

Filling the *vorfeld* in German written and oral discourse

Augustin Speyer

University of Pennsylvania

speyer@babel.ling.upenn.edu

German syntax requires that in archetypical declarative main clauses the finite verb is preceded by one constituent. This clause-initial position is often referred to as *vorfeld*. Syntax imposes no further constraints on the type or function of the constituent in the *vorfeld* (quite contrary to English, for example). So it is reasonable to assume that the choice of which constituent of the sentence is put to the front follows discourse structural requirements.

This paper has two parts: First it tries to identify the relevant discourse structural requirements from a corpus of monologic written text (494 sentences). Second it applies the results from this study to a corpus of informal dialogic spoken text (411 sentences). The main question is: Will there be differences between the two types of datasets, and are they related to the modular differences of the used media? Ideally, the analysis of the spoken corpus should show differences to the analysis of the written corpus, which should however be entirely due to special properties of spoken discourse. This paper argues that this is the case; especially two differences play a role, namely the distinctively oral property of intonation and the possibility – typically of written discourse – to backtrack the discourse structure and especially to go back and reconstruct the rhetorical relations in case they remain obscure after the first reading.

First to the written corpus study. The corpus used for this part of the study consists of written texts of varying degrees of formality. The study showed that mainly 3 types of elements can stand in the *vorfeld*: topics (understood as ‘centers’, see Grosz, Joshi, and Weinstein 1995), elements standing in a partially ordered set (= poset) relationship to some entity already evoked in the discourse (see Prince, 1999), and scene-setting elements. These elements are ranked as in (1)

(1) scene-setting >> poset >> topic

The topic never stands in the *vorfeld* if it is taken up again after a short sub-discourse has been inserted (a satellite of the elaboration type, but with topic shift, see Asher, and Lascarides (2003); here referred to as ‘insertion’). This indicates that in cases in which the topic is not easily recoverable the topic must be marked as such in order to

allow for an unambiguous interpretation, and a way to mark it is to leave it in its archetypical position.

In dealing with spoken discourse further characteristic properties have to be taken into account in addition to those given above: The use of pronouns to indicate the dialogue agent, the organisation in turns and a more erratic thematic structure.

The corpus used for this part of the study consisted of transcripts of informal dialogues. In principle a similar ranking as the one in (1) applies, with the addition of subject personal pronouns and the connective ‘dann’ (‘then’; ‘after that’). This corroborates the correctness of the analysis. The ranking is, however, not categorical. So, for instance, a poset-element may occupy the vorfeld even though the sentence contains a scene-setting element or a personal pronoun (2).

- (2) *so* *Komödien gefallen mir so am besten ... aber, äh, [poset Actionfilme]*
 such comedies please me so at best but uhm action movies

 mag ich nicht so arg
 like I not so much
 ‘I like comedies best, but action movies I don’t like so much.’

A possible explanation might come from the fact that prosodic well-formedness conditions play a more significant role in oral than in written discourse. One important condition is the Principle of Rhythmic Alternation (Lieberman, 1975), which applies also to the alternation of constituents in sentences (Speyer, in prep.). In German the tensed verb is typically unaccented, which means that with putting accented elements into the vorfeld-position the speaker can easily conform to the Principle of Rhythmic Alternation and minimise the danger of clash of two accented phrases in the remainder of the sentence.

Personal pronouns and ‘dann’ are ranked equally high as scene-setting elements, in that both are very prone to appear in the vorfeld (e.g. 71% of sentences with ‘dann’ have it in the vorfeld; compare: 69% of sentences with scene-setting elements in the vorfeld) and outrank poset cases and topics: If a sentence contains more than two phrases eligible for vorfeld-movement, it is always either the scene-setting element, the personal pronoun subject or ‘dann’ which stands in the vorfeld.

‘Dann’ and personal pronouns are on a par with scene-setting elements, as all are deictic and frame-setting. ‘Dann’ has in addition the function to clarify the temporal sequence (a ‘narration’-relation, in Asher, and Lascarides’ (2003) terminology). In written discourse, the relation typically stays unmarked as it can be recovered by re-reading the passage, if there is a misunderstanding. In oral discourse this is not possible and therefore the rhetorical relation needs to be marked explicitly.

Regarding another oral characteristic, the organisation in turns is irrelevant for constraints such as scene-setting and poset. It plays a role for placement and realisation of the topic however, especially after an insertion. If a topic is reintroduced after an insertion by the same dialogue agent who started the insertion,

it is realised in most cases as a pronoun in the archetypical topic position in the mittelfeld, as in monologic texts (3). If a topic is reintroduced across a turn, it is much more often in the vorfeld, always preceded by some discourse-structuring particle in the vorvorfeld (often items like *wobei* which marks ongoing coherence, see Günthner (2001)), and it is usually realised as a full noun phrase (4). In other words, it is introduced in a similar way as discourse-new entities would be introduced.

- (3) *Und dann hän [_{Topik} sie] halt immer die Bilder zammezeigt*
 and then have they ptc. always the pictures together-shown
 ‘And then they always showed the pictures together’
- (4) *Wobei, [_{Topik} strickende Männer] gabs ja eigentlich beim Grüne*
 But knitting men gave-it ptc. actually at-the green
 ‘But there were knitting men in the Green Party’

To summarise: A comparison of a written and an oral corpus with respect to information structural constraints on vorfeld-filling confirms the assumption that the same constraints play a role in principle in both corpora, but that there are slight differences which are consequences of media-specific requirements. In written, monologic discourse the *vorfeld* is filled according to a clear ranking of constraints shown in (1). In oral, dialogic discourse the same ranking applies but can be overrun by e.g. prosodic requirements, a purely oral matter of concern. The reintroduction of a topic after an insertion across turns deserves special attention, as here the reintroduction is done as if a discourse-new entity is introduced.

References

- Asher, N., and A. Lascarides (2003). *Logics of Conversation*. Cambridge University Press, Cambridge, Great Britain.
- Grosz, B.J., A.K. Joshi, and S. Weinstein (1995). Centering: A Framework for modelling the local coherence of discourse. *Computational Linguistics*, **21**: 203-225.
- Günthner, S. (2001). ‚wobei (.) es hat alles immer zwei seiten.‘ Zur Verwendung von *wobei* im gesprochenen Deutsch. *Deutsche Sprache*, **4**: 313-341.
- Liberman, M.Y. (1975). *The Intonational System of English*. PhD thesis, MIT.
- Prince, E. (1999). How not to mark Topics: ‚Topicalization‘ in English and Yiddish. In *Texas Linguistics Forum*, ch. 8.
- Speyer, A. (in prep.). A Prosodic Factor for the Decline of Topicalisation in English. In Reis, M., and S. Kepser, eds., *Linguistic Evidence*.