Investigating Word Order and Givenness Effects as Part of Expected-first Strategies in Information Structure

Torsten Kai Jachmann, Francesca Delogu, Heiner Drenhaus, Matthew W. Crocker
Language Science and Technology, Saarland University, Germany

In languages with free word order, non-canonical structures have been shown to induce increased processing costs relative to their canonical counterparts (e.g. [1], [2], [3]). It has also been shown that given information leads to reduced processing costs when appearing earlier in a structure, as predicted by the 'given-before-new' principle of information structure (e.g., [4], [5], [6]). When these two factors are combined, the cost for object-first structures has been suggested to be mitigated when that object is given as compared to new (e.g. [2], [7]). However, in some cases (e.g., [2]), the new information appears to be implied or related to the context rather than entirely new (e.g., [7]), possibly leading to differences in the modulations of the effects.

In the present set of three self-paced reading studies in German, we aimed to systematically investigate the interplay between Word Order preferences (SOV over OSV) and different grades of Givenness (Given – explicitly mentioned vs. Implied – inferable from the context vs. New – unmentioned and unrelated to the context), as illustrated in Example 1.

To assess the effect of Givenness on both NP1 and NP2, a total of 12 conditions were created and distributed across three experiments, in which the factor Givenness was split in a pairwise manner (Exp1-Given/New, Exp2-Given/Implied, Exp3-Implied/New). All three experiments manipulated Word Order (SOV/OSV) in the same way. As NP2 did not provide any insights beyond the findings on NP1, it is not discussed further in the scope of this abstract.

Analyses were conducted in Julia by fitting Linear Mixed-Effect Models using the MixedModels package. Models were further refitted in R to conduct Power analyses using the simr package, showing that all reported effects exceeded the 80% threshold. The results (see Fig.1 and Table 1) showed similar effects of Word Order across all three experiments, such that Object-first structures led to longer reading times (RTs) on the first NP and its spill-over region compared to Subject-first structures. Effects of Givenness could only be observed in Exp 1 and Exp 2, where both New and Implied entities resulted in increased RTs compared to Given entities. In Exp 3, no effect of Givenness was found. Further, no interactions were observed in any of the three experiments.

The additive effects of Word Order and Givenness in Exp 1 and Exp 2 did not replicate previous findings suggesting a mitigation of object-first difficulties when the object is given ([2], [7]), but are more in line with a previously proposed 'expected-first' strategy ([8]), according to which comprehenders incur increased processing costs from any kind of violation of expectations on the first NP (e.g.: "given-first", "subjects-first") independently.

Previous ERP studies (e.g., [9]) showed that Implied information is more accessible as compared to entirely New information – resulting in lower lexical retrieval costs – but still requires the integration of a new discourse referent. Consistent with this, the lack of an effect of Givenness in Exp 3, combined with the overall findings, suggests that implied entities are processed similarly to new ones regarding their integration costs (as RTs have been shown to be sensitive to integration rather than retrieval costs [10]).

Hence, our overall results suggest that even if the presence of implied entities is inferable from the context, they are not necessarily expected to be mentioned in the discourse, similar to new entities. They therefore result in increased integration costs compared to given entities, which are more expected to be mentioned next.

References
[1] Bader & Meng (1999), J. Psycholinguist. Res.; [2] Kaiser & Trueswell (2004), Cognition; [3] Bornkessel & Schlesewsky (2006), J. Ger. Linguistics; [4] Arnold et al., (2013), Wiley Interdiscip. Rev. Cogn. Sci.; [5] Primus (2017), Syntax; [6] Krifka, & Musan (2012), De Gruyter Mouton; [7] Yano & Koizumi (2018), Lang. Cogn. Neurosci.; [8] Borisenkov et al., (2022), AMLaP; [9] Burkhardt (2006). Brain Lang; [10] Aurnhammer et al. (2021), PloS One.

Example 1:

Ein Bäcker ging in eine Kneipe.
(A baker went to a pub.)

Canonical word order (SOV):
  a. Ich habe gesehen, dass <u>der Bäcker</u> dort ***NP2*** … **[GIVEN]**
     (I saw that <u>the baker</u> [NOM] there ***NP2*** ...)
  b. Ich habe gesehen, dass <u>der/ein Wirt</u> dort ***NP2*** ... **[IMPLIED]**
     (I saw that <u>the/an innkeeper</u> [NOM] there ***NP2*** ...)
  c. Ich habe gesehen, dass <u>ein Maurer</u> dort ***NP2*** ... **[NEW]**
     (I saw that <u>a mason</u> [NOM] there ***NP2*** ...)
Non-Canonical word order (OSV):
  d. Ich habe gesehen, dass <u>den Bäcker</u> dort ***NP2*** ... **[GIVEN]**
     (I saw that <u>the baker</u> [ACC] there ***NP2*** ...)
  e. Ich habe gesehen, dass <u>den/einen Wirt</u> dort ***NP2*** ... **[IMPLIED]**
     (I saw that <u>the/an innkeeper</u> [ACC] there ***NP2*** ...)
  f. Ich habe gesehen, dass <u>einen Maurer</u> dort ***NP2*** ... **[NEW]**
     (I saw that <u>a mason</u> [ACC] there ***NP2*** ...)

(Experiment 1: a, c, d, f; Experiment 2: a, b, d, e; Experiment 3: b, c, e, f)
* In conditions b and e (Implied) the article was chosen based on the other condition respectively (Given-Implied → indefinite; Implied-New → definite) in order to contrast the two conditions
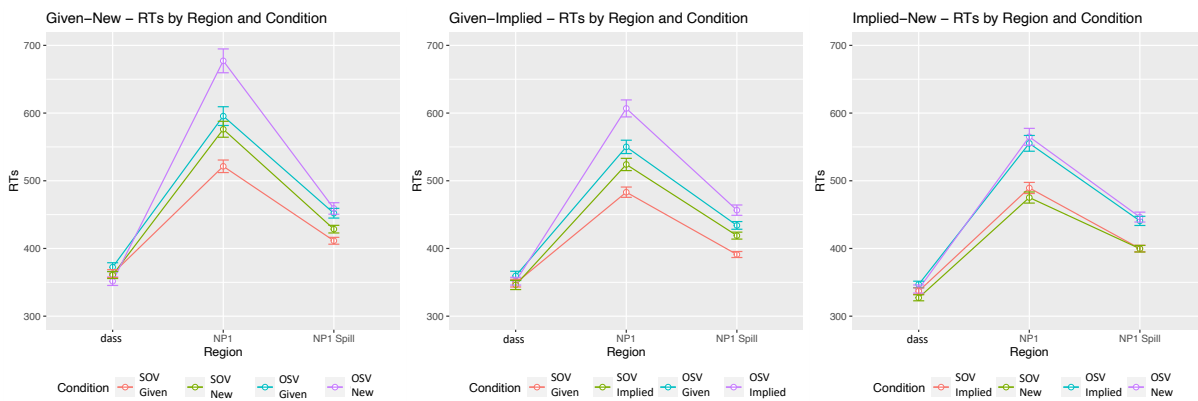

Fig.1

	Exp1- Given - New		Exp2- Given - Implied		Exp3- Implied - New	
	NP1	Spill-over	NP1	Spill-over	NP1	Spill-over
Word Order	***	***	***	***	**	**
Givenness	***	.	**	**	–	–
Interaction	–	–	–	–	–	–

Table 1 – Results   – - no effect , . – p < 0.1 , * - p < 0.05 , ** - p < 0.01 , *** - p < 0.001