How does information status affect sentence planning: An eye-tracking study

Lesya Ganushchak[1,2,3] & Yiya Chen[1,3]
[1]Leiden University Centre for Linguistics (LUCL), the Netherlands
[2]Education and Child Studies, Faculty of Social and Behavioral Sciences, Leiden University, the Netherlands
[3]Leiden Institute for Brain and Cognition (LIBC), the Netherlands

a.y.ganushchak@hum.leidenuniv.nl,
yiya.chen@hum.leidenuniv.nl

ABSTRACT

This study investigates online sentence formulation as a function of the information structure of an utterance. In an eye-tracking experiment, participants described pictures of transitive events. Information status of the message was manipulated in the discourse preceding each picture. In the Literal condition, the subject character was literally mentioned. In the Associative condition, the subject character was primed without literal mention. In the No Mention condition, the subject character was neither literally mentioned nor primed. The response was expected to be the same in all conditions. Results showed that participants looked first at the subject character and then the object character regardless of the condition. The time course of the gaze patterns, however, did differ across the three conditions, which reflects the availability of the to-be-encoded information as a function of the preceding discourse context.

Keywords: givenness, sentence formulation, eye-tracking, information status

1. INTRODUCTION

Producing a sentence starts with the preparation of a preverbal message and continues with encoding it linguistically [e.g., 8]. While much work has been done on sentence production in isolation, not much is known about planning of utterances as a function of the discourse context. We know that the phonetic shape of an utterance changes according to the information structure of the utterance ([4], [5], and references therein). Little, however, is known about the specific time course of online sentence formulation as a function of the information structure of an utterance which gives rise to the corresponding prosodic encoding. The goal of this project is to shed light on this issue via investigating the effect of one specific information structure notion: *givenness*. In particular, we will zoom in to examine how different levels of givenness affect the time course of sentence planning.

Speech production is not just the retrieval of lexical forms and syntactic structures. The planning of an utterance can be affected by the information status of the messages to be articulated (e.g., whether the information is new or known/given). Ganushchak, Konopka, and Chen [7] showed that information status affects planning of utterances. Participants were asked to describe pictures of two-character transitive events in Dutch and Chinese, while their eye-movements were recorded. Discourse focus (i.e. the *newness* of the information to be encoded) was manipulated by presenting questions before each picture. Speakers rapidly directed their gaze preferentially only to the character they needed to encode (i.e. the new information). The results suggest that information status of the message affects the time course of linguistic formulation in simple sentences

Not only focus, givenness also affects sentence formulation. A large number of studies have shown that speakers have a strong preference to begin sentences with accessible characters [e.g., 10]. Easy-to-name characters tend to serve as the subject of an utterance more often than harder-to-name characters. This is in accordance with the so-called *minimal load principle* [13] which states that completing easy processes before hard processes lightens the load on the production system and enables speakers to quickly encode individual increments [e.g., 6]. Recently, Konopka & Meyer [12] showed that utterance planning was affected when pervious discourse was semantically related and thus the target subject or object character of a simple Subject-Verb-Object utterance was semi-active.

It is important to note that accessible information cannot be treated as a uniform category [2]. There could be *a continuum of givenness*: a referring expression can be regarded as 1) completely given (i.e. already active in the speaker's mind at the time of utterance); 2) completely new (inactive), 3) accessible (semi-active). Baumann and Grice [2] have also shown that listeners are sensitive to the various prosodic cues that signal different levels of givenness.

To compare the time course of sentence formulation with various stages of givenness, eye-tracked participants were asked to describe pictures shown on a computer screen, as illustrated in Figure

1. Givenness was manipulated by means of a short story that preceded each picture. In the Literal condition, participants heard a story where subject character was literally mentioned (e.g., *frog*). In the Associative condition, the subject character was not mentioned but the story primed participants to think of the target subject character (e.g., story mentioned words such as: *quacking, pond, green, jump*). In the No Mention condition, stories did not literally mention nor prime the subject character depicted on the picture. The target response was expected to have the same structure and content in all conditions (*The frog catches the fly*). Differences in planning of the target response were evaluated by comparing speakers' eye-movements to the subject and object characters in the picture prior to speech onset.

Figure 1: Example of a target event.

When preparing to describe such pictures, speakers normally look at characters in the display in the order of mention [3] [9]. Formulation typically begins with a short apprehension phase (0-400 ms) during which speakers encode the gist of the event and during which fixations to characters usually do not differ. Event apprehension is then followed by a longer phase of linguistic encoding. Typically, easy-to-name characters are fixated for less time than harder-to-name characters [8][11][14].

If givenness does not influence the time course of sentence formulation, there should then be no difference in gaze patterns among conditions. However, if givenness does affect planning of an utterance during linguistic encoding, then viewing patterns in Literal/Associative conditions should differ from the No Mention condition after about 400 ms from the picture onset. Specifically, speakers should direct fewer fixation to the character that was mention/primed in the story. If levels of givenness also influence the planning, then in the Literal condition, there should be even less fixations to the subject character than in the Associative condition, since subject character is 'given' in its articulatory form in the Literal condition while it is only 'accessible' without being articulated in the Associative condition.

2. METHODS

2.1. Participants

Thirty native Dutch speakers (23 women) participated in the experiment (mean age: 21.2 years; SD = 1.6 years). All participants were students of Dutch universities. Participants gave written informed consent prior to participating in the study.

2.2. Materials

Seventy-eight colored pictures were used in the experiment. All pictures displayed simple actions (as exemplified in Figure 1). There were 25 target pictures, 50 fillers, and 3 practice pictures. Levels of givenness were manipulated by means of short stories (two sentences long) preceding each picture. For the Literal condition, subject character, depicted in the picture, was literally mentioned in the preceding story. Note, that the target subject character was always placed in the same grammatical role as in the intended target sentence and it was always placed in the second sentence of the story. For the Associative condition, the story was set up in such a way that it primed activation of the intended target word. All stories were pre-tested, to make sure that they did lead to the activation of the right word. In the No-Mention condition, the story did not include literal or associative mention of words that describe characters in the picture.

Expected target sentence: *De kikker vangt de vlieg* (The frog catches the fly)

(1) Literal story:
David gaat met zijn vader vissen. Een kikker springt opeens in de sloot. (David is going fishing with his father. A frog suddenly jumped into the ditch.)

(2) Associative story:
Koen hoort gekwaak bij de vijver. Als hij gaat kijken, ziet hij iets groens wegspringen.(Koen heard quacking near the pond. When he went to look, he saw something green jumping away.)

(3) No Mention story:
David gaat met zijn vader vissen. Ze gebruiken restjes van het avondeten als aas.(David is going fishing with his father. They use leftovers from dinner as bait.)

All stories were pre-recorded by a native Dutch female speaker and were presented auditorily prior to picture onset. After 40% of the trials, a yes-or-no

comprehension question was presented visually on the computer screen. The questions were presented only after filler trials and were meant to make sure that participants listened attentively to the presented stories.

Three lists of stimuli were created to counterbalance story types across target pictures. Each target picture occurred in each condition on different lists, so that each participant saw each picture only once.

2.3. Design and procedure

Participants were seated in a sound-proof room. They first heard a story and then saw the picture which was contextually related to the preceding story. The task started after three practice trials. All participants were instructed to describe the picture as briefly as possible but mentioning all the characters in the picture (e.g., *the frog catches the fly*). After some filler items, participants were asked a comprehension question between hearing a story and seeing a picture.

2.4. Data analysis

The time course of utterance formulation in the three conditions was compared with by-participant (β_1) and by-item (β_2) quasi-logistic regression analyses performed on agent-directed fixations [1]. We selected three time windows (0-400 ms, 400-800 ms, and 800-1400 ms) for analysis. Fixations were aggregated into a series of time bins of 200 ms each for the analysis for each participant and each item in each condition. The dependent variable in each time bin was an empirical logit indexing the likelihood of speakers fixating the agent out of the total number of fixations observed in that time bin. Time and Condition were entered as fixed effects into all models. All models included random by-participant and by-item random intercepts and slopes for the Time and Condition variables.

Fixations in the three experimental conditions were compared with two contrasts. The first contrast compared the No Mention condition against the Literal condition. The second contrast compared the No Mention condition against the Associative condition. Both contrasts thus assess how planning a sentence in response to the preceding discourse changes the overall distribution of attention to the two characters relative to the information status-neutral condition. Finally, separate analyses were run with new contrasts to compare agent-directed fixations in the Literal and Associative conditions against one another.

3. RESULTS

Figure 2 plots the proportions of fixations to the subject and object characters in target event pictures across condition.

3.1. First analysis (0 – 400 ms)

In all conditions, speakers rapidly directed their gaze to the agent in the picture within 400 ms after picture onset (main effect of Time: $\beta_1 = 6.09$, $\beta_2 = 6.07$, both $ts > 19$). All main effects and interactions in this time window were not significant (all $ts < 1$).

3.2. Second analysis (400 – 800 ms)

Within this window, there were no differences between the No Mention and Literal condition (all main effects and interactions: $ts < 1.5$).

Between the No Mention condition and the Associative condition, results showed that speakers were already less likely to fixate subject characters in the Associative condition than in the No Mention condition in the first 200 ms of the time window (400 ms – 600 ms) with a significant main effect of Condition ($\beta_1 = 0.97$, $\beta_2 = 1.08$, both $ts > 1.9$). Meanwhile, their fixation increased over the object character. The interaction between Time and Condition was also significant ($\beta_1 = -1.77$, $\beta_2 = -1.97$, both $ts < -2$). Within the later window, speakers also quickly re-directed their gaze to the subject character in the Associative condition while in the No Mention condition, fixations to the subject character remained stable.

Comparing the Literal and Associative conditions against one another in a separate analysis showed a significant interaction between Time and Condition ($\beta_1 = -1.37$, $\beta_2 = -1.68$, both $ts < -1.9$). As time progressed towards the end of the window, fixations to agents within this analysis window increased in the Associative condition but not in the Literal condition.

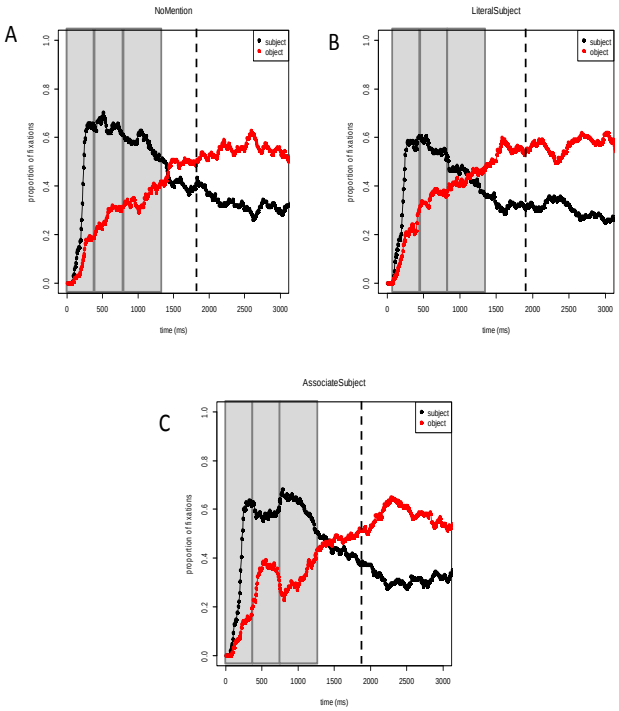
3.3. Third analysis (800 – 1400 ms)

Speakers began shifting their gaze away from the agent between 800 ms and 1400 ms (main effect of Time: $\beta_1 = -0.86$, $\beta_2 = -1.09$, both $ts < -3$). They were more likely to fixate subject characters in the No Mention condition than in the Literal condition ($\beta_1 = 0.57$, $\beta_2 = 0.55$, both $ts > 3.5$). However, the reverse was true for the No Mention condition vs. the Associative condition. Speakers were more likely to fixate subject characters in the Associative condition than No Mention condition ($\beta_1 = -0.38$, $\beta_2 = -0.31$, both $ts < -2$). No

interactions with Time were significant ($ts < 1$), as the decline in agent-directed fixations was comparable in all three conditions.

Comparison of the Literal and Associative conditions against one another in a separate analysis showed a significant main effect of Condition (β_1 = -0.48, β_2 = -0.43, both $ts < -3.5$): speakers were more likely to fixate subject character in the Associative than in the Literal condition. No interactions with time were significant ($ts < 1$).

Figure 2: Proportions of fixations to the subject and object characters in target event pictures: A. the No Mention condition; B. the Literal condition; C. the Associative condition. Dashed lines represent speech onset. Areas selected by rectangles depict the three time window (0 – 400, 400 – 800, and 800 – 1400 ms) used in the analyses.



4. DISCUSSION

The results of this experiment showed an intricate gaze pattern over the time course of utterance planning as a function of the discourse context. During the initial stage of planning (i.e. within 0 to 400 ms after picture onset), discourse context did not exert significant influence on the allocation of attention to the two event characters. This stage corresponds to the well-recognized message encoding stage. This is in line with previous findings [e.g., 7]. After 400 ms, our results replicated earlier findings showing that participants look at characters in the order of mention: first the subject character (*frog*) and then the object character (*fly*; [9]).

The Literal and Associative conditions, however, showed that fixations to the subject and object characters were influenced by the preceding discourse context. In the Literal condition, participants looked longer at the subject character but already shifted their gaze to the object character 1100 ms after the picture onset (see Figure 2B). The Associative condition showed a somewhat different pattern: speakers briefly looked at the subject character and then shifted their gaze to the object character within 500 ms after the picture onset. However, after 800 ms, speakers shifted their gaze back to the subject character.

How to account for our results? It is possible that in the No Mention and Literal conditions, the planning of the sentence occurs in a highly incremental and linear manner. That is, speakers fixated characters in order of mention. In the Literal condition, there was a significant reduction of fixations to the agent 800 ms after picture onset, compared to that in the No Mention and Associative conditions. This reduction can be explained by the fact that in the Literal condition, the subject character was explicitly mentioned in the preceding context, thereby reducing the costs of retrieving and phonetically encoding its name. This in turn led to reduced likelihood of speakers fixating this character [e.g., 7; 11].

In the Associative condition, speakers showed a preference to also fixate on the object character (which is *new* in discourse). This presumably is due to the concept that was strongly activated in the preceding discourse, which the speaker needs to verify during the formulation of the linguistic message. Therefore, there was an increased likelihood of attention allocation on the unmentioned object character. This verification process consequently delayed the phonological/phonetic encoding for the Associative condition, compared to the other two conditions.

To conclude, our results show that sentence planning in discourse context is affected not only by the information status of the message that needs to be encoded but also by the different levels of linguistic information that are accessible to the speaker.

5. ACKNOWLEDGEMENT

We thank Phoebe Ye for help with data collection. This research was supported by the European Research Council (ERC-206198 to YC).

6. REFERENCES

[1] Barr, D. J. 2008. Analyzing 'visual world' eyetracking data using multilevel logistic regression. *Journal of Memory and Language* 59, 457-474.

[2] Baumann, S., Grice, M. 2006. The intonation of accessibility. *Journal of Pragmatics* 38, 1636-1657.

[3] Bock, K., Irwin, D. E., Davidson, D., Levelt, W. J. M. 2003. Minding the clock. *Journal of Memory and Language* 48, 653-685.

[4] Chen, Y. 2012. Message-related variation. In Cohn, A., Fourgeron, C., and Huffman, M. (eds.), *Oxford Handbook of Laboratory Phonology*. Oxford: Oxford University Press.

[5] Cole, J. 2015. Prosody in Context: A review. *Language, Cognition and Neuroscience* 30, 1-31.

[6]Ferreira, F., Henderson, J.M. 1998. Syntactic reanalysis, thematic processing, and sentence comprehension. In J. D. Fodor & F. Ferreira (Eds.), Reanalysis in sentence processing (pp. 73–100). Dordrecht: Kluwer.

[7] Ganushchak, L.Y., A. Konopka, Chen, Y. 2014. What the eyes say about planning of focused referents during sentence formulation: a cross-linguistic investigation. *Frontiers in Psychology* 5, e1124.

[8] Griffin, Z. M. 2004. The eyes are right when the mouth is wrong. *Psychological Science* 15, 814-821.

[9] Griffin, Z. M., Bock, J. K. 2000. What the eyes say about speaking. *Psychological Science* 11, 274–279.

[10] Kempen, G., Hoenkamp, E. 1987. An incremental procedural grammar for sentence formulation. *Cognitive Science* 11, 201–258.

[11] Konopka, A. E. 2014. Speaking in context: Discourse influences formulation of simple sentences. Poster presented at *the 27th CUNY Human Sentence Processing Conference*, Columbus, OH, USA.

[12] Konopka, A. E., Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology* 73, 1–40.

[13] Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge. MA: MIT Press.

[14] Meyer, A. S., Lethaus, F. 2004. The use of eye tracking in studies of sentence generation. In J. M. Henderson & F. Ferreira (Eds.), The interface of language, vision, and action: Eye movements and the visual world (pp. 191–212). New York: Psychology Press.

[15] Rooth, Mats. 1996. Focus. In S. Lappin (ed.), The Handbook of Contemporary Semantic Theory (271-298). Oxford: Blackwell.