@mpi.nl

way down to articulation. In addition, as *one* increment (e.g., a word or phrase) is passed down to the next level of encoding, speakers may already begin planning the *next* increment. An incremental system of this kind makes sense in terms of both communicative and processing efficiency: Incrementality is argued to help to maintain fluency by allowing speech to be initiated without being preceded by long pauses and to reduce processing costs by allowing speakers to produce already-formulated pieces of an utterance instead of buffering them in working memory until the rest of the utterance is prepared.

A crucial question then is how large these planning units or increments actually are. Studies of planning scope to date have focused on the planning of simple and conjoined noun phrases (e.g., *the arrow and the bag*; Meyer, 1996; Smith & Wheeldon, 2001), modified noun phrases (e.g., *the blue cup*; Brown-Schmidt & Konopka, 2008; Brown-Schmidt & Tanenhaus, 2006) and transitive event descriptions (e.g., *The woman is chasing the chicken*; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Van de Velde, Meyer, & Konopka, 2014). In studies of noun phrase production, planning scope is normally operationalised in terms of the number of words activated before speech onset. In studies of more complex sentence production, the emphasis is primarily on the selection of *starting points* (Bock, Irwin, & Davidson, 2004; MacWhinney, 1977): When constructing a message and preparing to convey this information linguistically, what do speakers encode first? Different accounts of incrementality in production make different predictions in this regard, drawing on key theoretical distinctions in how lexical and structural processes can be coordinated during formulation. We review two accounts in this article and then outline a cross-linguistic comparison that provides new evidence to distinguish between these accounts.

Linear incrementality

The most radical version of incrementality assumes that speakers engage in little or no advanced planning prior to speech onset, even at the message level (Paul, 1886/1970). On this view, formulation begins by encoding the first available concept in the to-be-articulated message, which may then be immediately passed on to lexical encoding processes before speakers plan anything else in the message. For example, when preparing to convey the idea that a woman is chasing a chicken (Figure 1), speakers might begin formulation by conceptualising and lexically encoding the single character *woman*. At the message level, the size of the initial planning unit can therefore be as small as a single nominal concept (a unit isomorphic in size to a single noun; Brown-Schmidt & Konopka, 2008, 2014). Similarly, at the sentence level, sentence formulation may be a highly opportunistic, *lexically driven* process: The



Figure 1. What is happening here?

order of word retrieval (with planning of determiners such as *a* or *the*) is determined by the availability of individual concepts in a message, and the structure of the developing sentence is accordingly constrained by whichever word is retrieved first. Thus, theories that ascribe a pivotal role to lexical items in sentence production (Bock, 1982; Kempen & Hoenkamp, 1987; Levelt, 1989) suggest that linearisation is driven largely by factors influencing the *accessibility* of individual message entities.

Effects of accessibility on sentence form are among the most robust cross-linguistic findings: Speakers systematically make structural choices that allow them to position accessible information earlier in sentences. Accessibility may depend, for example, on a referent's *perceptual* salience and can be enhanced by exogenous, attention-grabbing cues (Gleitman, January, Nappa, & Trueswell, 2007; Ibbotson, Lieven, & Tomasello, 2013; Myachykov & Tomlin 2008; Tomlin, 1995, 1997). Referents can also differ in *conceptual* accessibility, including features such as imageability (Bock & Warren, 1985), givenness (Arnold, Wasow, Losongco, & Ginstrom, 2000) and animacy (Bock, Loebell, & Morey, 1992). Assigning perceptually and conceptually accessible referents to subject position instead of less accessible referents is compatible with the hypothesis that easy-to-name referents are encoded with priority.

The scope of early message and sentence planning has also been assessed more directly using visual-world eye-tracking paradigms, which provide a fine-grained temporal measure of the development of a message and sentence as it unfolds in real time (Gleitman et al., 2007; Griffin & Bock, 2000). In this paradigm, speakers' eye movements are tracked as they describe simple events. Because people tend to look at things they talk about, the timing of gaze shifts between characters in an event is a sensitive index of when the various increments of a message are encoded and how they are combined into a full sentence. Using this method, Gleitman et al. (2007) found that speakers of English can preferentially fixate a perceptually salient character within 200 ms of picture onset and that they tend to select it to be the first mentioned noun in their sentence. This suggests that sentence formulation in English can indeed begin with

priority encoding of as little as a single referent both conceptually and linguistically.

However, if we confine ourselves to English or other subject-initial languages, it is often unclear whether accessibility influences *linear word order* directly or whether it primarily influences *subject assignment* (Bock & Warren, 1985; McDonald et al., 1993), and thus only indirectly word order. A strong or ‘radical’ version of linear incrementality (Gleitman et al., 2007) would hold that accessibility directly drives lexical encoding and that subject assignment *follows* from an early choice to encode one message element linguistically before a different element (e.g., *woman* before *chicken*). The alternative view (described in more detail in the next section) would be that planning the first character and retrieving the first content word (*woman*) involves not only the lexical encoding of one message element (the *woman* character) but also the early selection of a subject – which requires some advanced planning of the relational structure of the event and some grammatical-level processing.

There is some support for both possibilities in studies of languages that allow scrambling and thus where subject position and sentence-initial position are potentially independent. Some studies of word order alternations have found that conceptual accessibility can directly affect word order, even when grammatical function (subjecthood) is controlled for (Branigan & Feleki, 1999, for Greek; Ferreira & Yoshita, 2003, for Japanese; Kempen & Harbusch, 2004, for German; MacWhinney & Bates, 1978, for Italian and Hungarian). Other work, by contrast, has found that accessible concepts are more likely to become subjects, rather than simply sentence-initial increments (Christianson & Ferreira, 2005, for Odawa). Yet other evidence suggests that within a language both word order *and* grammatical function assignment may be influenced by conceptual accessibility (Tanaka, Branigan, McLean, & Pickering, 2011, for Japanese).

Structural incrementality

While linear incrementality involves the piecemeal formulation of *parts* of messages, an alternative view postulates the upfront planning of the *relational wholes* of messages: Formulation begins with the generation of a larger conceptual representation of the message, where information is tied together by an abstract, relational scheme (Wundt, 1900/1970). In the current example (Figure 1), this view predicts that speakers first conceptualise a *chasing* event in which one character is acting on another, and defer linguistic formulation until after the relational structure of the message has been generated. Advance planning of the relational structure then allows for the early generation of a structural sentence frame, which in turn guides the order of subsequent lexical

retrieval processes (i.e., the retrieval of the words *woman*, *chase* and *chicken*).

Thus, like the linearly incremental account, the structural account assumes that sentence formulation can proceed incrementally (word by word), but rather than being driven by the availability of individual words, a sentence is built *out* from a structural plan that reflects the relational scheme of the message (Griffin & Bock, 2000; Lee, Brown-Schmidt, & Watson, 2013). This view accords with theories that assume that structure-building may operate independently of lexical processes (Bock, 1990; Chang, Dell, & Bock, 2006; Christianson & Ferreira, 2005; Dell, 1986; Fisher, 2002; Konopka & Bock, 2009; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010).

Empirical support for structure-driven formulation also comes from visual-world eye-tracking studies where speakers describe simple events. Griffin and Bock (2000) report evidence for an initial phase after picture onset (0–400 ms) during which speakers do not preferentially fixate either character in the depicted events. The authors interpret this as evidence of a non-linguistic ‘gist apprehension’ phase, in which speakers encode the relationship between event characters before directing their gaze preferentially to the first character they will mention. On this account, upfront gist apprehension allows for the generation of a structural frame, which in turn guides the order of lexical retrieval processes. Thus, speakers look to the character they will mention first not because their attention was initially drawn to it (contrary to Gleitman et al., 2007), but rather because their eyes were guided there by the structural framework generated shortly after picture onset (see also Bock, Irwin, Davidson, & Levelt, 2003).

The influence of word order on message and sentence formulation

In short, a variety of evidence has been brought to bear on the question of the time course of sentence formulation, but so far little consensus has been reached with respect to the size of planning units at the message and sentence levels, or on the temporal coordination of conceptual, lexical and structural processes. These conflicting findings suggest that the time course of message and sentence formulation may be flexible. In this regard, there is mounting evidence that multiple factors can influence breadth of planning (in English). Some of these are extralinguistic, relating, for example, to time pressure (Ferreira & Swets 2002) or to individual differences in working memory capacity (Swets, Jacobina, & Gerrig, 2008). Others concern production processes proper, for example, the relative ease of formulating a message plan (Kuchinsky & Bock, 2010) or resource constraints affecting the coordination of lexical and structural processes (Konopka, 2012; Konopka & Meyer, 2014).

Here we focus on an additional factor that might influence the time course of formulation: the grammatical structure of the language itself. To what extent might reliance on different planning strategies be driven by grammar? It is of course in some ways self-evident that planning processes must be affected by language-specific constraints, given that the target structures of linguistic encoding are language-specific. A key question for theories of incrementality, however, is how far up in the production system language-specific grammatical properties influence formulation.

To date, sentence production research has been undertaken on a limited group of languages, especially English (see Jaeger & Norcliffe 2009 for a review). Crucially, all languages investigated thus far share a common structural property: subjects come before verbs in simple sentences. It is therefore hard to empirically tease apart the two incrementality accounts outlined earlier, and to assess the extent to which a given formulation strategy might be more or less contingent on word order. Languages with *verb-initial* word order provide an important contrast: In order to produce a verb-initial sentence, relational information presumably *must* be planned early in order to retrieve an appropriate sentence-initial verb. Comparing the time course of sentence formulation for verb-initial and subject-initial sentences therefore allows us to assess how message-level and sentence-level encoding operations are affected by the position of the subject and the verb in a target sentence.

Current experiments

In two matched experiments we compare the time course of sentence formulation in two typologically different languages. In Experiment 1, we investigate whether the formulation of transitive sentences (e.g., a description of an event in which a woman is chasing a chicken; Figure 1) is influenced by linear word order in Tzeltal, a language whose basic word order is VOS: verbs are positioned *before* their arguments and subjects come last in the sentence. Tzeltal also optionally permits SVO word order, allowing for a within-language contrast of how sentence formulation can vary as a consequence of both subject position and verb position.

We outline the most relevant grammatical properties of Tzeltal for present purposes in more detail below (for a full grammatical description of the language, see Polian, 2013), and then describe the results of an eye-tracked picture description experiment (Experiment 1). The methodology is similar to that of earlier event description studies (Griffin & Bock, 2000; Konopka & Meyer, 2014) and allows for two types of analyses that, jointly, assess how word order affects the time course of formulation. In the first set of analyses, we examine how speakers' structural choices (voice and word order) are affected by

the conceptual and perceptual accessibility of event characters. This provides an initial measure of how speakers begin to formulate sentences. In the second set of analyses, we compare fixation patterns to event characters across different sentence types over time.

For a direct comparison against a subject-initial language, we then report results from the same production experiment carried out with native speakers of Dutch (Experiment 2). Together, the two experiments, carried out with two very different populations, provide a strong test of the effects of grammatical structure on the time course of sentence formulation.

Experiment 1

Tzeltal

Tzeltal is a Mayan language spoken in the Mexican state of Chiapas by over 400,000 people (Polian, 2013). In active sentences, Tzeltal's basic word order is VOS (or verb–patient–agent, VPA [1]): the grammatical subject comes sentence-finally. The grammar also permits subject-initial SVO word ordering (or agent–verb–patient, AVP [2]¹) where the grammatical subject comes sentence-initially. According to one small corpus study of Tzeltal based on a collection of spoken and written narrative texts (495 active transitive clauses in total; Robinson, 2002), VOS word order is twice as frequent as SVO order (66% vs. 31%).²

[1] ya s-nutz me'mut te antze
 ASP 3SG-chase chicken the woman
 The woman is chasing a chicken
 (VPA [VOS] word order)

[2] te antze ya s-nutz me'mut
 the woman ASP 3SG-chase chicken
 The woman is chasing a chicken
 (AVP [SVO] word order)

Tzeltal also has a passive voice construction, in which the verb is marked with the suffix *-ot*: the patient becomes the subject, while the agent becomes oblique and may or may not be marked by *yu'un*, a *by*-phrase. For the passive voice, the most typical word ordering is verb–agent–patient (VAP [3]) with sentence-final subject placement (the patient is now the subject). However, patient–verb–agent (PVA) word order [4] with sentence-initial subject placement is also possible.

[3] ya x-lek'-ot (y-u'un) tz'i' te kereme
 ASP ASP-lick-PASS (by) dog the boy
 The boy is being licked by a dog
 VAP [VOS] word order

[4] te kereme ya x-lek'-ot (y-u'un) tz'i'
 the boy ASP ASP-lick-PASS (by) dog
 The boy is being licked by a dog
 PVA [SVO] word order

Passives are less frequent than actives (Robinson, 2002). However, the passive has been described as being strongly preferred over the active for ‘non-canonical’ animacy configurations, that is, where the patient ‘out-ranks’ the agent in terms of animacy (when the patient is human or animate and the agent is non-human or inanimate; Polian, 2013; see also Aissen, 1997, for the closely related language Tzotzil).

Tzeltal does not mark case on verbal arguments. Rather, it is a ‘head-marking’ language: verbs carry agreement markers indexing the grammatical roles of their arguments. The agreement marking is sensitive to transitivity (it is ergatively aligned): e.g., third person subjects of transitive verbs are marked on the verb by the prefix *s-* (see examples [1] and [2]), while third person subjects of intransitive verbs and third person objects of transitive verbs are unmarked.

Task and predictions: how formulation of sentences with different word orders addresses questions about incremental planning

Native speakers of Tzeltal described pictures of simple transitive events involving familiar characters and actions (e.g., Figure 1) while their gaze and speech were recorded. They were instructed to produce a short description (the equivalent of one sentence) for each picture, but were otherwise free to produce any descriptions they wanted.

Analyses focused on three questions.

1. *Conceptual accessibility and structure choice: ‘radical’ linear incrementality or subject-selection?* First, we test how conceptual accessibility influences structure choices in Tzeltal by assessing the effects of character animacy on speakers’ choice of active or passive syntax in two analyses. As noted, the literature on English and subject-initial languages confounds assignment of a character to the first slot in the sentence (a strictly linearly incremental process) with selection of a sentence subject (a planning process requiring more extensive encoding of the entire event).

If Tzeltal speakers prefer to select accessible characters to be subjects (as do speakers of SVO languages), they should produce more active sentences to describe events with human agents and more passive sentences to describe events with human patients. Crucially, we test whether this applies regardless of word order (i.e., regardless of whether the subject comes first or last in the sentence). If conceptual salience influences the choice between active and passive syntax only in subject-initial sentences, this would indicate that conceptual accessibility only influences the timing of word retrieval, consistent with linear incrementality (Branigan & Feleki, 1999; Gleitman et al., 2007; Kempen & Harbusch, 2004). If, however, human characters are preferentially selected to be subjects in subject-final structures as well, this would indicate an

effect of conceptual accessibility on *subject selection* proper (Bock & Warren 1985; McDonald et al., 1993) and thus constitute evidence of advanced structural planning early in the formulation process. A third possibility is that conceptual accessibility influences word order *as well as* subject selection (see Tanaka et al., 2011, for Japanese). In this case, the effect of conceptual salience on active vs. passive syntax should be stronger in subject-initial sentences, where a salient character can be mentioned first, than verb-initial sentences, where both characters follow the verb. Thus, for example, events with a human agent and a non-human patient should be described more often with subject-initial active sentences than with subject-initial passive sentences; this difference should be smaller in verb-initial sentences.

We also test whether conceptual salience affects the choice *between* the dominant verb-initial and less frequent subject-initial word orders. One possibility is that speakers should produce more subject-initial sentences when the event contains a conceptually accessible referent (e.g., a human agent or patient) and more verb-initial sentences when the event contains referents that do not differ in conceptual accessibility (e.g., two human characters or two non-human characters). This is because the presence of *one* accessible referent should facilitate retrieval of one character name before the other character name and thus trigger a linearly incremental formulation (noun-first) strategy. In contrast, the presence of referents that do not differ in accessibility should favour an encoding strategy where speakers delay encoding of the two characters by producing the verb first. An alternative possibility is that Tzeltal speakers’ word order choices are affected by a preference for *minimising interference* (Gennari, Mirković, & MacDonald, 2012). Gennari and colleagues argue that similarity of two entities on a relevant conceptual dimension (such as animacy) increases the potential for interference and/or increases processing load, and suggest that speakers might prefer to reduce interference by making structural choices that avoid adjacent placement of conceptually similar elements. This makes the inverse prediction from the previous one: Events containing referents that do not differ in their conceptual features should be described *more often* with subject-initial sentences compared to events containing referents that are conceptually dissimilar. To decide between these possibilities, we test whether the different combinations of agent and patient animacy across target events influence the choice between subject-initial and verb-initial word order.

2. *Perceptual accessibility and structure choice: ‘radical’ linear incrementality or subject selection?* We next test whether sentence structure can be predicted from early attention shifts (i.e., the order in which speakers fixate the two characters at picture onset). As noted, linear incrementality predicts that speakers begin formulation by

prioritising conceptual and linguistic encoding of a single perceptually salient referent (Gleitman et al., 2007). Thus, we compare speakers' choice of sentence structure (active vs. passive syntax) on trials where first fixations are directed to agents and trials where first fixations are directed to patients. Analogous to the predictions listed earlier for conceptual accessibility, if fixation order influences subject selection, then speakers should select first-fixated characters to be sentence subjects more often than characters that are fixated later: i.e., speakers should produce more active sentences if they fixate the agent before the patient (e.g., the *woman* before the *chicken*) and more passive sentences if they fixate the patient before the agent (the *chicken* before the *woman*). Once again, we test whether this holds regardless of word order (i.e., regardless of whether the subject comes first or last in the sentence). We also assess the effects of character animacy and first fixations in a joint analysis to compare the relative strength of conceptual and perceptual accessibility in structure selection.

3. *Time course of formulation for verb-initial and subject-initial sentences: does grammatical structure determine when speakers encode the verb and the subject?* We examine the time course of formulation in active and passive verb-initial and subject-initial sentences by comparing the distribution of fixations to agents and patients in four sentence types [1–4] over a 3-second window. Within this window, we test whether early placement of the verb in verb-initial sentences changes the time of planning relational information compared to subject-initial sentences. If sentence structure mediates the relationship between the uptake of visual information in an event and the formulation of an event description, then early mention of the verb should result in earlier encoding of relational information ([1] and [3]) than in subject-initial sentences ([2] and [4]). We hypothesise that encoding the verb would require that speakers distribute their gaze between the two characters (as relational information is presumably 'distributed' across characters in an event), so differences in formulation of different sentence types can be investigated by examining patterns of divergence or convergence of fixations to agents and patients before speech onset.

Importantly, we investigate how early such effects arise. If formulation of verb-initial and subject-initial sentences differs from the outset of formulation, then the distribution of agent and patient fixations should show a high degree of compatibility with linguistic structure immediately after picture onset (0–400 ms): fixations to agents and patients should diverge slowly in verb-initial sentences and more rapidly in subject-initial sentences. For verb-initial sentences, this pattern would suggest that early verb mention rapidly induces or facilitates deployment of a processing strategy that prioritises encoding of relational information, in preparation for producing the verb. For subject-initial

sentences, rapid divergence of fixations to agents and patients would suggest that early subject mention favours a processing strategy where encoding of a single message element (agent if active, patient if passive) is sufficient. In contrast, if sentence structure does not influence early formulation, then the distribution of agent-directed and patient-directed fixations should not differ between verb-initial and subject-initial sentences in the first 400 ms of picture viewing. Word order should only shape the distribution of fixations after 400 ms, i.e., in time windows associated with linguistic encoding.

Method

Participants

Fifty-three native Tzeltal speakers from the indigenous Mayan community of Majosik' (Tenejapa, Chiapas, Mexico) participated for payment (27 female, mean age = 28, range = 16–47). Their educational background and level of bilingualism were assessed with a short questionnaire. 25 speakers reported receiving some primary school education (*primaria*, grades 1–6), 14 had completed middle school (*secundaria*, grades 7–9) and 14 had completed high school (*preparatoria*, grades 10–12). 18 participants described themselves as monolingual Tzeltal speakers, 20 claimed a little knowledge of Spanish and 15 described themselves as proficient in Spanish.

Materials and design

Target pictures consisted of 52 coloured line drawings of two-character transitive events (Figure 1). The animacy of the characters varied across events; the key contrast was between *human* and *non-human* characters: 14 events showed human agents acting on human patients, 12 showed human agents acting on animal patients, 16 showed non-human agents (11 animal agents, 5 inanimate agents³) acting on human patients and 10 showed non-human agents acting on animal patients (see Appendix). In 20 of the events, the agent carried an instrument (e.g. a woman tickling a girl *with a feather*).⁴ All action/agent/patient combinations were unique. There were two mirror-reversed versions of each target picture, one in which the agent appeared on the left-hand side and one in which it appeared on the right-hand side of the picture.

Two experimental lists were created by counterbalancing the two versions of the target pictures across lists and interspersing these pictures in a list of 90 unrelated filler pictures, for a total of 142 trials. Within lists, there was at least one filler picture between any two target trials.

Procedure

Participants were tested individually in a quiet room using a Tobii T120 eye tracker (120 Hz sampling frequency)

controlled by a Panasonic CF-FP computer. Instructions were provided in Tzeltal by a native speaker assistant. Participants were told that they would have to produce short descriptions of pictured events. Prior to the presentation of each picture, a fixation point appeared at the top of the screen: participants were instructed to look at the fixation point and the experimenter clicked with the mouse to continue.

To familiarise participants with the task, the experiment began with a training session. Participants saw nine filler pictures and heard pre-recorded Tzeltal descriptions of these events. They then saw the same pictures again and were asked to describe them aloud. The experiment began after the training session was completed. Responses were later transcribed by native speakers.

Sentence scoring

Sentences produced on target trials were scored as actives, full passives, truncated passives or responses with other constructions. The latter category included intransitive sentences and incomplete sentences, which were excluded from all analyses. Responses were also excluded if the first fixation in that trial fell on either the agent or the patient instead of the fixation point at the top of the screen (resulting in the exclusion of 427 responses) or if the first fixation directed to a character occurred only 400 ms or later after picture onset (177 additional responses).⁵ This left 1133 sentences for analysis. Among the four most common sentence types in which both characters were mentioned (951 sentences), responses were also excluded if onsets were longer than six seconds and three standard deviations from the grand mean (resulting in the exclusion of 43 sentences). The final data-set consisted of 908 sentences: 179 subject-initial actives (AVP word order), 392 verb-initial actives (VPA word order), 49 subject-initial passives (PVA word order) and 288 verb-initial passives (VAP word order).

Time course analyses were carried out for the subset of active and passive sentences where speakers mentioned both characters but omitted instruments (in all events, instruments were considered to be part of the agent interest area). Among the four most common sentence types (932 sentences), responses were also excluded if onsets were longer than three standard deviations from the grand mean (33 sentences). The final data-set for the time course analyses thus consisted of 899 sentences: 174 subject-initial actives (AVP word order), 382 verb-initial actives (VPA word order), 48 subject-initial passives (PVA word order) and 295 verb-initial passives (VAP word order).

Analyses

Analyses of structure choice (active vs. passive structures, and verb-initial vs. subject-initial structures) were

conducted with mixed logit models in R, after centring all predictors (Baayen, Davidson, & Bates, 2008; Jaeger, 2008). The models included Agent and Patient Animacy (human vs. non-human) as fixed factors and random intercepts for participants and items. The effect of first fixations on voice choice was tested in conjunction with these factors in separate models. Time course analyses are described in more detail below.

All models tested for theoretically relevant effects and interactions. Random slopes for fixed factors were included where mentioned only if they improved model fit (cf. Barr et al., 2013) at $p < .05$ (evaluated via backward model comparison).

Results

Distribution of responses

Speakers produced more active sentences than passive sentences (.63 vs. .37). Verb-initial sentences were also produced more often than subject-initial sentences (.77 vs. .23), consistent with the reported dominance of verb-initial syntax in Tzeltal. There were more actives both within the verb-initial and subject-initial sentence types (.58 and .79). Instruments were infrequently mentioned (.05 sentences).⁶

Conceptual accessibility and structure choice: 'radical' linear incrementality or subject-selection?

The first analysis compared the effects of Agent and Patient Animacy on voice (active vs. passive). As expected, speakers produced more active sentences to describe events with human agents than non-human agents (.94 vs. .31), and fewer active sentences to describe events with human patients than non-human patients (.25 vs. .92). Figure 2 shows that actives were especially infrequent for events featuring a human patient and a non-human agent (.05 vs. $\geq .76$ for all other event types; Polian, 2013).

A model that included humanness of Agent, humanness of Patient and Word Order (verb-initial vs. subject-initial) as predictors, together with all two-way interaction terms, showed effects of all three predictors on voice type (Table 1). Notably, there was no interaction between Word Order and Agent Animacy or between Word Order and Patient Animacy, indicating that speakers preferred to make human characters the subject of their sentence (i.e., choosing active constructions when the agent was human, and passive constructions when the patient was human), regardless of whether the subject was positioned first or last in the sentence (subject-initial or verb-initial).

The second analysis tested for the influence of conceptual accessibility on word order (verb-initial vs. subject-initial structures) by comparing production of verb-initial and subject-initial sentences for the different

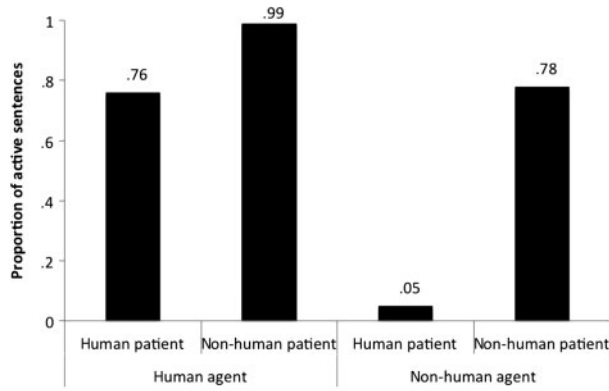


Figure 2. Proportions of active sentences produced in Tzeltal (Experiment 1), collapsing across verb position, as a function of agent and patient animacy.

agent-patient animacy configurations. Because we collapsed over voice type, this analysis assessed the effect of *subject* animacy (agents in active sentences and patients in passive sentences) and *object* animacy (patients in active sentences and agents in passive sentences) on word order choice. As established in the first analysis in this article, sentences combining a non-human subject and a human object were very infrequent (18 tokens in all), so the second analysis was restricted to the remaining three animacy combinations: Non-human subject + Non-human object; Human subject + Non-human object; and Non-human subject + Non-human object.

Figure 3 shows the proportions of verb-initial sentences for the different animacy combinations. Verb-initial word order was *most* frequently produced when the subject was human and the object was non-human (the rightmost bar of Figure 3), less frequently produced when neither character was human (the middle bar of Figure 3), and least frequently produced when both characters were human (the first bar of Figure 3).

Table 1. Results of regression comparing productions of active vs. passive sentences, given the animacy of the agent and the patient (human vs. non-human), and the word order produced (verb-initial vs. subject-initial) in Tzeltal (Experiment 1).

Effect	Est.	SE	z-value
Intercept	1.83	0.53	3.44*
Animacy of agent = <i>human</i>	5.57	0.95	5.82*
Animacy of patient = <i>human</i>	-6.42	0.97	-6.61*
Word order = <i>verb-initial</i>	-3.38	0.86	-3.90*
Animacy of agent = <i>human</i> * Animacy of patient = <i>human</i>	2.10	1.94	1.08
Animacy of agent = <i>human</i> * Word order = <i>verb-initial</i>	-0.17	1.28	-0.13
Animacy of patient = <i>human</i> * Word order = <i>verb-initial</i>	-0.86	1.48	-0.58

Note: * $p < .05$

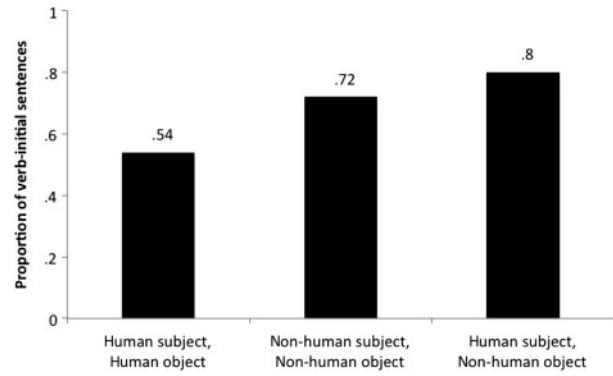


Figure 3. Proportions of verb-initial (vs. subject-initial) sentences in Tzeltal (Experiment 1) with respect to subject and object animacy.

Differences across items were assessed in a new model including a three-level treatment-coded animacy factor. In the model, animacy-matched events (i.e., events with Human subjects + Human objects and events with Non-human subjects + Non-human objects) were significantly less likely to be described with verb-initial word order compared to Human subject + Non-human object events (Table 2). Thus, contrary to the predictions of linear incrementality, subject-initial sentences were not produced more often when there was a single accessible referent in the event to facilitate word retrieval. Instead, the choice to position one character at the beginning of the sentence appears to be conditioned by whether or not it *matched in animacy* with the other character. This result is consistent with Gennari et al.'s (2012) proposal that speakers may prefer to separate conceptually similar referents to reduce interference. We return to this point in the General Discussion section.

Perceptual accessibility and structure choice: 'radical' linear incrementality or subject-selection?

Speakers were more likely to direct their attention to agents than patients at picture onset (.74 vs. .25). Human agents attracted more early fixations (.81) than non-human

Table 2. Results of regression comparing productions of verb-initial vs. subject-initial sentences in Tzeltal (Experiment 1), given the combined animacy features of the subject and object (reference level: human subject + non-human object).

Effect	Est.	SE	z-value
Intercept	11.34	2.27	4.98*
Non-human subject + Non-human object	-2.69	0.89	-3.00*
Human subject + Human object	-2.55	1.01	-2.53*

Note: * $p < .05$

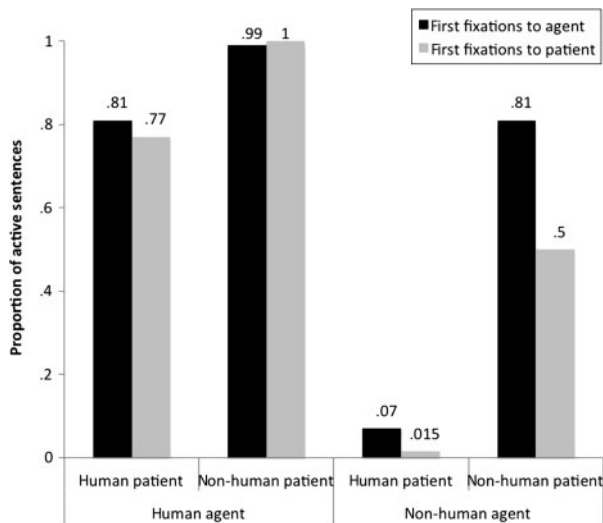


Figure 4. Effect of first fixations on production of active and passive sentences in Tzeltal (Experiment 1).

agents (.67) and human patients attracted more early fixations (.38) than non-human patients (.14).

Speakers produced more active sentences when the first character fixation was directed to the agent than when it was directed to the patient (.70 vs. .39). However, first fixations were not reliable predictors of sentence form when character animacy was taken into account (Figure 4; Table 3). Specifically, testing all two-way interactions between Agent and Patient Animacy, First Character

Table 3. Results of regression comparing productions of active vs. passive sentences in Tzeltal (Experiment 1), given first fixation (to agent vs. to patient), animacy of the agent and the patient (human vs. non-human), and the word order produced (verb-initial vs. subject-initial).

Effect	Est.	SE	z-value
Intercept	1.77	0.50	3.50*
Animacy of agent = <i>human</i>	5.34	0.90	5.90*
Animacy of patient = <i>human</i>	-6.12	0.92	-6.59*
Word order = <i>verb-initial</i>	-3.27	0.85	-3.82*
First fixation = <i>agent</i>	0.47	0.64	0.73
Animacy of agent = <i>human</i> * Animacy of patient = <i>human</i>	1.56	1.84	0.85
Animacy of agent = <i>human</i> * Word order = <i>verb-initial</i>	-0.06	1.30	-0.05
Animacy of patient = <i>human</i> * Word order = <i>verb-initial</i>	-1.04	1.49	-0.69
Animacy of agent = <i>human</i> * First fixation = <i>agent</i>	-1.74	1.49	-1.16
Animacy of patient = <i>human</i> * First fixation = <i>agent</i>	0.83	1.57	0.53
Word order = <i>verb-initial</i> * First fixation = <i>agent</i>	-1.26	1.31	-0.96

Note: * $p < .05$

Fixations and Word Order showed the expected main effects of Agent and Patient Animacy but no effect of First Character Fixations. Moreover, including First Character Fixations in the model did not reliably improve model fit. Thus, while accessible agent and patient characters were more likely to become subjects, the order in which they were fixated did not additionally influence their assignment to subject position, for either word order.

Time course of formulation

Figure 5a and Figure 5b show the time course of formulation for subject-initial and verb-initial active and passive sentences. Formulation of *subject-initial* sentences (Figure 5a) was similar to earlier results obtained with SVO languages. When producing active AVP sentences, Tzeltal speakers quickly directed their gaze to the agent (the grammatical subject) and continued fixating this character preferentially until speech onset; shifts of gaze to the patient (the grammatical object) occurred only after speech onset. Despite sparse data, a similar pattern was observed with passive sentences. Speakers first directed their gaze to the patient and were generally more likely to fixate the patient than the agent before speech onset. Shifts of gaze to the agent occurred again after speech onset. Thus, in both active and passive sentences, the subject character was the initial focus of attention. In contrast, Figure 5b shows that formulation of *verb-initial* sentences deviates dramatically from this pattern: speakers' attention and gaze was more evenly distributed across the two characters before speech onset, with an advantage for the agent regardless of voice.

Three sets of analyses were carried out to compare formulation of subject-initial and verb-initial sentences. Voice and the sequential order of the agent and patient are confounded in this data-set, so the analyses first compared active and passive sentences with different agent-patient word orders, and then sentences with similar agent-patient word order but different voice. Specifically, effects of early verb production within each sentence type were first tested by comparing the distribution of agent-directed fixations across the two types of active sentences (VPA and AVP word orders) and the two types of passive sentences (VAP and PVA word orders) separately. Second, to compare sentences with the same linear order of the two characters, complementary analyses were carried for the two types of sentences with agent-patient word order (active AVP sentences and passive VAP sentences) and the two types of sentences with patient-agent word order (active VPA sentences and passive PVA sentences). Third, we compared formulation of the two types of verb-initial sentences (active VPA vs. passive VAP) to test when speakers begin to encode agents and patients when the verb is produced first.

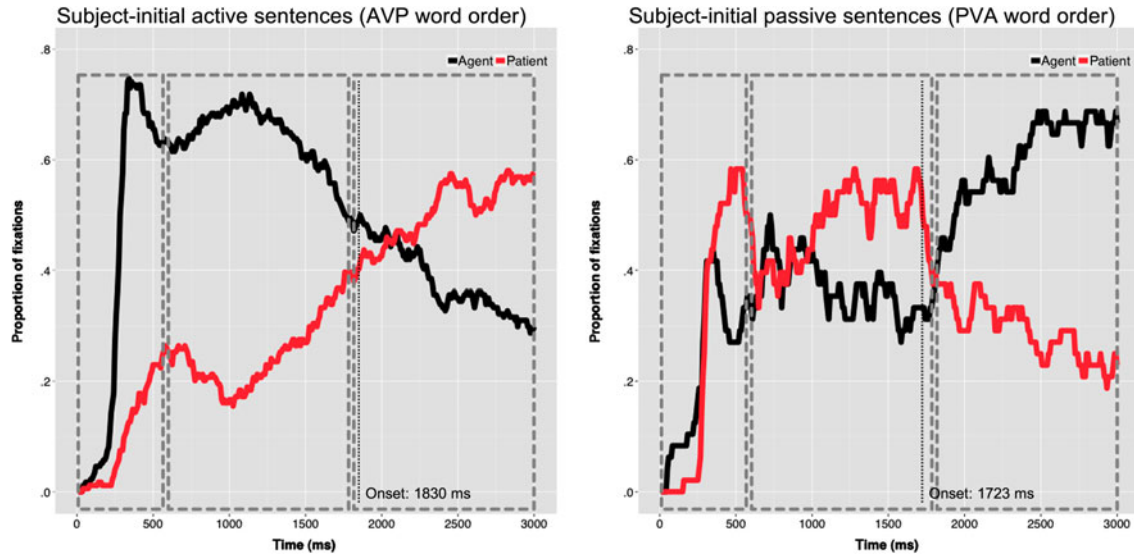


Figure 5a. Proportion of agent-directed and patient-directed fixations in active and passive subject-initial sentences in Tzeltal (Experiment 1).

Analyses were by-participant and by-item quasi-logistic regressions (Barr, 2008). Eye position was sampled every 8.3 ms, and samples were then aggregated into 200 ms time bins for the analyses. An empirical logit was calculated reflecting the log odds of speakers fixating agents in each time bin from the total number of fixations observed in that bin (fixations to the agent, patient, and to empty areas on the screen). Each analysis was performed over three time windows, chosen based on three theoretically important processing distinctions. The first time window included the period between 0 ms (picture onset) and 600 ms that arguably corresponds to event apprehension (encoding of the relational structure of the event;

Griffin & Bock, 2000).⁷ Fixations in this time window were aggregated into three consecutive 200 ms bins. The two subsequent time windows included the period between 600 ms and 3000 ms that is normally associated with linguistic encoding: 600–1800 ms (speech onset) and then 1800–3000 ms, after aggregating data into six consecutive 200 ms bins for each analysis.

All models included the Time variable (Time Bin) and either Word Order (character order) or Voice (active vs. passive). In all cases, to arrive at the simplest best-fitting models, full models including all interactions between factors were simplified to leave only interactions that improved model fit relative to an additive model at $p < .10$

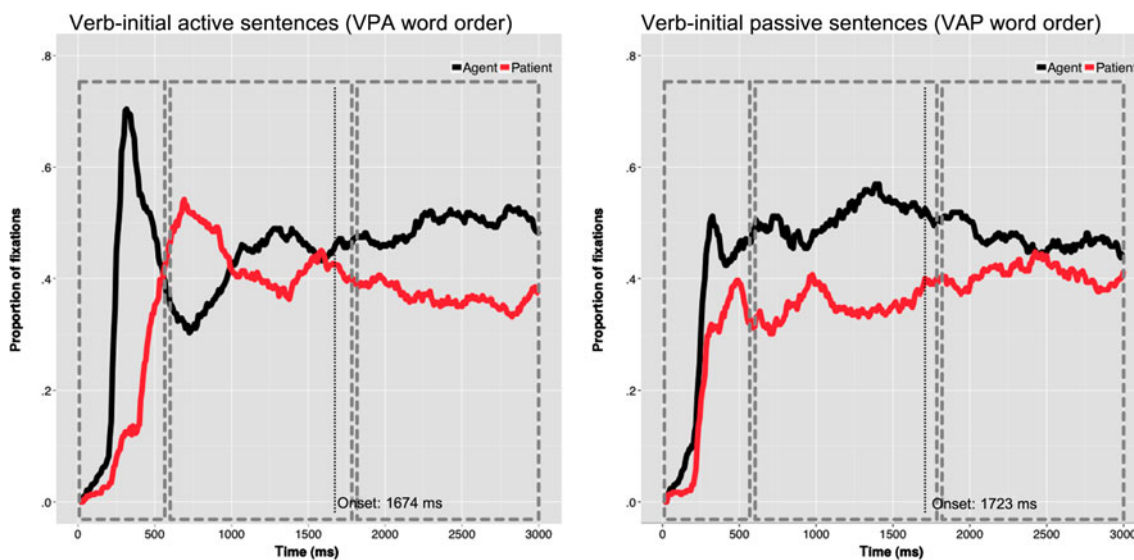


Figure 5b. Proportion of agent-directed and patient-directed fixations in active and passive verb-initial sentences in Tzeltal (Experiment 1).

Table 4. Results of regressions comparing fixations to the agent in verb-initial and subject-initial active sentences (VPA and AVP word orders, respectively) in Tzeltal (Experiment 1). (*s*) indicates the inclusion of random slopes.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i> -value	Est.	SE	<i>t</i> -value
<i>a) 0–600 ms</i>						
Intercept	-.87	.03	-36.61*	-.98	.04	-27.36*
Time bin	2.93	.08	37.42*	3.56	.12	30.06*
Word order	.50	.05	9.64*	.40	.07	6.01*
Time bin * Word order	-2.92	.17	-16.98*	-2.39	.23	-10.46*
<i>b) 600–1800 ms</i>						
Intercept	-.25	.02	-16.15*	.18	.02	9.79*
Time bin	.34 (<i>s</i>)	.02	15.06*	-.06 (<i>s</i>)	.03	-2.15*
Word order	-1.24	.03	-38.07*	-1.14	.03	-33.36*
Time bin * Word order	.81	.05	16.65*	.28	.02	11.79*
<i>c) 1800–3000 ms</i>						
Intercept	-.16	.01	-13.80*	-.10	.01	-6.68*
Time bin	.02 (<i>s</i>)	.02	1.08	-.20 (<i>s</i>)	.02	-10.51*
Word order	-.10	.02	-4.25*	-.04	.03	-1.40*
Time bin * Word order	.96	.04	23.35*	.63	.04	16.69*

Note: * $p < .05$

and that were reliable at $p\text{MCMC} < .05$ (for models without random slopes). Random slopes for fixed factors were included only if they improved model fit. Main effects of the Word Order and Voice variables indicate differences across conditions at the start of a given time window; interactions with Time show whether or not the slope of the fixation functions changed over time in subsequent bins in that time window. Results from the 0–600 ms time window are interpreted primarily in terms of the presence or absence of interactions with Time as theoretically interesting differences began emerging after the first 0–200 ms bin.

First analysis: comparing formulation of active and passive sentences

Active sentences (AVP vs. VPA word orders). Speakers rapidly directed their gaze to the agent after picture onset. When the agent was produced first (AVP), fixations to the agent remained stable until 600 ms in subject-initial sentences (Figure 5a); in contrast in verb-initial sentences, where the agent was produced later in the sentence (VPA), looks to the agent declined rapidly after 300 ms (Figure 5b). This resulted in an interaction between Time Bin and Word Order in the analysis of the 0–600 ms time window (Table 4a).

Carrying over from the first time window, there were more fixations to the agent in subject-initial than verb-initial sentences at 600–800 ms (a main effect of Word Order; Table 4b). Between 600 ms and 1800 ms (speech onset), speakers then continued fixating the agent in subject-initial sentences, suggesting preferential linguistic

encoding of the subject character, and shifted their gaze away from this character around speech onset. In contrast, formulation of verb-initial sentences continued with speakers distributing their attention roughly equally between the agent and the patient, suggesting that they continued encoding information about both characters to select a suitable verb. The sharp decline in fixations in subject-initial but not verb-initial sentences before 1800 ms resulted in an interaction between Time Bin and Word Order.

Finally, fixations observed between 1800 and 3000 ms showed that speakers fixated the two characters in the order of mention in both sentence types: fixations were directed to the patient in subject-initial sentences (AVP) and to the agent in verb-initial sentences (VPA), resulting again in an interaction of Time Bin and Word Order (Table 4c).

Passive sentences (PVA vs. VAP word orders). Formulation of passive sentences showed similar, but numerically smaller effects. As expected, early fixations (0–600 ms) were directed to the patient in subject-initial sentences (PVA; Figure 5a). Compared to formulation of subject-initial active sentences, the preference for fixating the first-mentioned character over the second character was smaller – likely due to sparse data as well as to the fact that patients are generally fixated less often than agents at the outset of formulation (e.g., Konopka & Meyer, 2014; Van de Velde et al., 2014; also see Cohn & Paczynski, 2013, for a review). More importantly, formulation of verb-initial passive sentences showed a different pattern,

Table 5. Results of regressions comparing fixations to the agent in verb-initial and subject-initial passive sentences (VAP and PVA word orders, respectively) in Tzeltal (Experiment 1). (*s*) indicates the inclusion of random slopes.

Effect	By-participants			By-items		
	Est.	SE	<i>t</i> -value	Est.	SE	<i>t</i> -value
<i>a) 0–600 ms</i>						
Intercept	–1.53	.03	–52.78*	–1.46	.04	–33.23*
Time bin	3.69	.10	38.61*	3.87	.12	33.33*
Word order	.24	.04	8.09*	.19	.05	3.51*
Time bin * Word order	–	–	–	–	–	–
<i>b) 600–1800 ms</i>						
Intercept	–.12	.01	–8.10*	–.20	.02	–8.74
Time bin	.12 (<i>s</i>)	.02	5.01*	.14 (<i>s</i>)	.03	4.65*
Word order	.50	.02	23.47*	.54	.03	17.41*
Time bin * Word order	–	–	–	–	–	–
<i>c) 1800–3000 ms</i>						
Intercept	–.05	.02	–2.74*	–.11	.03	–4.16*
Time bin	.05 (<i>s</i>)	.03	1.99*	.27 (<i>s</i>)	.03	8.48*
Word order	.05	.05	1.06	.14	.04	3.43*
Time bin * Word order	–.95	.07	–13.77*	–1.07	.06	–17.99*

Note: **p* < .05

with speakers fixating the agent more often than the patient (VAP; Figure 5b). This difference was present at the beginning of the 0–600 ms time window (a main effect of Word Order; Table 5a) and did not change over time (there was no interaction with Time Bin). The same pattern was observed between 600 and 1800 ms (main effect of Word Order but no interaction with Time Bin; Table 5b).

Finally, speakers showed a strong preference for fixating the two characters in the order of mention after speech onset (1800–3000 ms): They quickly directed more fixations to the agent when it was mentioned last (PVA, subject-initial sentences) than when the patient was mentioned last (VAP, subject-initial sentences), producing an interaction of Time Bin with Word Order (Table 5c).

Second analysis: comparing formulation of sentences with the same order of arguments

Sentences with agent–patient word order (actives vs. passives). The second set of analyses compared formulation of subject-initial and verb-initial sentences with the same relative ordering of agents and patients, i.e., active AVP sentences and passive VAP sentences. Analyses of the first time window were restricted to 200–600 ms and showed main effects of Word Order (all *t*s > 13) and no interactions with Time: speakers were more likely to fixate agents within 200 ms of picture onset when agents were produced at the beginning of the sentence (AVP) than when they were produced after the verb (VAP), and this difference persisted over the entire time window.

Between 600 and 1800 ms, speakers were also more likely to fixate agents in AVP than VAP sentences. There were large differences in agent-directed fixations in the two types of sentences at 600–800 ms (all *t*s > 19 for the main effect of Word Order). Fixations to agents then declined rapidly in AVP sentences by 1800 ms (all *t*s > 6 for the interaction with Time Bin).

Together with the separate analyses of active and passive sentences outlined earlier, these results provide converging evidence that early placement of the verb influences the degree to which speakers prioritise encoding of one character over information about both characters before speech onset.

Sentences with patient–agent word order (actives vs. passives). Similarly, in sentences where patients were produced before agents (active VPA sentences and passive PVA sentences), speakers were less likely to look at agents before 600 ms when the verb was produced at the beginning of the sentence (VPA) than when it was produced later (PVA; the main effect of Word Order was reliable by-participants, *t* > 22, but marginal by-items), and there were no interactions with Time Bin (this analysis was performed over a 200–600 ms time window due to sparse data for passives). In the 600–1800 ms time window, speakers were also less likely to look at agents in PVA sentences than in VPA sentences (all *t*s > 5 for the main effect of Word Order; the interaction with Time Bin was reliable only in the by-participant analysis, *t* > 11). Thus again, verb placement influenced the timing of encoding information about the two characters before speech onset.

Third analysis: comparing formulation of verb-initial active and passive sentences

The final analysis compared formulation of verb-initial active and passive sentences (VPA actives vs. VAP passives). On the hypothesis that early production of the verb results in encoding of relational information over an extended window, this analysis tested whether fixations after 600 ms are consistent only with encoding of the verb, or whether they also show allocation of resources to the first-mentioned argument. The results were consistent with the latter hypothesis. Specifically, at 600 ms, speakers directed more fixations to the agent in VAP passive sentences, where the agent is produced after the verb, than VPA passive sentences, where the agent is produced last (all $t_s > 11$ for the main effect of Voice in the 600–1800 ms time window). An interaction with Time Bin was present only in the by-participant analysis ($t = 15.42$), showing that the difference in agent-directed and patient-directed fixations increased over time.

Discussion

In describing pictures of simple events, Tzeltal speakers' choice of voice and word order were influenced by the animacy of the characters shown in the target events, but not by where speakers first directed their attention. Speakers were more likely to describe events with active sentences when the agent in the event was human, and more likely to describe events with passive sentences when the patient in the event was human. The animacy effects held across both word orders (verb-initial and subject-initial sentences), showing that accessible entities tend to be selected to be *subjects* in Tzeltal, even when subjects are produced last in the sentence. In terms of planning scope, this implies that early sentence formulation in Tzeltal involves a high degree of advance planning, requiring identification of both characters and determining their animacy as well as selection of one of the two characters as the subject of the sentence (consistent with structural incrementality).

With respect to the choice *between* subject-initial and verb-initial word orders, speakers did not show a preference for subject-initial structures when one or both of the characters in the depicted events was human, compared to events in which neither character was human. This suggests that the choice to utter a subject-initial sentence instead of a verb-initial sentence is not driven primarily by the availability of a nominal concept that could trigger early word retrieval. Instead, subject-initial constructions were produced more often when the subject and the object had matching animacy features (i.e., when both were either human or non-human). Besides providing an explanation for why Tzeltal speakers might switch between verb-initial and subject-initial constructions, this result also speaks against the possibility of a radically

linearly incremental production process. It suggests that one factor driving the choice of word order (subject-initial vs. subject-final) is a preference to avoid interference (by separating entities with similar conceptual features; Genari et al., 2012). This, again, implies a degree of planning of *both* entities at the outset of formulation, consistent with structural incrementality.

In contrast to the effects of conceptual accessibility, perceptual accessibility (i.e., first fixations) did not influence structure choice. Tzeltal speakers were not more likely to begin their sentences with whichever referent had first attracted their attention, further supporting the view that structure choice in Tzeltal is not the outcome of a radically linear incremental formulation process.

Finally, time course analyses showed effects of verb placement on formulation from the earliest time windows. Subject-initial sentences were formulated in a similar way to English sentences with the same word order (Gleitman et al., 2007; Griffin & Bock 2000; Kuchinsky & Bock, 2010): Formulation began with fast divergence of fixations to the two characters, and was followed by a wide time window in which speakers fixated preferentially the first-mentioned character before speech onset, and ended with preferential fixations to the second character after speech onset. Formulation of verb-initial sentences deviated from this pattern, showing that early production of the verb in a sentence called for earlier encoding of relational information.⁸ Compared to subject-initial sentences, speakers showed a smaller preference for the subject character in verb-initial sentences both in the early time window, associated with gist apprehension (0–600 ms), and in later time windows, associated with linguistic encoding (600–1800 ms, 1800–3000 ms).

Within the two types of verb-initial sentences, there were also more fixations to the first-mentioned than the second-mentioned character in the 600–1800 ms time window, indicating that linguistic encoding of the first character had also begun before speech onset. Importantly, the likelihood of fixating the first-mentioned character in verb-initial sentences was still smaller than in subject-initial sentences, confirming that early production of the verb enforced a structure-specific formulation strategy.

Experiment 2

For a direct comparison of sentence formulation in Tzeltal to formulation in a subject-initial language, Experiment 2 examined performance of native Dutch speakers in the same task. We first examine the effects of character animacy and first character fixations on structure choice, and then compare formulation of subject-initial Dutch and Tzeltal sentences.

Method

Participants

A total of 21 native speakers of Dutch from the Nijmegen area participated for payment.

Materials, design, and procedure

The experiment and procedure were identical to that of Experiment 1.

Sentence scoring and analyses

Sentences produced on target trials were scored as actives, full passives, truncated passives and other constructions. Analyses were carried out on the smaller data-set consisting of actives and full passives.

For all analyses, trials were excluded if the first fixation in that trial fell on either the agent or the patient (this resulted in the removal of 91 responses) or if the first fixation directed to a character occurred 400 ms after picture onset (38 additional trials). This left 905 sentences, of which 656 were transitive descriptions. Responses were then also excluded if onsets were longer than three standard deviations from the grand mean (12 sentences). The final data-set consisted of 644 sentences (561 actives, 64 full passives, 19 truncated passives).

Results

Sentence structure

Speakers produced overwhelmingly more active than passive descriptions (.90 active sentences). Sentence structure again depended on character animacy: Events with human agents elicited more active sentences than events with non-human agents (.95 vs. .83), and conversely, events with human patients elicited fewer active sentences than events with non-human patients (.83 vs. .96). The interaction between Agent and Patient Animacy was reliable ($\beta = -3.90$, $z = -2.83$), showing that properties of the agent exerted a stronger influence on sentence form than properties of the patient (Figure 6): Production of active sentences did not vary with patient animacy when events included a human agent, but was more sensitive to patient animacy for events with non-human agents. The presence of this interaction in the Dutch data-set but not the Tzeltal data-set may be due to the fact that Dutch speakers demonstrated a larger preference for active syntax overall.

First character fixations

Speakers directed more first fixations to agents than patients (.71 vs. .29). Human agents attracted only numerically more fixations (.72) than non-human agents (.69), suggesting that the two types of agents did not differ in overall salience.

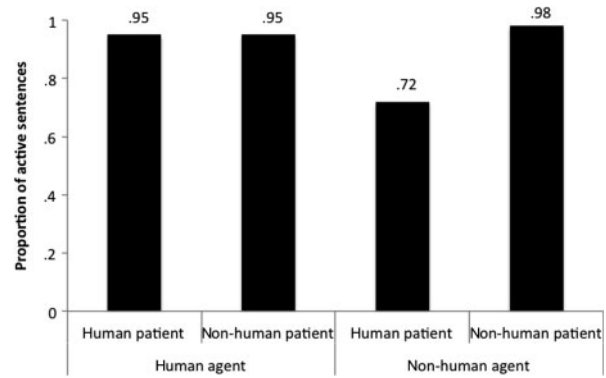


Figure 6. Proportions of active sentences produced in Dutch (Experiment 2) as a function of agent and patient animacy.

More importantly, the influence of first fixations on sentence form was relatively weak. Speakers produced more active sentences when they first fixated the agent (.93) than when they first fixated the patient (.81; Figure 7). This resulted in a main effect of First fixations ($\beta = 1.12$, $z = 2.26$) in a full model including all two-way interaction between First Fixations, Agent Animacy and Patient Animacy, as well as by-participant random slopes for Agent Animacy. However, as in Experiment 1, including First Fixations in the model did not reliably improve model fit, confirming that properties of the two characters were stronger predictors of sentence form than early attention shifts.

Time course of formulation

Figure 8 shows the expected pattern for the time course of formulation of active and passive sentences in an SVO language. When producing active sentences, speakers

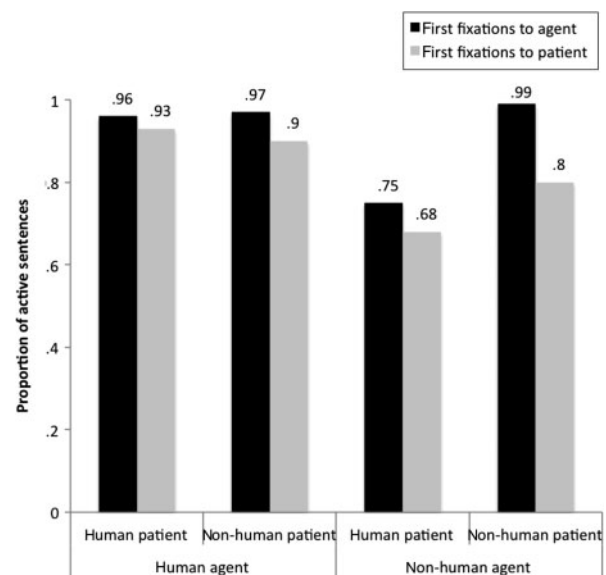


Figure 7. Effect of first fixations on production of active and passive sentences in Dutch (Experiment 2).

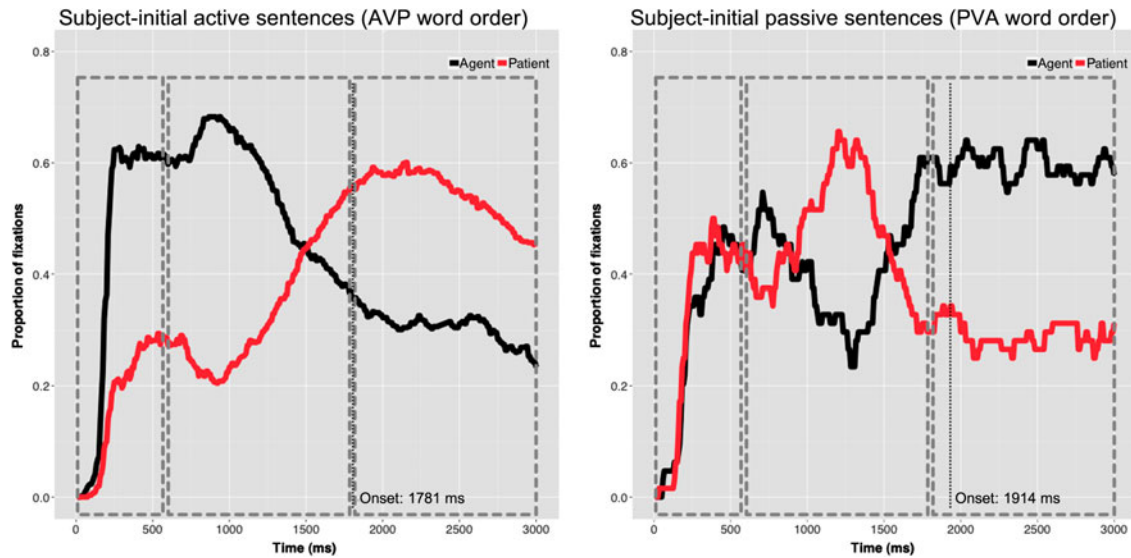


Figure 8. Proportion of agent-directed and patient-directed fixations in active and passive subject-initial sentences in Dutch (Experiment 2).

looked quickly at the agent, continued fixating this character until speech onset, and finally shifted their gaze to the patient. When producing passive sentences, they looked preferentially at the patient before speech onset and shifted their gaze to the agent after speech onset (high variability in the early 0–600 ms time window is due to sparse data and to the fact that sentences with a dispreferred structure are generally harder to generate).

Time course analyses compared formulation of subject-initial active sentences in Dutch and Tzeltal across the two experiments before speech onset (0–600 ms, 600–1800 ms). In the 0–600 ms time window, Dutch speakers were somewhat more likely to fixate the agent within 200 ms of picture onset than Tzeltal speakers, but both groups fixated the agent at comparable rates between 200 and 600 ms (resulting in an interaction between Time Bin and Language: all $t_s < -13$). The difference between groups prior to 200 ms is due to the fact that, on average, first fixations to the agent occurred earlier in the Dutch data-set than the Tzeltal data-set ($M = 208$ vs. 251 ms, respectively). Between 600 ms and 1800 ms, Dutch speakers also looked away from the agent earlier than Tzeltal speakers (resulting in an interaction between Time Bin and Language: all $t_s > 16$). The difference was again likely due to the fact that Dutch speakers initiated their sentences faster than Tzeltal speakers.

To control for overall differences in production speed, complementary analyses were carried out after normalising the durations of all trials (such that a time of 0 corresponds to picture onset and a time of 1 corresponds to speech onset). These analyses showed no difference between agent-directed fixations in the window corresponding to the first 600 ms of each trial in the two groups of speakers (including an interaction between Time Bin

and Language group did not improve model fit compared to an additive model: $\chi^2(1) = .49$, $p = .49$ by-participants; the by-items analysis showed a marginally reliable interaction). The analysis of the time frame corresponding to the 600–1800 ms time window again showed that Dutch speakers began shifting their gaze away from the agent somewhat faster than Tzeltal speakers (the interaction of Time Bin and Language group was reliable by participants, $t < -17$, but not by items).

Discussion

Experiment 2 highlights several important similarities and differences in sentence formulation between Dutch and Tzeltal. First, in both languages, structure choice was sensitive to character animacy: highly accessible (human) characters were more likely to become sentence subjects than less accessible (non-human) characters. This demonstrates that across typologically very different languages, the same conceptual features exert a similar kind of influence on voice choice (influencing subject selection in both cases). At the same time, Dutch speakers showed a greater overall preference for active syntax by comparison with Tzeltal speakers, for whom voice choice was more sensitive to the relative animacy of the agent and the patient. This may imply cross-linguistic differences in the extent to which accessible message elements drive choices between structural options (see also Gennari et al., 2012).

By comparison with conceptual accessibility, early shifts of visual attention exerted a very weak effect on structure choice. Dutch speakers showed only a weak tendency to begin their sentences with the character that first attracted their attention; in Tzeltal, there was no discernible effect of first fixations on sentence voice or

word order at all. Again, these (weak) differences may reflect cross-linguistic differences in how linguistic encoding processes are influenced by the availability of message-level information.⁹

Importantly, time course analyses revealed remarkable similarities in the formulation of SVO sentences in Dutch and in Tzeltal in this item set. Speakers looked preferentially at the character that would become the sentence initial subject before speech onset, and then preferentially fixated the second character. Thus, across languages, sentences that are structurally similar were formulated in similar ways. This cross-linguistic parallelism is particularly striking, given that the two populations under study differ along a number of non-linguistic dimensions that could, in principle, have influenced gaze behaviour: perhaps most relevantly the Tzeltal participants in our study had little to no prior experience with computers or with participating in experiments. Nevertheless, such differences do not appear to have influenced gaze patterns, allowing us to be fairly confident that the relationship between *looking* and *speaking* is stable across the two languages. Thus, since formulation of Tzeltal *verb-initial* sentences deviated markedly from both the Dutch and Tzeltal SVO pattern, the two experiments together provide converging evidence that, within and across languages, differences in the linear order of words in sentences affect the order of encoding operations throughout formulation.

General discussion

Message and sentence formulation involve closely coordinated conceptual and linguistic operations that transform conceptual representations into linear sequences of words. Here we tested how the preparation of conceptual and linguistic material before articulation may be influenced by the grammatical properties of the target language, using the contrast between verb-initial and subject-initial structures in Tzeltal and Dutch.

Accessibility effects on sentence formulation

As a first measure of how language structure influences information flow at the interface between message conception and linguistic formulation, we examined the effects of conceptual and perceptual accessibility on voice choice (active vs. passive) and word order (verb-initial vs. subject-initial). Speakers in both languages were sensitive to conceptual accessibility (character animacy). Importantly, in Tzeltal, where subjects may be produced sentence-initially as well as sentence-finally, conceptual accessibility influenced subject selection regardless of word order. This demonstrates that conceptually available information is not necessarily seized ‘on the fly’ by lexical retrieval processes, setting in motion an opportunistic, linearly incremental formulation process whereby the

most available nominal concept is the first to be encoded and articulated. Rather, a referent’s animacy may influence the *mapping* between message-level event roles (agent, patient) and grammatical roles (subject, object), implying a wider scope of planning at the message level. This finding is consistent with previous studies showing that accessibility may influence subject selection, rather than (or in addition to) linear order (Christianson & Ferreira, 2005; Tanaka et al., 2011). The Tzeltal results represent perhaps the most dramatic demonstration of this phenomenon to date, given that in verb-initial structures, the subject is positioned *last* in the sentence. For Tzeltal, it is likely that early subject selection in verb-initial sentence production is also necessitated by the fact that verbs carry subject agreement marking, which enforces a syntactic commitment at the outset of formulation. An important question for future cross-linguistic research is whether conceptual accessibility influences subject selection to the same extent in VOS languages in which verbs do not carry agreement information.

In addition, the results show that formulation of subject-initial structures in Tzeltal, which are in principle compatible with a linear (word-driven) formulation process, was also not strictly linear. Analyses of character animacy effects on word order choice showed that speakers did *not* automatically assign the most accessible referent to a sentence-initial position (thereby producing a subject-initial sentence), indicating that the choice to utter a subject-initial sentence was not immediately driven by the availability of a nominal concept that could trigger early retrieval of a single character name. Rather, Tzeltal speakers’ preference for selecting a subject-initial over a verb-initial structure was sensitive to the *match* in animacy of the two arguments. Speakers produced verb-initial structures more often when the two characters had different features (e.g. a human and an animal), but preferred to separate two arguments with matching conceptual features (e.g. two humans, or two animals) by selecting a subject-initial structure. Notably, similar preferences have been described for K’iche’, another verb-initial Mayan language, for the feature of definiteness rather than animacy: England (1991) observes that speakers of K’iche’ strongly prefer SVO structures when both the arguments are either indefinite or definite.

Speakers’ departure from verb-initial structures may reflect a general preference to avoid interference that might otherwise arise from the adjacency of two similar elements (e.g., Bock, 1987; Dell, Oppenheim, & Kitredge, 2008; Gennari, et al., 2012; Jaeger, Furth, & Hilliard, 2012). In support of this, Gennari and colleagues found that speakers of English, Spanish and Serbian are less likely to produce active object relative clause constructions with two adjacent noun phrases (*the man (who/that) the woman is punching*) when the two entities are human, and hence conceptually similar. Alternatively,

there may be a communicative explanation for this result. In several experimentally-elicited pantomime studies, it has been found that participants prefer to pantomime SVO structures (e.g. *girl kicks boy*) over SOV structures (*girl boy kicks*) when describing ‘semantically reversible’ transitive events (i.e., events involving two human participants, where either could be interpreted as the agent; Meir, Lifschitz, Ilkbasaran, & Padden, 2010; Gibson et al., 2013). Gibson et al. (2013) explain these results in the context of rational communicative behaviour over a noisy channel: rational producers should avoid SOV structures for reversible events because if either argument were lost due to noise, this would hinder communication (e.g., if either noun in the sentence *girl boy kicks* were lost, it becomes unclear whether the remaining argument was the agent or the patient). SVO word order minimises the communicative consequences of such an ambiguity or uncertainty because the partial structure is still interpretable (e.g. *kicks boy*). While Gibson et al.’s proposal was developed to account for the avoidance of verb-final structures, the same argument can be applied to verb-initial structures, because here too, the loss of one of two post-verbal arguments would lead to problems of recoverability for reversible events. Ultimately, whether communicative efficiency or avoidance of semantic interference turns out to be the correct explanation for Tzeltal speakers’ choice to produce subject-initial or subject-medial sentences, the fact that this choice is influenced by the *combined* conceptual features of the two event characters allows us to conclude that the production of subject-initial sentences in Tzeltal is not typically the outcome of a linear incremental (word-driven) formulation process.

In contrast to the influence of animacy on structure choice, the effect of early gaze shifts (i.e., gaze shifts resulting from differences in early fixation order across characters in an event) was weak in Dutch and non-existent in Tzeltal. This result is consistent with recent work focusing on the relationship between early fixations and sentence form in SVO languages, which suggests that low-level perceptual properties may generally be subordinate to conceptual factors in their capacity to affect formulation (Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Van de Velde et al., 2014). The fact that our data showed a weak effect of perceptual accessibility on structure choice in Dutch (first-fixated characters were more likely to become sentence subjects), but no effect in Tzeltal may also indicate that the extent to which perceptual accessibility affects linguistic formulation differs across languages as a function of language-specific grammatical properties. Support for this possibility comes from studies of case marking languages showing that, by comparison with English, perceptual salience exerts little or no effect on structure choice (Hwang & Kaiser, 2009, for Korean; Myachykov, Garrod, & Scheepers, 2010 for

Russian and Finnish; Myachykov & Tomlin, 2008 for Russian). Myachykov et al. (2010) speculate that obligatory case-marking enforces a structural commitment at the outset of sentence formulation. Similarly in Tzeltal, the overall tendency to begin sentences with morphologically complex verbs, which necessitates early relational encoding and an upfront syntactic commitment, may, in general, attenuate a reliance on perceptual accessibility at the outset of formulation.

Effects of sentence structure on sentence formulation

Taken together, the effects of accessibility on voice and word order choice in Tzeltal argue against a radically incremental formulation process both for the production of verb-initial sentences and subject-initial sentences. While this implies a certain similarity with respect to the nature of early message preparation for both word orders (e.g., some degree of processing of agents and patients in the target events), time course analyses showed that, from a very early stage of formulation, the word order that was under production exerted a strong effect on the way that speakers assembled their sentences online. The pattern of fixations observed in earlier studies with English and Dutch speakers (Gleitman et al., 2007; Griffin & Bock, 2000; Konopka & Meyer, 2014; Kuchinsky & Bock, 2010) was fully replicated with both Tzeltal and Dutch speakers for subject-initial sentences: Event characters were fixated in a predictable, sequential order, anticipating order of mention. This cross-linguistic similarity in the formulation of subject-initial sentences demonstrates that when the linear order of constituents used by speakers is the same, so is the time course of formulation. The striking contrast with formulation of *verb-initial* sentences suggests that early production of the verb changed the order of encoding operations: relational information received priority over encoding of either character, as shown by a convergence of fixations to agents and patients over a nearly two-second time window in these sentences.

We note that this result rules out the possibility, sometimes advocated in the literature, that a verb lemma is necessarily retrieved at the outset of sentence formulation (Bock, 1987; Bock & Levelt, 1994; Ferreira, 2000). In our two experiments, the order of visual uptake of information from an event differed between SVO and VOS sentence types, both within the earliest time window (associated with conceptual encoding), and in later time windows (associated with linguistic encoding), implying that the timing of both conceptual and linguistic encoding required for verb retrieval differed as a consequence of the word order of the to-be-uttered sentence. Similarly, there is some evidence suggesting that verbs are not planned early in verb-final structures either (Hwang & Kaiser, 2014, for Korean and Schriefers, Teruel, & Meinhausen,

1998, for German; but see Kurumada & Jaeger, 2015, for evidence of some advanced planning of the verb in Japanese verb-final structures).

How is it that the order of encoding operations, as reflected in eye movement patterns, so closely anticipates the word order of the to-be-uttered sentence? One logical possibility, compatible with the linear incrementality view, is that the eye is drawn first to some element in the visual scene, causing speakers to start their sentence with the information that first attracted their attention (Gleitman et al., 2007). The other, causally inverse possibility is that the eye is directed to attend first to certain aspects of the scene as a consequence of having *already* generated a structural plan for the sentence (Bock et al., 2004; Griffin & Bock, 2000). As previously discussed, in the Tzeltal experiment, speakers' structural choices were not affected by where they first directed their gaze. The different patterns of fixations we find for verb-initial and subject-initial structures in the early stages of formulation are therefore likely to *reflect* rather than precede the formulation of a structural sentence frame.

Moreover, the fact that already within 600 ms of picture onset we see structure-specific differences in the uptake of visual information implies that a rudimentary sentence frame can be generated within a very rapid time frame, already within the first few hundred milliseconds of picture viewing. This possibility is supported by recent studies showing that very brief presentations (40–300 ms) of event pictures are sufficient for speakers to identify event categories, as well as the role and identity of characters in the event (Dobel, Gumnior, Bölte, & Zwitserlood, 2007; Hafri et al., 2013). Connecting these results to our animacy effects on structure choice, the picture that emerges is thus one in which rapid gist extraction allows for the quick identification of the two characters' event roles and their animacy features, on the basis of which a rudimentary structural frame is generated. This structural frame in turn serves to guide subsequent conceptual and linguistic encoding operations, leading the eye to sample information from the visual scene in the order that the structure calls for it.

Further cross-linguistic research will need to clarify whether the extent of early relational encoding for verb-initial structures differs as a function of the properties of the verbs themselves: In Tzeltal, the extensive prioritising of early relational encoding may be driven not only by the verb's placement, but also by its complex morphology, which specifies information about both participants in the event. For recent evidence supporting the possibility that verbal morphology can affect the early stages of formulation in a verb-initial language, see Sauppe, Norcliffe, Konopka, Van Valin and Levinson's (2013) study of sentence production in Tagalog.

Finally, we also note that differences in formulation of subject-initial and verb-initial sentences within and across

languages are not all or none: they do not imply categorical differences in the underlying planning strategies but rather point to *shifts* in the planning strategies that speakers employ to formulate the two types of sentences. Indeed, recent studies of sentence formulation in SVO languages (English and Dutch) show that formulation of subject-initial sentences can involve a fair degree of relational planning as well (contrary to the strong version of linear incrementality advocated by Gleitman et al., 2007). For example, speakers are more likely to begin formulation by encoding the relational structure of the target event when this information is easy to express linguistically and when a suitable syntactic structure is easy to generate (see discussions in Konopka & Meyer, 2014; Kuchinsky & Bock, 2010; Van de Velde et al., 2014). The results of Experiment 1 in this paper show that the shift towards priority encoding of relational information at the outset of formulation is considerably *larger* when the structure of the sentence explicitly requires early encoding of relational information. An important avenue of future research will be to determine the extent to which such early structure-mediated effects on formulation are attenuated or heightened under different circumstances, for example, given varying degrees of event codability, or depending on the nature of the speech context (isolated sentence production vs. connected discourse).

In short, the existence of a consistent relationship between the order in which information is viewed and the order in which it is expressed demonstrates that sentence structure and online processing are tightly coupled from the earliest stages of formulation. This calls into question the idea that message formulation is necessarily encapsulated from linguistic formulation (Levelt, 1989). Rather, the results suggest that there may be no strict separation between processes related to conceptualisation and those related to linguistic formulation in spontaneous speech.

Acknowledgements

We thank Antun Guzman Osil and Manel Guzman Osil in Tenejapa and Juan Méndez Girón at CIESAS, San Cristóbal de las Casas, for research assistance and transcriptions. We also thank Gilles Polian at CIESAS for research support. In Nijmegen, we thank Emiel van den Hoven and Frédérique Schless for help with the Dutch data collection and transcriptions, Gabriela Garrido for additional experimental assistance, Francisco Torreira for Praat scripting and Tilman Harpe for drawing a selection of the picture stimuli. Finally, we are grateful for feedback from audiences at the Max Planck Institute for Psycholinguistics, the Department of Comparative Linguistics at the University of Zurich, the 2013 Annual CUNY Conference on Human Sentence Processing and at the symposium on 'The Large Cognitive Implications of Small Languages' at the 2014 American Academy of Arts and Sciences Annual Meeting.

Disclosure statement

No potential conflict of interest was reported by the authors

Funding

This work was conducted within the framework of the ERC Advanced Grant [#269484] INTERACT, awarded to SCL.

Notes

1. VSO and OVS word orders are possible but very rare (.9% and 3%, respectively, in Robinson's, 2002, corpus).
2. Preverbal subjects are described as having pragmatic functions to do with topicality (Robinson, 2002; Polian 2013). In Tzeltal discourse, the sentence-initial position is often used for topic resumption or topic initiation. However, topics do not have to occur pre-verbally: sentence final subjects can also be topical (Robinson, 2002).
3. Events with inanimate agents were included to increase production of passives. In all analyses, inanimate agents are grouped together with animal agents as 'non-human' (the results for items with inanimate agents and animal agents did not differ).
4. Instruments were included to increase the range of identifiable action types.
5. Since fixations occurring before 400 ms are critical for evaluating patterns in the early scan paths, we excluded trials where speakers' deployment of attention to the picture was delayed beyond this window.
6. Analyses of structure choice included responses where instruments were mentioned. To rule out a possible influence of instrument mention on our results we also repeated all analyses excluding trials with instruments mentioned. This exclusion did not change any of the results we report.
7. Allowing for the use of a different experimental set-up and a different population than in earlier studies, we chose a wider time window for the first analysis (0–600 ms) than normal (0–400 ms). However, carrying out analogous analyses for active sentences on the smaller time window (0–400 ms) largely showed the same results. These analyses were not carried out for passive sentences due to sparse data.
8. There may be alternative explanations for the convergence of fixations to agents and patients in verb-initial sentences, but we do not find them compelling. The first explanation concerns potential information structural differences between VOS and SVO word orders in Tzeltal. Specifically, because sentence-initial subjects function as sentence topics in Tzeltal (see Footnote 2), the pattern of fixations we find for verb-initial sentences could simply reflect a failure to identify an appropriate topic to select as the sentential starting point. We regard this as unlikely for several reasons. First, sentence-final subjects can also be topics (Robinson, 2002), so there is no reason to assume that the production of a verb-initial sentence is necessarily the outcome of a failure to identify a topic. Second, if speakers chose to produce a subject-initial structure because they had identified a topic-worthy entity in the event, then this would predict that animate entities would be selected preferentially to be sentence-initial subjects, given that animacy is known to contribute to a referent's 'topic-worthiness' (Givón, 1976; Mak, Vonk, & Schriefers., 2006). However, our structure choice analyses showed that speakers preferred to select animate entities to be the subject, regardless of the word order produced. Third, if the convergent fixation patterns in verb-initial sentences reflected a failure to identify a topic, this would predict longer speech onset

latencies for verb-initial sentences compared to subject-initial sentences, yet verb-initial sentences are produced more quickly on average (1674 ms vs. 1830 ms; see Figure 5a and 5b). Finally, our task involved the description of a series of unconnected pictures that were not embedded in any larger discourse context. As such, each picture consistently presented an 'all-new' context for speakers, rendering discourse-level influences less of a potential concern. Another possible explanation for the convergence of fixations in verb-initial sentences is that the results average over items that differ in the extent to which speakers need to process both characters to encode a suitable verb (see Hafri, Papafragou & Trueswell, 2013). To test this hypothesis, we compared the time course of formulation for events where the action was primarily 'carried' by the agent and events where the action was primarily 'carried' by the patient (determined via a norming study completed by a different group of Dutch participants). While speakers tended to direct more fixations before speech onset to the character that was more 'informative' for the purposes of encoding the verb, this factor alone did not account for the large difference in fixations observed between subject-initial and verb-initial sentences. These results confirm that, when encoding a verb first, speakers do prefer to fixate both the agent and the patient.

9. An anonymous reviewer observes that the failure to find a robust effect of first fixations on structure choice in Tzeltal could simply reflect a lack of statistical power: In Gleitman et al.'s (2007) study of attentional effects on English word order choices, the observed effect was small (speakers were only 10% more likely to produce passive structures when their attention was directed to the patient in the scene). We believe this is unlikely: The Tzeltal experiment involved substantially more participants than the Dutch experiment, so if anything, we had more power to detect an effect of first fixations in Tzeltal than in Dutch.

References

- Aissen, J. (1997). On the syntax of obviation. *Language*, 73, 705–750.
- Arnold, J. E., Wasow, T., Losongco, T., & Ginstrom, R. (2000). Heaviness vs. Newness: The effects of structural complexity and discourse status on constituent ordering. *Language*, 76(1), 28–55. doi:10.1353/lan.2000.0045
- 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. doi:10.1016/j.jml.2007.12.005
- Barr, D. J. (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language*, 59, 457–474. doi:10.1016/j.jml.2007.09.002
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. doi:10.1016/j.jml.2012.11.001
- Bock, K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89(1), 1–47. doi:10.1037/0033-295X.89.1.1
- Bock, K. (1987). Co-ordinating words and syntax in speech plans. In A. W. Ellis (Ed.), *Progress in the psychology of language* (pp. 337–390). London: Erlbaum.
- Bock, K. (1990). Structure in language: Creating form in talk. *American Psychologist*, 45, 1221–1236. doi:10.1037/0003-066X.45.11.1221

- Bock, 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, NY: Psychology Press.
- Bock, K., Irwin, D. E., Davidson, D. J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, 48, 653–685. doi:10.1016/S0749-596X(03)00007-X
- Bock, K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945–984). London: Academic Press.
- Bock, K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: Bridging the syntactic cleft. *Psychological Review*, 99, 150–171. doi:10.1037/0033-295X.99.1.150
- Bock, K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, 21(1), 47–67. doi:10.1016/0010-0277(85)90023-X
- Branigan, H. P., & Feleki, E. (1999). Conceptual accessibility and serial order in Greek speech production. *Proceedings of the 21st Cognitive Science Society Conference*, Vancouver, WA.
- 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. doi:10.1016/j.cognition.2008.07.011
- Brown-Schmidt, S., & Konopka, A. E. (2014). Processes of incremental message planning during conversation. *Psychonomic Bulletin & Review*. Advance online publication. doi:10.3758/s13423-014-0714-2
- Brown-Schmidt, S., & Tanenhaus, M. K. (2006). Watching the eyes when talking about size: An investigation of message formulation and utterance planning. *Journal of Memory and Language*, 54, 592–609. doi:10.1016/j.jml.2005.12.008
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, 113, 234–272. doi:10.1037/0033-295X.113.2.234
- Christianson, K., & Ferreira, F. (2005). Conceptual accessibility and sentence production in a free word order language (Odawa). *Cognition*, 98(2), 105–135. doi:10.1016/j.cognition.2004.10.006
- Cohn, N., & Paczynski, M. (2013). Prediction, events, and the advantage of Agents: The processing of semantic roles in visual narrative. *Cognitive Psychology*, 67(3), 73–97. doi:10.1016/j.cogpsych.2013.07.002
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283–321. doi:10.1037/0033-295X.93.3.283
- Dell, G. S., Oppenheim, G. M., & Kittredge, A. K. (2008). Saying the right word at the right time: Syntagmatic and paradigmatic interference in sentence production. *Language and Cognitive Processes*, 23, 583–608. doi:10.1080/01690960801920735
- Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, 125(2), 129–143. doi:10.1016/j.actpsy.2006.07.004
- Ferreira, F. (2000). Syntax in language production: An approach using tree-adjoining grammars. In L. Wheelton (Ed.), *Aspects of language production* (pp. 291–330). London: Psychology Press.
- Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, 46(1), 57–84. doi:10.1006/jmla.2001.2797
- Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of Psycholinguistic Research*, 32, 669–692. doi:10.1023/A:1026146332132
- Fisher, C. (2002). The role of abstract syntactic knowledge in language acquisition: A reply to Tomasello (2000). *Cognition*, 82, 259–278. doi:10.1016/S0010-0277(01)00159-7
- Gennari, S. P., Mirković, J., & MacDonald, M. C. (2012). Animacy and competition in relative clause production: A cross-linguistic investigation. *Cognitive Psychology*, 65, 141–176. doi:10.1016/j.cogpsych.2012.03.002
- Gibson, E., Piantadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A noisy-channel account of cross-linguistic word order variation. *Psychological Science*, 24, 1079–1088. doi:10.1177/0956797612463705
- Givón, T. (1976). Topic, pronoun and grammatical agreement. In C. N. Li (Ed.), *Subject and topic* (pp. 149–188). London: Academic Press.
- Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57, 544–569. doi:10.1016/j.jml.2007.01.007
- Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279. doi:10.1111/1467-9280.00255
- Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, 142, 880–905. doi:10.1037/a0030045
- Hwang, H., & Kaiser, E. (2009, March). *The effects of lexical vs. perceptual primes on sentence production in Korean: An on-line investigation of event apprehension and sentence formulation*. Paper presented at the 22nd CUNY Conference on Human Sentence Processing, Davis, CA.
- 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*, 40, 1363–1376. doi:10.1037/a0036797
- Ibbotson, P., Lieven, E. V. M., & Tomasello, M. (2013). The attention-grammar interface: Eye-gaze cues structural choice in children and adults. *Cognitive Linguistics*, 24, 457–481. doi:10.1515/cog-2013-0020
- 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. doi:10.1016/j.jml.2007.11.007
- Jaeger, T. F., Furth, K., & Hilliard, C. (2012). Phonological overlap affects lexical selection during sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 1439–1449. doi:10.1037/a0027862
- Jaeger, T. F., & Norcliffe, E. (2009). The cross-linguistic study of sentence production: State of the art and a call for action. *Language and Linguistics Compass*, 3, 866–887. doi:10.1111/j.1749-818X.2009.00147.x
- Kempen, G., & Harbusch, K. (2004). A Corpus study into word order variation in German subordinate clauses: Animacy affects linearization independently of grammatical function assignment. In T. Pechmann (Ed.), *Language production* (pp. 173–181). Berlin: Mouton.
- Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science*, 11, 201–258. doi:10.1207/s15516709cog1102_5
- 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. doi:10.1016/j.jml.2011.08.003

- Konopka, A. E., & Bock, K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58(1), 68–101. doi:10.1016/j.cogpsych.2008.05.002
- Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, 73, 1–40. doi:10.1016/j.cogpsych.2014.04.001
- Kuchinsky, S. E., & Bock, K. (2010, March). *From seeing to saying: Perceiving, planning, producing*. Paper presented at the 23rd meeting of the CUNY Human Sentence Processing Conference, New York, NY.
- Kurumada, C., & Jaeger, T. F. (2015). *Communicative efficiency in language production: Optional case-marking in Japanese*. Manuscript submitted for publication. Retrieved from https://www.academia.edu/5142210/Kurumada_C_and_Jaeger_T.F._resubmitted_Communicative_efficiency_in_language_production_Optional_casemarking_in_Japanese._submitted_for_publication
- Lee, E. K., Brown-Schmidt, S., & Watson, D. W. (2013). Ways of looking ahead: Incrementality in language production. *Cognition*, 129, 544–562. doi:10.1016/j.cognition.2013.08.007
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- MacWhinney, B. (1977). Starting points. *Language*, 53, 152–168. doi:10.2307/413059
- 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. doi:10.1016/S0022-5371(78)90326-2
- Mak, W. M., Vonk, W., & Schriefers, H., (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466–490. doi:10.1016/j.jml.2006.01.001
- McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, 25, 188–230. doi:10.1006/cogp.1993.1005
- Meyer, A. S. (1996). Lexical access in phrase and sentence production: Results from picture – word interference experiments. *Journal of Memory and Language*, 35, 477–496. doi:10.1006/jmla.1996.0026
- Meir, I., Lifschitz, A., Ilkbasaran, D., & Padden, C. (2010, April). *The interaction of animacy and word order in human languages: A study of strategies in a novel communication task*. Paper presented at the 8th International Conference on the Evolution of Language, Utrecht.
- Myachykov, A., Garrod, S., & Scheepers, C. (2010). Perceptual priming of structural choice during English and Finnish sentence production. In R. K. Mishra & N. Srinivasan (Eds.), *Language & cognition: State of the art* (pp. 54–72). Munich: Lincom Europa.
- Myachykov, A., & Tomlin, R. S. (2008). Perceptual priming and structural choice in Russian sentence production. *Journal of Cognitive Science*, 6, 31–48. Retrieved from https://www.researchgate.net/publication/44842550_Perceptual_priming_and_structural_choice_in_Russian_sentence_production
- Paul, H. (1886/1970). The sentence as the expression of the combination of several ideas. In A. L. Blumenthal (Ed. & Trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 34–37). New York: Wiley.
- Polian, G. (2013). *Gramática del tseltal de Oxchuc* [Grammar of Oxchuc Tzeltal]. Mexico, DF: Centro de Investigaciones y Estudios Superiores en Antropología Social.
- Robinson, S. (2002). Constituent order in Tenejapan Tzeltal. *International Journal of American Linguistics*, 68(1), 51–81. doi:10.1086/466479
- Sauppe, S., Norcliffe, E., Konopka, A. E., Van Valin, R. D. Jr., & Levinson, S. C. (2013). Dependencies first: Eye-tracking evidence from sentence production in Tagalog. In M. Knauff, M. Pauen, N. Sebanz, & E. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the Cognitive Science Society* (pp. 1265–1270). Austin, TX: Cognitive Science Society.
- Schriefers, H., Teruel, E., & Meinhausen, R. M. (1998). Producing simple sentence: Results from picture-word interference experiments. *Journal of Memory and Language*, 39, 609–632. doi:10.1006/jmla.1998.2578
- Smith, M., & Wheeldon, L. (2001). Syntactic priming in spoken sentence production – An online study. *Cognition*, 78, 123–164. doi:10.1016/S0010-0277(00)00110-4
- Swets, B., Jacovina, M. E., & Gerrig, R. J. (2008, November). *Individual differences in the planning scope of language production*. Paper presented at the 49th meeting of the Psychonomic Society, Chicago, IL.
- Tanaka, M. N., Branigan, H. P., McLean, J. F., & Pickering, M. J. (2011). Conceptual influences on word order and voice in sentence production: Evidence from Japanese. *Journal of Memory and Language*, 65, 318–330. doi:10.1016/j.jml.2011.04.009
- Tomlin, R. (1995). Focal attention, voice, and word order: An experimental, cross-linguistic study. In M. Noonan & P. Downing (Eds.), *Word order in discourse* (pp. 521–558). Amsterdam: John Benjamins.
- Tomlin, R. (1997). Mapping conceptual representations into linguistic representations: The role of attention in grammar. In J. Nuyts & E. Pederson (Eds.), *Language and conceptualization* (pp. 162–189). Cambridge: Cambridge University Press.
- Wundt, W. (1900/1970). The psychology of the sentence. In A. L. Blumenthal (Ed. and Trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31). New York: Wiley.
- Van de Velde, M., Meyer, A. S., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing “easy” and “hard” events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, 71, 124–144. doi:10.1016/j.jml.2013.11.001

Appendix

Target events used in both experiments, listed in alphabetical order by agent–patient animacy combination. Instruments carried by agents are added in parentheses.

Human agents acting on human patients

1. Army general kicking boy
2. Barber cutting man’s beard (with scissors)
3. Boxer punching man
4. Doctor vaccinating boy (with needle)
5. Girl pushing boy
6. Girl tripping construction worker
7. Masseur massaging man
8. Man throwing child up in the air
9. Nun braiding girl’s hair
10. Nurse bathing child (with sponge)
11. Policeman stopping woman
12. Priest pulling donkey (with rope)
13. Soldier shooting man (with gun)
14. Woman ticking girl (with feather)

Human agents acting on animal patients

15. Boy trapping frog (with jar)
16. Cook carrying pig
17. Cowboy lassoing bull (with lasso)
18. Farmer whipping donkey (with whip)
19. Hunter dragging deer
20. Man catching fish (with fishing rod)
21. Man hunting pig (with spear)
22. Security guard pulling dog (with rope)
23. Toddler poking lizard (with stick)
24. Vet examining horse
25. Woman chasing chicken
26. Woman petting sheep

Non-human agents acting on human patients

27. Ball hitting teacher
28. Bird carrying child (with hook)
29. Boulder crushing man
30. Bull attacking girl
31. Cart hitting street vendor
32. Cat scratching girl
33. Crocodile biting man

34. Deer lifting man
35. Dog licking boy
36. Duck splashing bride
37. Fire burning child
38. Horse dragging man
39. Kitten tying girl (with string)
40. Light bulb electrocuting electrician
41. Monkey painting boy (with paintbrush)
42. Mosquito biting football player

Animal agents acting on animal patients

43. Bird catching insect (with stick)
44. Cow pulling dog (with rope)
45. Dog catching butterfly (with net)
46. Dog chasing squirrel
47. Donkey pulling cow
48. Eagle grabbing rabbit
49. Jaguar attacking armadillo
50. Monkey grabbing crab
51. Monkey hooking snake (with stick)
52. Pig sniffing cat