The mouse is pulling the hedgehog. Or the other way around?
Non-canonical Word Order Comprehension in Czech and German Four-Year-Olds

Keywords: word order; case marking; first language acquisition; Czech; German

The preference for canonical word order (often SVO) has been shown in many languages that exhibit some word-order variation, including Czech and German (Slobin & Bever 1982; Ditmar et al. 2008; Smolík 2015). However, Czech-learning children appear to comprehend the non-canonical OVS order earlier than their German-learning peers.

Since the existing studies were not directly comparable, we present a cross-linguistic design allowing for a direct comparison. Sentences with the same pattern of case ambiguity and word order were constructed in Czech and German to describe an identical set of pictures while maintaining the structural difference: case marking in Czech is achieved by noun endings, while German noun phrases are inflected on articles.

Each stimulus consisted of two pictures with reversed role assignment (the mouse pulling the hedgehog vs the mouse being pulled by the hedgehog) and one sentence corresponding to one of the two pictures. The sentences always included one noun (phrase) ambiguous between the nominative and accusative, the other was unambiguous. Combining the order of nouns and roles resulted in four conditions (Table 1). Each of our eight picture pairs was presented once in each condition, resulting in 32 stimuli. Each child was presented with all the stimuli in a pseudorandomized order. While looking and listening, children's gaze was tracked, and children were also asked to point to the correct picture.

The sample consists of 30 German-learning and 30 Czech-learning children (GER mean age (sd) = 51.4 (6.6), CZE 51.5 (4.8) months) with comparable results of a standardized vocabulary test (GER median percentile 78.8, CZE 70.0).

For each stimulus, pictures were coded as representing the SVO or OVS interpretation of the sentence and the proportion of SVO interpretations was applied as the dependent variable. We fitted four nested mixed-effect regression models for Czech and German data separately, each with random intercepts for items and participants: two logit models of SVO-pointing responses, and two linear models of SVO-looking proportion split into five time-segments. The results are summarized in Figures 1 and 2.

In both the pointing and the looking measure, German four-year-olds showed frequent inappropriate SVO interpretations in the non-canonical OVS condition (consistent with Dittmar et al. 2008 and others), however, in the stimuli starting unambiguously, they clearly discriminated OVS from SVO, showing comprehension in SVO but random performance in OVS (consistent with Özge et al. 2022). Czech four-year-olds discriminated well OVS from SVO regardless the initial ambiguity, both in pointing and gaze direction. In stimuli with unambiguous initial nouns, they showed a clear preference for correct SVO and OVS interpretations, although SVO was interpreted more reliably.

The results indicate that Czech-learning children are strikingly better in extracting case information for sentence interpretation, compared to their peers learning German. This can be attributed to various factors. One factor can be the structural difference in case marking (suffixes vs prenominal articles). Differences in the type and token frequency of morphologically ambiguous subjects and objects could play a role, as well as overall greater saliency of case marking in Czech.

Condition	German stimulus	Czech stimulus	English translation
SVO initially unambiguous	Der Igel zieht die Maus.	Ježek táhne myš.	*The hedgehog is pulling the mouse.*
OVS initially unambiguous	Den Igel zieht die Maus.	Ježka táhne myš.	*The mouse is pulling the hedgehog.*
SVO initially ambiguous	Die Maus zieht den Igel.	Myš táhne ježka.	*The mouse is pulling the hedgehog.*
OVS initially ambiguous	Die Maus zieht der Igel.	Myš táhne ježek.	*The hedgehog is pulling the mouse.*

Table 1 The four conditions in one example item.

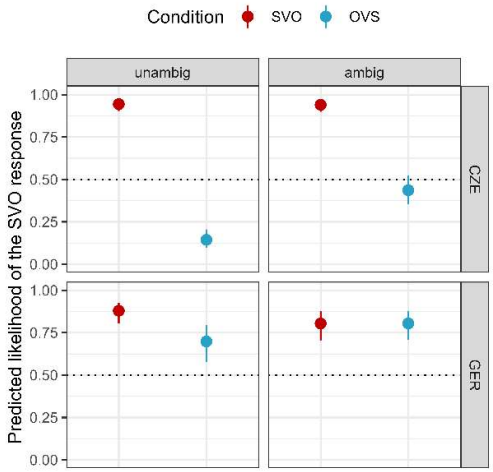

Figure 1 Conditional predictions and their 95% CI from the logit Pointing Models – predicted values from the separate CZE and GER models were bound for the purpose of this plot.

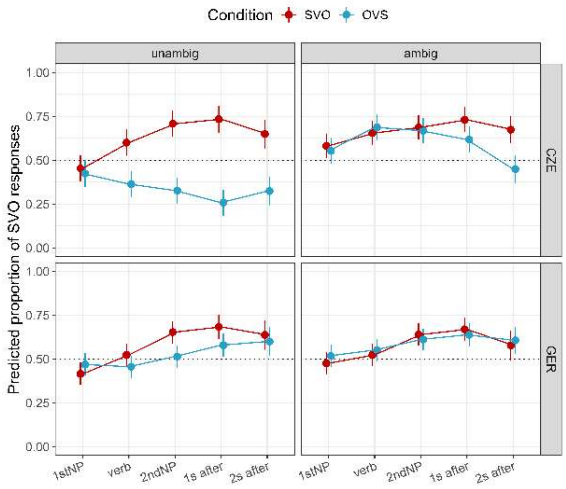

Figure 2 Conditional predictions and their 95% CI from the linear Looking Models – predicted values from the separate CZE and GER models were bound for the purpose of this plot.

References

Dittmar, M., Abbot-Smith, K., Lieven, E., & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child development, 79(4)*, 1152-1167.

Özge, D., Kornfilt, J., Maquate, K., Küntay, A. C., & Snedeker, J. (2022). German-speaking children use sentence-initial case marking for predictive language processing at age four. *Cognition, 221, 104988*.

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition, 12(3)*, 229-265.

Smolík, F. (2015). Word order and information structure in Czech 3-and 4-year-olds' comprehension. *First language, 35(3)*, 237-25.