

## Offline Task Demands Shape Online Parsing: Evidence from German Garden-Path Effects

Pia Schoknecht (pia.schoknecht@uni-potsdam.de)<sup>1</sup>, Shravan Vasishth<sup>1</sup>

<sup>1</sup>University of Potsdam, Karl-Liebknecht-Straße 24-25, 14476 Potsdam, Germany

Offline tasks such as comprehension questions and acceptability judgments are often assumed to induce different levels of processing depth, with deeper processing under questions [e.g., 2]. The present study tests whether processing depth is indeed the main difference between these frequently used offline tasks. For this purpose, we investigated whether task demands modulate garden-path effects [e.g., 4], using a classic German garden-path design [5] (see p. 3). Garden-path effects were chosen as a test case because they (arguably) cannot be attributed to shallow processing. Garden-path effects arise when in the face of temporary structural ambiguity, the parser chooses an ultimately incorrect, but at that moment fully licensed, parse. This choice is typically based on general parsing principles (e.g., minimal attachment [4]) or frequently occurring structures and meanings [e.g., 1]. If the garden-path effect is modulated by task demands, this suggests that offline tasks differ in more than just processing depth.

We conducted two word-by-word moving-window self-paced reading experiments that only differed in the offline task: binary acceptability judgments vs. yes/no comprehension questions. All participants were native speakers of German, and no participant took part in both experiments. After exclusion, the final samples comprised 227 (mean age: 27 years, 87 female, 140 male) and 241 participants (mean age: 26 years, 60 female, 181 male), respectively. Each participant read 72 sentences in total (24 critical items in a Latin Square design, 24 well-formed fillers and 24 ill-formed fillers). Ill-formed fillers (missing words, syntactically incorrect or semantically incongruent) were included in both experiments to render the judgment task meaningful and maintain consistent materials across experiments. In both experiments, a binary response was required after each sentence. The data from both experiments were jointly analyzed using maximal Bayesian hierarchical models. The garden-path effect interacted with the offline task in both the offline responses (mean estimate of the interaction: -0.81 log-odds, CrI: [-1.10, -0.51],  $BF_{10} > 800$  across investigated priors) and the online reading times of the post-critical region (mean estimate of the interaction: 50 ms, [CrI: 3, 96],  $N(0, 0.05)$ :  $BF_{10} = 3.13$ ,  $N(0, 0.1)$ :  $BF_{10} = 2.48$ ). Crucially, the garden-path effect was larger under acceptability judgments than under comprehension questions (see Table 1 and Figure 1).

In the present study, the strength of the garden-path effect was modulated by the offline task. This finding challenges the assumption that task effects primarily stem from differences in processing depth. Because garden-path effects are not typically attributed to shallow processing but instead reflect commitment to a specific syntactic analysis, their modulation implies that task demands influence parsing decisions. The larger garden-path effect under judgments suggests that this task—likely because of its focus on structural well-formedness—led to stronger commitment to an initial syntactic analysis, increasing disruption when disambiguation contradicted it. In contrast, under comprehension questions, readers showed reduced disruption upon encountering the disambiguating word. Reduced processing difficulty at disambiguation implies that, under that task, grammatical relations might have remained underspecified [6, 7].

Overall, these results suggest that offline task demands influence parsing decisions and systematically shape the magnitude of garden-path effects.

## References

- [1] K. Christianson et al. *Cogn Psychol*, 42(4):368–407, 2001.
- [2] F. Delogu, C. Aurnhammer, H. Brouwer, and M.W. Crocker. *Brain Cognition*, 186:106293, 2025.
- [3] L. Frazier. *Nat Lang Linguist Th*, 5(4):519–559, 1987.
- [4] L. Frazier and K. Rayner. *Cogn Psychol*, 14:178–210, 1982.
- [5] M. Meng and M. Bader. *Lang Speech*, 43(1):43–74, 2000.
- [6] A. J. Sanford and A. C. Graesser. *Discourse Process*, 42(2):99–108, 2006.
- [7] B. Swets, T. Desmet, C. Clifton, and F. Ferreira. *Mem Cognition*, 36:201–216, 2008.

## Figures and Tables

|                            | ambiguous     |              | unambiguous   |              |
|----------------------------|---------------|--------------|---------------|--------------|
|                            | subject-first | object-first | subject-first | object-first |
| Experiment using judgments | 81            | 55           | 84            | 83           |
| Experiment using questions | 84            | 77           | 83            | 86           |

Table 1: Raw response accuracy (% correct) in the critical trials of both experiments

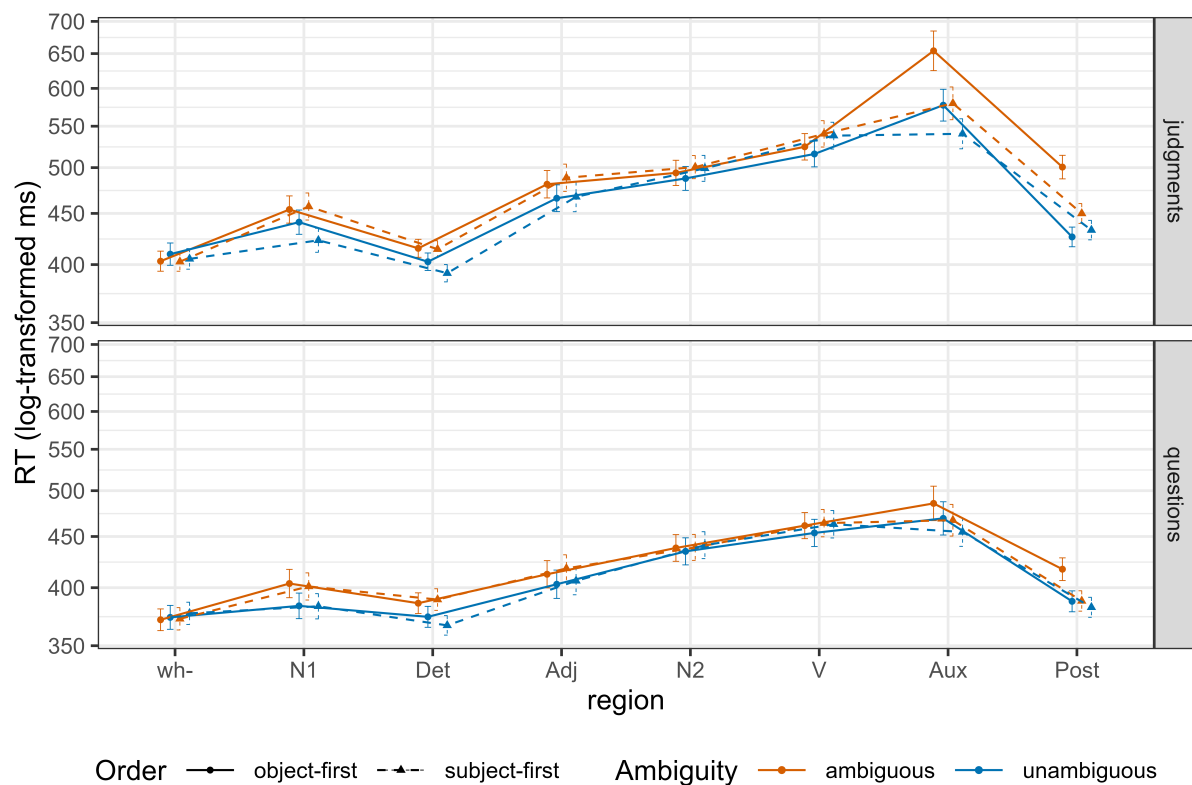


Figure 1: Reading times with 95% confidence intervals in selected consecutive regions (wh-: interrogative pronoun, N1: first noun, Det: determiner, Adj: adjective, N2: second noun, V: first part of verbal phrase, Aux: auxiliary, Post: post-critical region) of both experiments.

## Optional Supplemental Information

The design by [5] is a  $2 \times 2$  design, crossing the factors ambiguity (ambiguous, unambiguous) and word order (subject-first, object-first) and consist of four conditions, see (1).

- (1) a. Alle wollten wissen, welche Lokalspolitikerin die amtierenden  
all wanted to.know which<sub>FEM, NOM/ACC</sub> local.politician<sub>FEM, NOM/ACC, SG</sub> the sitting  
Minister<sub>MASC, NOM/ACC, PL</sub> gelobt hat<sub>SG</sub>, als die Pressekonferenz begann.  
ministers praised has as the press.conference started  
'Everyone wanted to know which local politician had praised the sitting ministers as the press conference started.'
- b. Alle wollten wissen, welche Lokalspolitikerin die amtierenden  
all wanted to.know which<sub>FEM, NOM/ACC</sub> local.politician<sub>FEM, NOM/ACC, SG</sub> the sitting  
Minister<sub>MASC, NOM/ACC, PL</sub> gelobt haben<sub>PL</sub>, als die Pressekonferenz begann.  
ministers praised have as the press.conference started  
'Everyone wanted to know which local politician the sitting ministers had praised as the press conference started.'
- c. Alle wollten wissen, welcher Lokalspolitiker die amtierenden  
all wanted to.know which<sub>MASC, NOM</sub> local.politician<sub>MASC, NOM, SG</sub> the sitting  
Minister<sub>MASC, NOM/ACC, PL</sub> gelobt hat<sub>SG</sub>, als die Pressekonferenz begann.  
ministers praised has as the press.conference started  
'Everyone wanted to know which politician had praised the ministers as the press conference started.'
- d. Alle wollten wissen, welchen Lokalspolitiker die amtierenden  
all wanted to.know which<sub>MASC, ACC</sub> local.politician<sub>MASC, ACC, SG</sub> the sitting  
Minister<sub>MASC, NOM/ACC, PL</sub> gelobt haben<sub>PL</sub>, als die Pressekonferenz begann.  
ministers praised have as the press.conference started  
'Everyone wanted to know which local politician the sitting ministers had praised as the press conference started.'

Conditions (1a, b) are temporary ambiguous. Syncretism of the nominative and accusative forms of the wh-phrase *welche Lokalspolitikerin* in its feminine form allows the wh-phrase to be interpreted as either the subject or the object of the clause. According to the subject preference [e.g., 3], comprehenders will opt to interpret the first noun in a clause as the subject of the clause, if not stated otherwise. The temporary ambiguity persists until the disambiguating auxiliary *hat/haben* is encountered. The number of the auxiliary, or more precisely the agreement of the number of the auxiliary with the nouns in the sentence, indicates the grammatical function of the nouns.

Conditions (1c, d) are unambiguous and function as control conditions. The masculine forms of the wh-phrase (*welcher Lokalspolitiker* vs. *welchen Lokalspolitiker*, see 1c, d) overtly mark nominative and accusative case. The overt case marking provides an early cue of grammatical function.

Conditions (1a, c) have subject-first order and (1b, d) have object-first order. Both orders are acceptable in German, but subject-first order is preferred. The order manipulation is confounded by the difference in word length between the critical auxiliaries *hat* and *haben*. However, this confound cancels out, because the garden-path effect in this design is captured by the interaction of ambiguity and order ( $b-a > d-c$ ). The post-critical region in all conditions, i.e., the region that might capture spill-over effects, is the word *als*.