Journal of Memory and Language 71 (2014) 124–144

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

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

Maartje van de Velde [a,b], Antje S. Meyer [a,c], Agnieszka E. Konopka [a,c,*]

[a] Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[b] International Max Planck Research School for Language Sciences, P.O. Box 310, 6500 AH Nijmegen, The Netherlands
[c] Donders Institute for Brain, Cognition and Behavior, P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

ARTICLE INFO

Article history:
Received 4 June 2012
revision received 30 October 2013
Available online 10 December 2013

Keywords:
Timecourse of sentence formulation
Incrementality
Structural processing

ABSTRACT

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

Introduction

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

Much of what we know about message and sentence formulation comes from eye-tracking studies showing a tight link between gaze and speech. When asked to describe pictured events (e.g., a cat catching a mouse; Fig. 1), speakers normally direct their gaze to the charac-

* Corresponding author. Address: Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands.
E-mail address: agnieszka.konopka@mpi.nl (A.E. Konopka).

0749-596X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jml.2013.11.001

ters in the event in the order of mention (Gleitman et al., 2007; Griffin, 2004; Griffin & Bock, 2000). Gaze shifts from the first character to the second character are typically initiated when speakers finish retrieving the name of the first character. Despite the apparent systematicity of this process, however, it is still debated how speakers *begin* to direct their gaze to the two characters in such an orderly fashion. This question concerns the way speakers prioritize encoding the different types of information that become available to them during early viewing of an event.

On one proposal, speakers may immediately direct their gaze to the character that attracts their attention at picture onset, irrespective of its status in the event, and begin encoding it linguistically. For example, when perceptual cues are presented in the location of the agent or the patient in an upcoming picture, speakers tend to shift their gaze to this location within 200 ms of picture onset and to begin their sentences with this character rather than the uncued character (Gleitman et al., 2007; Kuchinsky & Bock, 2010; Myachykov, Tomlin, & Posner, 2005; Tomlin, 1997). Thus studies employing visual cuing manipulations have suggested that speakers can begin message and sentence formulation by encoding as little as one character (e.g., "*cat*"). Planning of the *second* character ("*mouse*"), as well as its relationship to the first character ("*catching*"), must then happen when speakers shift their gaze to the second character. This type of planning is referred to as *linear incrementality* (Bock, Irwin, & Davidson, 2004): in each increment, speakers plan only enough information to describe the referent they are fixating and the sentence is built up from the addition of these increments, one by one, to the developing message.

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

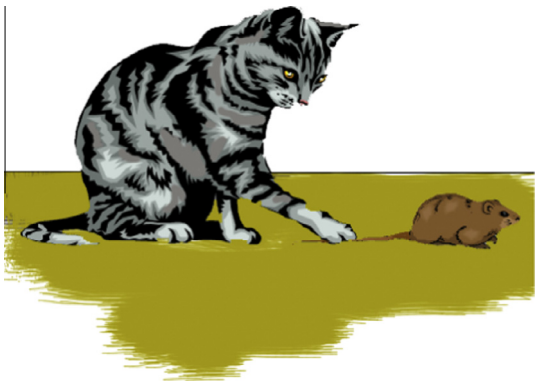

Fig. 1. Example of a target item. The modal active description of this event is "The cat is catching the mouse."

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

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