Through narrative planning towards the preverbal message: A DRT-based approach[*]

Katja Jasinskaja
Universtität Heidelberg/IMS, Universität Stuttgart
katja.jasinskaja@ims.uni-stuttgart.de

Antje Roßdeutscher
IMS, Universität Stuttgart
antje@ims.uni-stuttgart.de

Abstract

This paper develops a model of narrative production which incorporates language-independent and language-specific principles of discourse planning, concentrating particularly on the differences in narrative planning between English and German speakers documented in a number of psycholinguistic studies (e.g. Carroll et al., 2008). For example, one of the most striking features of German narratives is a global preference for encoding the protagonist of the story as the grammatical subject of the sentence, which shapes the narrative structure at all levels, from selection of events for verbalization, to decisions related to foregrounding and backgrounding of information, to actual subject selection on a sentence by sentence basis. In this paper preferences of this kind are implemented as systems of hierarchically organized rules which are applied to a knowledge base encoded in the framework of Discourse Representation Theory to derive typical German and English narrative solutions.

1 Introduction

The focus of this paper is on modeling human discourse production, in particular, the production of spoken narrative. Previous research in psycholinguistics has given rise to elaborate language production models (such as Levelt, 1999), which describe this process as taking a number of steps, from conceptual preparation, over grammatical (syntactic) and morpho-phonological encoding, down to phonetic encoding and articulation. Decisions related to discourse planning—what to
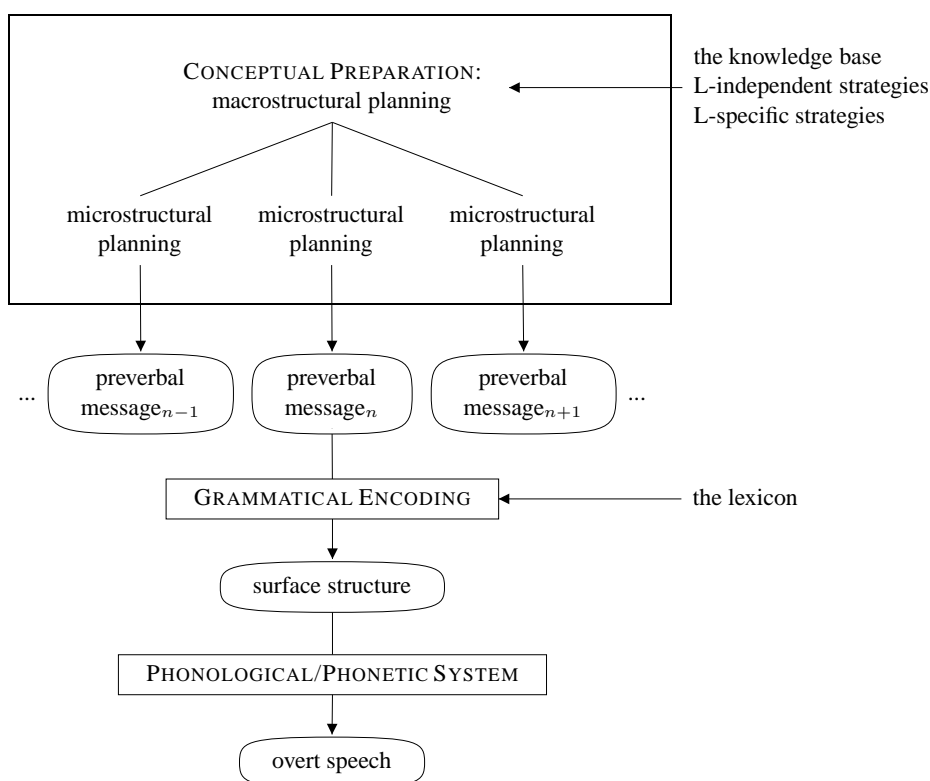

Figure 1: Macro- and microsctructural planning in production

say next—take place at the level of *conceptual preparation*, which selects fragments of a representation of knowledge about the world and turns them into a *preverbal message*. It is this module of the production process that we will concentrate on. Figure 1 shows the place occupied by this module in the overall architecture of Levelt's (1999) model of the speaker.

The preverbal message is a conceptual structure that can be expressed by words and grammar of a particular language. The standard assumption seems to be that this structure is of a size to be expressible by a single sentence. Constructing it involves *macrostructural* (what to say next?) and *microstructural planning*—the choice of lexical concepts and conceptual categories necessary for fixing the grammatical features. Obviously, since languages differ in their range of lexical concepts and (obligatory) grammatical categories, microstructural planning is not language-independent. After Slobin (1987) the language-specific aspects of conceptual preparation are referred to as 'thinking for speaking.'

Macrostructural planning deals with splitting up the speaker's overall communicative intention and content to be communicated into pieces small enough to be captured by a single sentence, and sequencing those pieces. This is the proper domain of theories of discourse structure. There is a widely-spread assumption both among psycholinguists and among discourse structure theorists that planning at

this level is governed by language-independent principles, based on the principles of general pragmatics and general cognitive mechanisms. Indeed, these general mechanisms play the most central role in discourse planning. However, there is a growing body of evidence that considerations of microstructural planning can influence the decisions at the macro-level (Carroll et al., 2008, and references therein). That is, during macroplanning the speaker already takes into account the availability of lexical concepts and the demands of the grammatical system of her language in order to make her task at the microlevel easier. In other words, macrostructural planning has a language-specific component as well.

The goal of the present paper is to develop this line of thought and to integrate the language-specific aspects of discourse planning into a general theory of discourse structure. We will concentrate especially on modeling the differences in narrative planning between English and German speakers based on the findings of Carroll et al. (2008). This task issues at least three challenges.

First, theoretical research in discourse structure and semantics has developed high standards of formalization. The integration task requires meeting comparable standards in modeling 'thinking for speaking.' We will build on our previous efforts (Jasinskaja and Roßdeutscher, 2008) of implementing this component in the framework of Discourse Representation Theory (DRT, Kamp and Reyle, 1993).

Second, studies in discourse, pragmatics and context-sensitive semantics have made substantial progress in understanding the hierarchical organization of discourse, and narrative in particular, as well as the phenomena of anaphora and presupposition. The general (language-independent) mechanisms that stand behind them play a central role in macrostructural planning. The division of labor between language-independent and language-specific principles will be somewhat redefined in this paper as compared to the versions presented in Carroll et al. (2008) and Jasinskaja and Roßdeutscher (2008), granting a more prominent place to language-independent aspects and the appreciation of the extent to which the two aspects of macroplanning are interwoven. However, the existing theories of discourse structure, anaphora and presupposition usually take the interpretation perspective, and it is not so trivial to turn them around in the direction of production. For instance, it is one thing to specify the way in which a presupposition introduced by some linguistic trigger finds its antecedent or is accommodated, but it is another thing to define the conditions under which a certain bit of content may or must be asserted, presupposed in the sense that it is verbalized and marked as a presupposition, or presupposed in the sense that it need not be verbalized at all. In the ideal case we should have an explicit specification of the latter set of conditions which we can directly combine with the relevant language-specific conditions in conceptual preparation. It will not be possible to provide a comprehensive specification of this sort in this paper, so we will have to take rather rough shortcuts and mostly take the effects of the language-independent principles for granted, concentrating primarily on spelling out the language-specific component.

Finally, the task of modeling conceptual preparation presupposes that there is an initial knowledge representation which serves as the starting point of the con-

ceptualization process. It does not necessarily mean that this representation is absolutely pre-conceptual, but we will least assume that it is language-neutral. There is not so much work in formal discourse semantics that goes beyond assigning lexical concepts to words. Some DRT-based approaches to lexical semantics (e.g. Kamp and Roßdeutscher, 2005) analyze lexical concepts in terms of more basic categories that will be useful for a language-neutral representation of the kind we need. However, there are still gaps to be bridged. For our present purposes, an appropriate notion of event needs to be developed.

The structure of this paper is as follows. Section 2 recapitulates the findings of Carroll et al. (2008) on language-specific aspects of macrostructural planning in English and German. Section 3 presents the initial knowledge representation that serves as input of the conceptual preparation module, whereas section 4 models language-specific aspects of narrative planning in interaction with general language-independent principles of discourse organization.

2 Background

In our assumptions about the role of language-specific considerations at the level of macrostructural planning we will refer primarily to the evidence collected in a number of studies represented by Carroll et al. (2008, and references therein). These studies are based on a multilingual corpus of spoken narratives in which the participants all retell the same animation film. The film shows a clay figure searching for water in a series of desert-like environments: a sand desert, a paper desert, a stone desert, etc. We will concentrate on the sequence taking place in the paper world, where the earth's surface is made of pieces of paper and some pieces are carried around by the wind. On his way through this desert, the clay man is twice in danger of being overrun by a large flying piece of paper, whereas the third time he is actually knocked down by one piece. He finally finds a small puddle of water, but while he is trying to collect the water from the ground, the wet paper breaks in and he falls through the hole into the next world. Typical solutions of an English and a German speaker are given in (1) and (2).

(1)   a. and you see him land on a new plane [...]
      b. and eh he wakes up
      c. and eh just as he is opening his eyes from the fall [...]
      d. he looks in front of him
      e. and there's this piece of paper coming straight for him
      f. and he's scared
      g. and he jumps up onto his knees
      h. and the piece of paper misses him

First, the solutions differ in that the English speaker establishes a causal chain in (1), whereas this is not so in (2), e.g. doing some paces in (2b) does not naturally

4

(2) a. als nächstes sieht man das männchen fallen [...]
        as   next   sees   one   the   man   fall

        'next one sees the man fall'

     b. und Ø geht n' paar   schritte
        and   goes a   couple steps

        'and [he] makes a few steps'

     c. Ø   wird dann von so   nem wehenden blatt umgeworfen
            gets then by   such a   flying   sheet knocked over

        'then [he] gets knocked over by such a flying sheet'

     d. und Ø hört   auf einmal wieder dieses tropfen
        and   hears suddenly   again   this   dripping

        'and [he] suddenly hears again this dripping of water'

     e. und Ø sieht dann auch n' äh n stück papier
        and   sees   then also a   eh a piece paper

        'then [he] also sees a piece of paper'

     f. wo   also so   ne kleine pfütze schon   is
        where thus such a   small   puddle already is

        'where there is already a little puddle'

     g. und er kniet   sich   dann davor
        and he kneels himself then in front of it

        'then he kneels in front of it'

lead to be knocked down, cf. (2c).[1] It is the use of sentence-internal *dann* 'then' that establishes narrative connection between the two event descriptions, see also (2g). Second, the German speaker skips talking about all the dangerous situations except the one where the protagonist gets directly affected, i.e. knocked over by the piece of paper, cf. line (2c), whereas the English speaker also describes in detail the first situation where the man manages to avoid the collision. If German speakers tried to do the same, we would have discourses like (3), but we do not find them in the corpus, although (3) is coherent. It is very untypical for German speakers to fill the subject position with an object other than the protagonist as in (3c) and (3e).

(3) a. und er   wird dann wach
      and he   gets   then   awake

      'and then he wakes up'

   b. und öffnet dann die Augen
      and opens   then   the eyes

      'and then he opens his eyes'

   c. und da   kommt gerade ein Papier auf ihn zu
      and there comes   right   a   paper   up   him to

      'and there is a paper coming towards him'

   d. und er   springt auf
      and he   jumps   up

      'and he jumps up'

   e. und das Papier fliegt vorbei
      and the paper   flies   past

      'and the paper flies past him'

This illustrates the generalization proposed by Carroll et al. (2008) that German speakers select the protagonist as a global topic for the whole narrative, which then systematically surfaces as the grammatical subject of the clauses, whereas the events in which the protagonist is not or only marginally involved (i.e. without dramatic consequences for himself) are simply not narrated. Thus the typical German narrative can be grosso modo characterized as following the instruction: "Select events for description that are changes of properties of the protagonist; connect those descriptions by anaphoric temporal adverbials." In contrast, the majority of English speakers produce solutions that could be characterized as following the

---

[1]The symbol ∅ in (2) and in other German examples represents the omitted subject in a non-technical sense. It is not to be understood in the syntactic sense of null subject in prodrop languages. We follow here the widely spread tradition in discourse analysis to treat each event description in a narrative as a separate utterance, or elementary discourse unit, with no implications for the syntactic analysis intended. Whether the utterances (2a)–(2e) constitute coordinated clauses with proper null subjects, or coordinated VPs with one clause spanning over several utterances is immaterial for the present discussion.

instruction: "Introduce some fictitious witness time T_{now} of a narrator witnessing the story; fill the period T_{now} by producing continuous event descriptions, either 'shot by shot' or causally connected."

Why do the English and the German speakers assess the narration task differently? First of all, one should take into account that it is remarkably difficult to produce a coherent narrative spontaneously. Therefore experienced story-tellers develop narration strategies that make their task easier. What counts as an "easy option" in turn depends on the grammar of the language in question, i.e. it is crucial which grammatical features require obligatory marking:

(4) **Leading Hypothesis**. If it is obligatory to mark a grammatical feature sentence by sentence, the speaker will choose its value globally.

This means that the speaker will set the value of the obligatory feature as a default for all the sentences in the discourse, although it can be overridden locally for the sake of coherence.

In verb-second languages the preverbal position must be filled by exactly one constituent and that constituent may, but need not be the subject of the sentence. A German speaker must decide not only which salient discourse referent to select for subject but also *where* to place it. A global decision is advantageous here because it reduces the set of decisions to be made for each particular sentence to either placing the globally selected subject or a temporal adverb in the sentence-initial position. Choosing the subject is ranked higher in the decision process than event selection. Whether or not an event is selected for description depends on whether or not that event is suitable for being described with the protagonist as subject. If the selected events turn out not to be causally connected the temporal adverbial guarantees establishing the story line, cf. (2g).

Speakers of a strict SVO language have no choice as to where to place the subject, therefore a global selection of the subject has no strategic advantage. However, in contrast to German speakers, who neither have the obligation nor the possibility to choose aspect, English speakers have to make that choice in each sentence. Some of them opt globally against progressive and at the same time against the introduction of the narrator and his witness-time T_{now} to be filled by a causal chain. Their narration solutions are similar to those of the German speakers. The majority of English speakers, however, make their global choice for the progressive and the deictic narrator-centered perspective on the story events. For example, the speaker of (1) starts and ends episodes in the progressive (not documented in the lines). Of course, a causal chain cannot be established unless the speaker locally switches to perfective aspect. E.g. in (1f), (1g) and (1h) the global default decision for the progressive is overridden for the sake of providing causal antecedents for their successors.

If event selection follows the causal chain, this is unsuitable for if not incompatible with any further restriction on subject selection. As a consequence English speakers simply select the salient agent of the event for subject and change the

topic if the salient agent changes. To sum up: (4) leads to different *hierarchies of decisions* necessary for describing or non-describing a particular event.

The decision hierarchy rules out discourses like (3) because the events in (3c) and (3e), where the protagonist is neither the agent nor is strongly affected by the action, are either filtered out by global event selection criteria and will not be verbalized, or, if the speaker locally undoes that global decision and verbalizes them for the sake of coherence, the piece of paper will be demoted to a non-subject position, the subject position being reserved for the protagonist. Thus wordings like *er sieht ein Papier auf sich zukommen* (he sees a paper coming) and *er weicht dem Papier aus* (he avoids the paper) would be generated for (3c) and (3e) instead. The passive voice in (2c) shows a further typical demotion operation.

This is a summary of the generalizations proposed by Carroll et al. (2008) which are adopted for the most part in this paper, with some adjustments concerning the division of labor between the language-specific and language-independent principles in narrative planning which will be discussed in section 4.

3 Knowledge representation

The production process begins with a knowledge representation describing the participants' knowledge of the content of the film, which we assume to be shared by all speakers. At one level, it contains referents for objects that can be seen in the film, as well as time points and intervals of the film time, and DRS-conditions for propositions that can be judged true on the basis of the picture and the sound track. At another level that representation is enriched with all sorts of common-sensical inferences—the "pragmatic" interpretation of the film—in particular causal relations between states and changes of state, inferences about the normal consequences of some events (e.g. if this sheet of paper continues moving and the man does not move, it will crash on him); and finally, ascriptions of perception and mental states—beliefs, emotions, intentions—to the protagonist. A fragment of this representation is shown in Fig. 2. The DRS describes the episode where the protagonist (m) wakes up, sees a sheet of paper ($sheet_1$) flying towards him, and avoids it by sitting up, cf. (1).

States, events, and causality: For our purposes we identify states as time-dependent propositions, i.e. a pair of a (definite) proposition and a time interval during which it holds, e.g. s_1, s_2 in Fig. 2. A special kind of "state" is a gradual change of some property over a period of time, cf. s_6, s_{47}. Events are instant changes from a proposition p to $\neg p$, identified by a pair of states whose first elements are p and $\neg p$ respectively, and the second elements are time intervals such that the end point of one interval is the starting point of the other, cf. e_1 in Fig. 2. As an abbreviation, events will sometimes be represented as: $e : (s_x, s_y) : t : \textsc{Become}(p)$, meaning that a change from state s_x characterized by $\neg p$ to state s_y characterized by p occurs at time point t (the information on the starting time of s_x and the end time of

$m, head, eyes, spine, legs, sheet_1, w_2, t_1...t_{21}, t^{1:25}...t^{2:27}, s_1...s_{57}, e_1...e_{18}$

$s_1 : \left(\neg open(eyes), \langle t^{1:28}, t_1^{1:31} \rangle\right)$

$s_2 : \left(open(eyes), \langle t_1^{1:31}, t^{1:37} \rangle\right)$

$e_1 : (s_1, s_2), \quad e_1 \Rightarrow s_2$

$t_1 \le t^{1:29}, \quad t^{1:37} \le t_2$ (the whole movement)

$s_6 : (move(sheet_1), \langle t_1, t_2 \rangle)$ (the whole movement)

$s_7 : (move(sheet_1), \langle t_2^{1:31}, t^{1:34} \rangle) ; s_7 \subset s_6$ (part 1 visible to m)

$s_8 : (\textsc{PathSurf}(sheet_1, w_2), \langle t_1, t_2 \rangle)$

$s_{10} : (\textsc{Obs}(m) \parallel \langle m, sheet_1 \rangle), \langle t_3, t^{1:34} \rangle), \quad t_1 \le t_3 \le t_2^{1:31}$

$s_{11} : (\textsc{Align}(\textsc{Obs}(m), \langle m, sheet_1 \rangle), \langle t_3, t^{1:34} \rangle)$

$s_{14} : \left(\neg see(m, s_7), \langle t^{1:25}, t_2^{1:31} \rangle\right)$

$s_{15} : \left(see(m, s_7), \langle t_2^{1:31}, t^{1:34} \rangle\right)$

$e_3 : (s_{14}, s_{15}); \quad s_2 \Rightarrow e_3; \quad s_{11} \Rightarrow e_3; \quad s_7 \Rightarrow e_3; \quad e_3 \Rightarrow s_{15}$

$s_{16} : \left(\textsc{IncSurf}(head, w_2), \langle t_2^{1:31}, t_2^{1:33} \rangle\right)$

$t_6 \le t_2^{1:32}, t^{1:34} \le t_7 \le t^{1:37}, t_{20} \approx t^{1:34}, t_{20} \le t_{21}$

$s_{22} : \left(\neg believe\left(m, [s_{29} : (\textsc{ApplyForce}(sheet_1, m), \langle t_{20}, t_{21} \rangle)]\right), \langle t^{1:25}, t_6 \rangle\right)$

$s_{23} : \left(believe\left(m, [s_{29} : (\textsc{ApplyForce}(sheet_1, m), \langle t_{20}, t_{21} \rangle)]\right), \langle t_6, t_7 \rangle\right)$

$e_5 : (s_{22}, s_{23}); \quad s_8 \Rightarrow ... \Rightarrow e_5; \quad s_{15} \Rightarrow ... \Rightarrow e_5; \quad s_{16} \Rightarrow ... \Rightarrow e_5; \quad e_5 \Rightarrow s_{23}$

$t_{12} < t_{13}$

$s_{32} : (\neg scared(m), \langle t^{1:25}, t_{12} \rangle)$

$s_{33} : (scared(m), \langle t_{12}, t_{13} \rangle)$

$e_8 : (s_{32}, s_{33}); \quad s_{23} \Rightarrow e_8; \quad e_8 \Rightarrow s_{33}$

$t_6 \le t_{14}; \quad t_{14} < t_{15}; \quad t_{14} \le t_1^{1:33}; \quad t^{1:34} \le t_{15}$

$s_{36} : (\neg intend(m, \textsc{Cause}(m, \neg \textsc{ApplyForce}(sheet_1, m))), \langle t^{1:25}, t_{14} \rangle)$

$s_{37} : (intend(m, \textsc{Cause}(m, \neg \textsc{ApplyForce}(sheet_1, m))), \langle t_{14}, t_{15} \rangle)$

$e_{10} : (s_{36}, s_{37}); \quad s_{23} \Rightarrow e_{10}; \quad e_{10} \Rightarrow s_{37}$

$t_{14} \le t_{18}; \quad t_{18} < t_{19}; \quad t_{18} \le t_1^{1:33}; \quad t^{1:34} \le t_{19}$

$s_{40} : (\neg intend(m, \textsc{Cause}(m, \textsc{Align}(spine, \textsc{vert}))), \langle t^{1:25}, t_{18} \rangle)$

$s_{41} : (intend(m, \textsc{Cause}(m, \textsc{Align}(spine, \textsc{vert}))), \langle t_{18}, t_{19} \rangle)$

$e_{12} : (s_{40}, s_{41}); \quad s_{37} \Rightarrow ... \Rightarrow e_{12}; \quad e_{12} \Rightarrow s_{41}$

$s_{46} : (\neg [\angle(spine, \textsc{hor}) \text{ changes from } 0° \text{ to } 90°], \langle t^{1:25}, t_1^{1:33} \rangle)$

$s_{47} : (\angle(spine, \textsc{hor}) \text{ changes from } 0° \text{ to } 90°, \langle t_1^{1:33}, t^{1:34} \rangle)$

$e_{15} : (s_{46}, s_{47}) \quad s_{41} \Rightarrow e_{15}; \quad e_{15} \Rightarrow s_{47}$

$s_{50} : (\neg \textsc{Align}(spine, \textsc{vert}), \langle t^{1:25}, t^{1:34} \rangle)$

$s_{51} : (\textsc{Align}(spine, \textsc{vert}), \langle t^{1:34}, t_1^{1:55} \rangle)$

$e_{17} : (s_{50}, s_{51}) \quad s_{41} \Rightarrow e_{17}; \quad s_{47} \Rightarrow e_{17}; \quad e_{17} \Rightarrow s_{51}$

$s_{53} : (\neg \textsc{IncSurf}(head, w_2), \langle t_2^{1:33}, t^{2:27} \rangle)$

$e_{18} : (s_{16}, s_{53}) \quad s_{47} \Rightarrow e_{18}; \quad e_{18} \Rightarrow s_{53}$

$s_{55} : (legs \parallel \textsc{hor}, \langle t^{1:25}, t_1^{1:44} \rangle)$

$s_{56} : (\neg \textsc{ApplyForce}(sheet_1, m), \langle t^{1:25}, t^{2:27} \rangle)$

$s_{57} : (\neg \textsc{ApplyForce}(sheet_1, m), \langle t_{20}, t_{21} \rangle)$

$s_{53} \Rightarrow s_{57}$

Figure 2: A fragment of the DRS representing the interpretation of the film

9

s_y is lost in this notation). For example, the event e_1 of the man's eyes going open can be described as: $e_1 : (s_1, s_2) : t_1^{1:31} : \textsc{Become}(open(eyes)).$[2]

The relation of immediate cause $\Rightarrow$ holds between a state and an event if the state is one of the necessary preconditions for the event to happen, and overlaps with the event in time, e.g. the protagonist m looking in the direction of the sheet of paper $sheet_1$ (s_{11}), as well as his eyes being open (s_2) are preconditions for him getting to see $sheet_1$ ($s_2 \Rightarrow e_3$; $s_{11} \Rightarrow e_3$). An event immediately causes its poststate, e.g. $e_1 \Rightarrow s_2$. A state can also directly cause another state, e.g. the wind causes the movement of the sheets of paper.

Spacial objects and relations: The language for describing spacial objects and relations is largely borrowed from Kamp and Roßdeutscher (2005). The space is defined by the horizontal plane HOR and the axis VERT (directed upward). The participants of the scene can be represented as 1-dimensional (the man's body parts, e.g. $spine$, $legs$), 2-dimensional (the sheets of paper) and 3-dimensional objects (e.g. the protagonist). Partly depending on the number of dimensions, the objects can have a number of functional projections such as $\textsc{Front}(\cdot)$, $\textsc{Vert}(\cdot)$, and $\textsc{LeftR}(\cdot)$, so the orientation of objects in the space and with respect to other objects can be characterized in terms of parallelism ($\parallel$) and alignment ($\textsc{Align}(\cdot, \cdot)$) between functional projections and the absolute HOR and VERT. For example, the man's spine is a 1D-object that only has a vertical projection, for brevity we will write $spine$ for $\textsc{Vert}(spine)$, so its vertical orientation in space is represented as $\textsc{Align}(spine, \textsc{Vert})$, cf. state s_{51}. Rectilinear movement is characterized in terms of paths, in particular *surface paths*, e.g. w_2, cf. $\textsc{PathSurf}(sheet_1, w_2)$—directed line segments that characterise the projection of the object on a horizontal surface during its movement. The position of the man's head on the path of $sheet_1$ is represented as $\textsc{IncSurf}(head, w_2)$, cf. s_{16}. Finally, the observer axis $\textsc{Obs}(\cdot)$ defines the perspective upon the spacial setting, $\textsc{Obs}(m)$ corresponds to the protagonists perspective and in practice is identified with the direction of his look. Thus $\textsc{Align}(\textsc{Obs}(m), \langle m, sheet_1 \rangle)$ states that the man's look is pointing in the same direction as the line segment from the man to $sheet_1$, i.e. the man is looking in the direction of $sheet_1$.

---

[2]This definition of events and states is based entirely on their temporal structure in the film world and is determined independently of the way they are expressed in language. This is different from the Davidsonian notions of event and state as they are usually applied in formal semantics, including DRT. There it depends partly on the aspectual form of the main verb whether the sentence describes an event or a state . Thus what counts as a state or an event depends ultimately on what relevant distinctions are made by the grammar of a particular language. It is unclear that such a notion is suitable for representing non-linguistic knowledge (interpretation of a film), which we assume to be language-neutral and shared by all participants of the study regardless of their first language.

4 Strategies in narrative planning

In the next step, the above knowledge representation is turned into the preverbal message by a set of rules that implement the narrative planning module, and determine whether a particular event is narrated or not, whether it appears in the foreground (e.g. a main clause), or in the background (secondary predication), which in turn has consequences for the subject and predicate selection.

4.1 General remarks

Both universal (pragmatic) narrative planning principles and language-specific techniques are at work here. This paper is concerned primarily with the language-specific side, but a few words should be said about the universals since they set the frame for cross-linguistic variation. First of all, the story of the film naturally contains more central and more peripheral (clusters of) events. In the paper scene, the most central elements are the protagonists' desire to find water, his attempt to dig for it in the paper which leads to the paper breaking through and him falling into the hole. Even the least verbose speakers recount these eventualities, regardless of their first language. Example (5) shows such a minimal solution in English.

(5) a. poor guy
   b. he gets eh/he lands softly in the world of paper
   c. and eh he walks through the wishy surface
   d. until he hears again the drop of water
   e. and he sees a puddle on the paper
   f. and tries to find more
   g. and digs through
   h. and falls through again

It is not our job here to discuss why exactly these events are so prominent in the story structure. Presumably, it is because the search for water and failure to find it is the theme going through all the five scenes of the film. The prominent status of these events and their higher likelihood of being mentioned can be accounted for in almost any theory of discourse that assigns it a hierarchical structure, e.g. Story Grammar (Rumelhart, 1975) or Rhetorical Structure Theory (RST, Mann and Thompson, 1988). The central events of the story will occupy a relatively high position in that hierarchy (be nuclei). Then assuming that a certain numerical variable x reflects the degree of detail a particular speaker adopts in their retelling, with lower values indicating less and higher values indicating more detail, speakers with lower x will only access events at higher levels of the structure, whereas speakers with greater x will both access the higher and the deeper levels. This idea is the basis of various approaches to text summarization (Rumelhart, 1977; Lehnert, 1981; Marcu, 1997).[3] For our present purposes let's assume that certain

---

[3]The application of this idea to our task presupposes that the units in the knowledge base described in section 3, cf. figure 2, are organized in some kind of hierarchical structure similar to story

events are obligatorily selected by these general mechanisms, so their mention is not subject to individual or cross-linguistic variation.

The second language-independent principle that appears to play a central role in event selection in our data is (6). By contraposition, if you have skipped an important cause you cannot narrate the consequence.

(6) If an event is selected for narration then (some of) its causal antecedent(s) must be selected as well.

To some extent this idea is also reflected by Story Grammar where the structure of an episode has 'event' and 'reaction' as obligatory parts (e.g. Rumelhart, 1975). However, we assume a rather broad notion of causality (cf. section 3) as a relation between a precondition and a change of state, and a change of state and its post-state. Under this view, an event can have lots of causes and not all of them necessarily have to be named explicitly—some can be presupposed. Which ones can, depends in part on the general conditions for presupposition accommodation (Beaver and Zeevat, 2007). For example, the event where the protagonist sees the piece of paper rolling at him (e_3 in figure 2) is caused by at least three states: the piece of paper actually rolling in his direction (s_7), the man's eyes being open (s_2), and his face being turned in the direction of the moving paper (s_{11}). There is nothing unusual about s_{11} so it can be easily accommodated. Similarly, a normally functioning person keeps her eyes open most of the time, so s_2 need not be mentioned either, unless the speaker had previously stated that the protagonist was asleep or unconscious at the beginning of the scene, as some speakers do, cf. (1). In that case s_2 must be mentioned (usually by reference to the event e_1 of the man opening his eyes), otherwise the state of the man's eyes being closed (s_1) persists, so accommodation of s_2 fails for inconsistency. Finally, papers do not normally move about by themselves in our world, so the fact that the paper is moving towards the man (s_7) cannot be easily accommodated and must be mentioned (though it is enough if it appears as the object of *see*: 'He sees a piece of paper flying at him'); moreover, the cause of that process must normally be named as well (the wind), or at least it must be introduced as a peculiar property of the film world.

One practical consequence of the rule in (6) is that an event cannot be picked

grammar or RST-style structures. But how should the fact that these structures characterize *texts* be reconciled with the fact that our knowledge base is intended as a representation of the film *content*, and in that sense is prior to text? On the one hand, one should not forget that the knowledge of the film content is the result of the interpretation of the film, which can be seen as a kind of non-verbal discourse with its own structure. Aspects of that structure should then be reflected in our knowledge base. On the other hand, Wilensky (1983) has argued that certain aspects of hierarchical "story structure" characterize primarily the mental or conceptual representation, rather than text, and are even used to structure personal experience which is not acquired through discourse interpretation. Theories like Segmented Discourse Representation Theory (SDRT, Asher and Lascarides, 2003) also locate discourse structure at the conceptual level, i.e. in the DRS. However our task would require a much more coarse-grained structure than what is suggested by the standard SDRT to leave space for the great amount of individual variation in the discourse structure at local levels.

out by whatever event selection principles happen to apply without the whole causal chain leading up to it (restricted to the "important" causes in the above sense). Such chains can be thought of as episodes, so either a whole episode is selected for narration, or the whole episode is skipped.

Finally, where the universal principles leave space for variation the language-specific strategies apply, such as (7) for German or (8) for English.[4] These strategies have already been described informally in section 2. Here we formalize them in terms of decision hierarchies:

(7) German:
$$\left\{ \begin{array}{l} \{ \langle \text{Subj(e)=m} \rangle, \langle \boxed{t_{dann}} \mid V \boxed{\text{Subj}}, \boxed{\text{Subj}} \mid V \boxed{t_{dann}} \rangle \} \prec \\ \{ \langle \text{select(e)}, \neg \text{select(e)} \rangle, \langle \text{predicate(e,m)} \rangle \} \end{array} \right\}$$

(8) English:
$$\left\{ \begin{array}{l} \left\{ \left\langle \begin{array}{c} \text{select}(e) \\ \neg\text{select}(e) \end{array} \right\rangle, \left\langle \begin{array}{c} \text{add}(s_p : (perceive(n, e), T_{now})), \text{select}(s_p) \\ \neg\text{add}(s_p : (perceive(n, e), T_{now})) \end{array} \right\rangle \right\} \prec \\ \{\langle \text{Subj}(e)\rangle, \langle \text{predicate}(e)\rangle\} \end{array} \right\}$$

In particular, the German decision hierarchy (7) says that the selection of the protagonist (m) as subject and of the word order options is prior to event and predicate selection. In contrast, for a typical English speaker event selection is prior to and independent of the choice of subject, cf. (8). Instead, the speaker takes an early decision on whether the selected eventualities should be presented as perceived by the narrator (n). If yes, the knowledge representation is enriched with a perception state s_p extending over a time interval T_{now} (extended 'now', witness time), and that state is selected for narration. Since s_p is a state extended in time the eventuality being perceived (e) also has to be extended in time.[5] This creates a preference for applying the narrator's perceptual frame to states and processes which typically surface in the progressive, e.g. *you can hear winds blowing*. The proper punctual events then have to be assembled into blocks embedded in T_{now}. If the speaker decides against introducing the narrator's perspective, it makes sense to give preference to the punctual events since they drive the narrative forward. They will be typically expressed in simple present and the overall structure of the narrative ends up being rather similar to the German one. In any case, the global choice to introduce the narrator's perspective or not goes together with a global choice between the progressive or the simple as the dominant aspectual form, which is advantageous for speakers of a language with obligatory grammatical aspect marking, as was pointed out above. The results of applying these strategies to our knowledge base with regard to event selection, foregrounding/backgrounding of events, and subject and predicate selection are presented in detail in the next sections.

---

[4]Read e.g. $\boxed{t_{dann}} \mid V \boxed{\text{Subj}}$ as a temporal adverbial in the prefield (before V), the subject in the middle field.

[5]Obviously, something that takes a single instant to happen, normally also takes a single instant to see.

4.2 Event selection

As was pointed out above, the causal chain leading up to the resolution of the scene (the man falling through the hole), illustrated by the minimal English solution in (5), is always selected because of its prominence in the story structure. The episodes in which the protagonist is fighting with the flying sheets of paper are causally related to that sequence but deliver optional causes which can be accommodated. One of the preconditions for the hero being able to pursue his goal of finding water is that he is properly functioning. On the one hand, proper functioning is a property of human beings that we assume by default. On the other hand, it is the result of the fact that the man is successful in overcoming the difficulties he encounters. This leaves the speaker with two options: either not narrate those episodes at all, in which case the proper functioning state of the protagonist is accommodated, or if the speaker has mentioned the occurrence of one of the dangerous situations then proper functioning cannot be accommodated anymore, so the speaker also has to state that the man manages to avoid the dangerous consequences or regains his health to a sufficient degree after being affected. In other words, the universal narrative planning principles such as prominence in the story structure and causality leave it open whether the episodes with the flying sheets of paper will be selected for narration or not. Thus this is where we expect to and do find the most individual and cross-linguistic variation in event selection.

Next, it can be shown that the German speakers will encounter more difficulties in recounting these episodes than the English speakers given the differences in the language-specific narrative strategies (7) vs. (8). The crucial role is played by the fact that the Germans select the protagonist as the grammatical subject globally and prior to making other structural decisions, cf. $\langle Subj(e) = m \rangle$ in (7). It is as if they assessed the narration task by reducing it to answering the question *What does m do?* Of course, events where the man does not do anything, i.e. where he is not involved at all or where he has a low proto-agent status and is therefore not likely to surface as subject (Dowty, 1991) do not conform to this question and will be filtered out.

Notice that the structure of the relevant episodes is such that their causal chains are initiated by one or the other sheet of paper moving in the hero's direction—a process that does not involve the hero himself, e.g. s_7 in the episode with $sheet_1$, cf. Fig. 2, and so does not pass the $\langle Subj(e) = m \rangle$ test. That movement can lead to perception by the protagonist (e_3, s_{15}) and controlled action, e.g. sitting up to avoid $sheet_1$, which is represented by the causal chain of e_{10} through e_{18}, from forming the intention to prevent collision to the change of location of the man's head. In these events the man has relatively high proto-agent status so they should be selected for narration as relevant answers to the *What does m do?* question. But according to the causality principle (6), if these events are selected their initial cause s_7, the movement of $sheet_1$, should be selected as well, which is in conflict with the result of applying the $\langle Subj(e) = m \rangle$ filter.

How is this conflict resolved? One possibility is that it depends on the look-

ahead capacity of the speaker. Assuming that the knowledge base is processed eventuality by eventuality in the chronological order,[6] the speaker first activates the cause s_7, so if her look-ahead window is small she will not simultaneously consider the consequences that qualify for narration. Thus the cause gets filtered out, so by the time the speaker activates the consequences, principle (6) forces her to skip them as well. The data suggest that speakers employ very small if any look-ahead, since the German speakers almost never mention the episode with $sheet_1$ and similar episodes.

In contrast, for a typical English speaker event selection is prior to and independent of the choice of subject, cf. (8). The speaker has an option to choose the temporal frame covering the episode in Fig. 2 for their witness time t_{now}, and if they do so, the causally connected events within that frame are selected for narration. This is why the English speakers, such as (1), have less difficulty in recounting this and similar episodes and do so significantly more often than the Germans.

However, there is one important exception to this generalization. If the protagonist is properly affected by some external cause, e.g. if he is actually knocked down by a flying sheet of paper rather than avoiding it, the episode has overall higher chances of being narrated and there is no significant difference between the English and the German speakers in the selection frequency. This is unexpected given the considerations presented above, because the man is an even less prototypical agent in such cases and shows more proto-patient properties. For example, in the episode where he is knocked down by the sheet of paper, he does not perceive the danger coming, he does not perform any volitional action, his movement (falling) is caused by the movement of the paper. Thus, the global subject selection in the German decision hierarchy (7) should result in even lower selection chances for the events in this episode and an even more drastic difference to the frequency among the English narrations.

This discrepancy is probably to be blamed on the universal narrative planning principles, which have precedence over the language-specific ones in our view. Since the situations where the protagonist is properly affected by an external force more severely jeopardize the fulfillment of his ultimate goal to find water, they receive a higher prominence in the story structure, though not as high as the "search for water" sequence itself. Therefore speakers that implement a higher level of detail in their retellings will select such episodes for narration regardless of their language-specific preferences.[7] This is why we find a high degree of individual variation in the selection of this episode (related to the individual level of detail),

---

[6] The relevant chronological order is formed primarily by events in the narrow sense (changes of state). States, including dynamic processes extended in time, are linked to this chronology by association with events that they cause or whose a pre- or poststate they constitute.

[7] This does not mean that the German speakers will not have difficulties once they face the necessity of building those events into their narrative. However, since they cannot step back from mentioning them anymore, they have to use other techniques to extricate themselves from the difficult situation and end up producing a coherent discourse. Some of these techniques are described in the next sections.

$$m, spine, legs, sheets, sheet_1, sheet_3, gro,$$
$$t_1...t_{11}, t^{1:20}...t^{2:27}, s_1...s_{240}, e_1...e_{128}$$

$$t_{50} < t_{110} < t^{1:20} < t^{2:27} \leq t_{51} \leq t_{111}$$

$$s_3 \Rightarrow s_{240}$$

$$s_7 \subset s_{240}; \quad s_{208} \subset s_{240};$$

$$s_3 : (wind, \langle t_{50}, t_{51} \rangle)$$

$$s_{240} : (move(sheets), \langle t_{110}, t_{111} \rangle)$$

$$s_7 : (move(sheet_1), \langle t_2^{1:31}, t^{1:34} \rangle);$$

$$s_{208} : (move(sheet_3), \langle t^{1:54}, t_1^{1:55} \rangle)$$

$$t_2^{1:31} \leq t_{14} \leq t_{18} \leq t^{1:34}$$

$$s_3 \Rightarrow s_7 \Rightarrow e_3 \Rightarrow s_{15} \Rightarrow ... \Rightarrow e_{10} \Rightarrow ... \Rightarrow e_{12} \Rightarrow ... \Rightarrow e_{17} \Rightarrow ... \Rightarrow s_{57}$$

$$e_3 : (s_{14}, s_{15}) : t_2^{1:31}; \quad \text{BECOME}(see(m, s_7))$$

$$e_{10} : (s_{36}, s_{37}) : t_{14} :$$
$$\text{BECOME}(intend(m, \text{CAUSE}(m, \neg \text{APPLYFORCE}(sheet_1, m))))$$

$$e_{12} : (s_{40}, s_{41}) : t_{18} :$$
$$\text{BECOME}(intend(m, \text{CAUSE}(m, \text{ALIGN}(spine, \text{VERT}))))$$

$$e_{17} : (s_{50}, s_{51}) : t^{1:34} : \quad \text{BECOME}(\text{ALIGN}(spine, \text{VERT}))$$

$$s_{57} : (\neg \text{APPLYFORCE}(sheet_1, m), \langle t^{1:34}, t^{2:27} \rangle)$$

$$s_3 \Rightarrow s_{208} \Rightarrow e_{120} \Rightarrow s_{215} \Rightarrow ... \Rightarrow s_{222} \Rightarrow$$
$$\Rightarrow [e_{126} \oplus e_{127} \oplus e_{128}] \Rightarrow [s_{225} \oplus s_{227} \oplus s_{229}] \Rightarrow e_{140} \Rightarrow s_{243};$$

$$e_{120} : (s_{214}, s_{215}) : t_1^{1:55} : \quad \text{BECOME}(\text{APPLYFORCE}(sheet_3, m))$$

$$s_{222} : ([m \text{ is falling}], \langle t_1^{1:55}, t_2^{1:55} \rangle)$$

$$e_{126} : (s_{224}, s_{225}) : t_2^{1:55}; \quad \text{BECOME}(\text{VERT}(m) \parallel \text{HOR})$$

$$e_{127} : (s_{226}, s_{227}) : t_2^{1:55}; \quad \text{BECOME}(\text{CONTACT}(spine, gro))$$

$$e_{128} : (s_{228}, s_{229}) : t_2^{1:55}; \quad \text{BECOME}(\text{CONTACT}(legs, gro))$$

$$e_{140} : (s_{242}, s_{243}) : t_2^{1:55} : \quad \text{BECOME}(\neg [m \text{ is properly functional}])$$

Figure 3: Episodes with $sheet_1$ and $sheet_3$ in the knowledge base

but the cross-linguistic difference does not reach statistical significance (Carroll et al., 2008).

What follows gives a brief illustration of how event selection is determined technically by applying the German (7) and the English (8) decision hierarchies and the universal principles discussed in section 4.1 to the knowledge base. Fig. 3 shows a (severely abridged) fragment of the knowledge base comprising states and events that play a role in the causal chains associated with the $sheet_1$-episode (see also Fig. 2) and the $sheet_3$-episode, where the protagonist gets knocked down by the flying sheet of paper. Examples (9) and (10) present an English and a German solution typical in the sense that both of them narrate the episode with $sheet_3$, cf. clauses (9n)–(9o) and (10k), but only the English speaker talks about $sheet_1$, cf. (9e)–(9h).

(9) a. second sequence opens with again a s / strange desert like landscape
     b. this time eh it seems to be made of sheets of paper [...]
     c. and then we see the figure falling onto the paper
     d. we can hear winds blowing in the background / a eh / wind like sounds
     e. the / the figure ehm sits up
     f. as a piece of paper sort of rolls itself
     g. and hurls towards him
     h. ehm he managses / manages to avoid it
     i. and then stands up
     j. ehm this paper pieces of paper / sheets of paper sort of swirling about here there and everywhere
     k. and as he walks
     l. he realizes
     m. he's sinking into this paper like landscape
     n. ehm a piece of paper then suddenly hits him (square) on
     o. in sense of flat onto his back

To model event selection by these speakers let's assume that they implement the same level of detail which is high enough to select the globally more prominent $sheet_3$-episode (the third block of clauses in Fig. 3), but not high enough for selecting the $sheet_1$-episode (the second block of clauses in Fig. 3). Thus the process s_{208} ($sheet_3$ flying) and the chain of events from e_{120} ($sheet_3$ hitting m) to e_{140} (m losing his normal functionality) as well as their connecting result states are "preselected" by the universal narrative planning principles on the basis of their prominence in the hierarchical story structure and causality. The resulting DRS for both speakers will contain conditions of the form select(e_{120}) applying to each of these eventualities.[8] Presumably, the selection of s_3 (wind blowing) and s_{240} (unindividuated sheets flying) is similarly triggered by the selection of the $sheet_3$ episode via the causality principle (6): s_3 causes s_{240} and consequently also s_{208} as its mereological part. So select(s_3) and select(s_{240}) are also added to the resulting DRS, ultimately surfacing in the clauses (9d), (9j) and (10e)–(10f).

The language-specific decision hierarchies are then applied to eventualities whose selection is not yet decided: s_7, e_3, $e_{10}-e_{17}$ and s_{57}. The German decision hierarchy (7) skips over eventualities where the protagonist m cannot be made subject, thus s_7 ($sheet_1$ rolling) is not selected since m does not participate in it, i.e. the condition ¬select(s_7) is added. Subsequent events are first submitted to the higher ranked universal principles and are filtered out by principle (6) as consequences of the deselected state s_7, so the language-specific preferences do not

---

[8]"Preselected" is only intended to mean that the universal principles at work here are ranked higher than the language-specific ones. In terms of the on-line production process the selection of these events might just as well happen as the speaker comes to them in the course of the chronological processing of the story line. Of course, this presupposes that the prominence values of the events (not represented here) are assigned globally on the basis of the overall structure of the story.

apply and the relatively high proto-agent status of the protagonist in those events (cf. above) does not play a role. So the whole episode is skipped by the German speaker. The English decision hierarchy (8) chooses freely among these events. The speaker of (9) must have selected all the eventualities upto s_{57}—the result state of successfully avoiding the collision, which appears in (9h)—otherwise s_{57} would have to be deselected as well.

4.3 Foregrounding and backgrounding

The eventualities selected for narration can be put in the *foreground* or in the *background*. In accordance with the common terminology, foreground events constitute the main line of the narrative, whereas the background is the related commentary. Prototypically, the foreground of the narrative is formed by chronologically ordered *events* in the narrow sense, i.e. punctual changes of state, which is expressed by the use of completive, perfective or punctual aspect and lexical material with intrinsic punctual *Aktionsart* (achievements and accomplishments, cf. Vendler, 1967), whereas the background is constituted by clauses describing states and processes extended in time (durative, progressive, imperfective, iterative) or evaluative comments that are not located in the story time (Tomlin, 1987). However, another linguistic dimension that is used to encode the foreground/background distinction is the choice between putting an event in a main clause or a subordinate construction of some sort (e.g. subordinate finite clause, infinitival or participial construction, nominalization, etc.). The aspectual and the subordination dimension may correlate, but they are *a priori* independent, which means that states and processes can be foregrounded by appearing in a main clause, while proper events can be backgrounded by subordination (Chvany, 1985). In other words, the distribution of eventualities between foreground and background is not given by their aspectual class, but is a matter of presentation by the narrator.

Since the sentence scheme for the German speakers is set globally by the $\langle Subj(e) = m \rangle$ constraint in (7), i.e. in constructing their narrative German speakers are basically answering the question *What did the protagonist do?*, events where the protagonist is agent or experiencer normally surface as main clauses. Since these events also constitute the story line for the most part, the German narratives overwhelmingly show the prototypical mapping of story line events to main clauses, as in (10j). It should be noted that story line events, such as *das männchen fallen* in (10a), *und dann weich landend in / auf ner fläche* in (10b) and *als würde er sich wundern* in (10h) presented as being perceived by the narrator (cf. *man sieht* in (10a), *sieht so aus* in (10g)), will be treated as foregrounded for our present purposes despite the fact that they systematically appear in syntactically subordinate constructions both in German and English narratives. The narrator's perspective on the film events constitutes a separate layer which is not part of the film as such. Since it does not belong to the story-line, it is neither part of the foreground by nature, nor does the speaker have any reason to make it foreground as long as his task is to report the events of the film. The $\emptyset$-subject in (10j) which is strictly speak-

(10) a. als   nächstes   sieht   man   das   männchen   fallen [...]
        as   next   sees   one   the   man   fall

     'next one sees the man fall'

   b. und   dann   weich   landend   in /   auf   ner   fläche
       and   then   softly   land   in   on   a   surface

     'and softly land on a surface'

   c. die   mit   papierstücken   so   belegt   is
       that   with   pieces of paper   so   covered   is

     'that is covered with pieces of paper'

   d. es   ist   also   ne   ne   ganze   weite   landschaft
       it   is   thus   a   a   whole   wide   landscape

     'so this is a wide area'

   e. wo   es   windig   ist
       where   it   windy   is

     'where it is windy'

   f. und   einzelne   papierblätter   durch   die   luft   fliegen
       and   single   pieces of paper   through   the   air   fly

     'and some pieces of paper are flying in the air'

   g. und   sieht   so   aus
       and   looks   so   out

     'and it looks'

   h. als   würde   er   sich   wundern
       as   would   he   himself   wonder

     'as if he wondered'

   i. wie   es   hier   aussieht
       how   it   here   looks

     'how it looks here'

   j. und   ∅   geht   n'   paar   schritte
       and   goes   a   couple   steps

     'and [he] makes a few steps'

   k. ∅   wird   dann   von   so   nem   wehenden   blatt   umgeworfen
           gets   then   by   such   a   flying   sheet   knocked over

     'then [he] gets knocked over by such a flying sheet'

ing ungrammatical in this position, but would be grammatical if the clause (10h) containing its antecedent *er* were a main clause, could be seen evidence for (10h) being part of the foreground.

In contrast, if an eventuality (event or state) is selected by higher ranked universal principles despite the fact that m is absent or has a low proto-agent status in it, one of the possible ways for a German speaker to stay in accordance with the $\langle Subj(e) = m \rangle$ constraint at the global level, is to put that eventuality in the background.

(11) $[\text{select}(e) \wedge Subj(e) \neq m] \Longrightarrow \text{background}(e)$

Since German does not have grammatical aspect, the preferred backgrounding strategy would be to use syntactic subordination.[9] Presumably, this is why s_3 (wind blowing) and s_{240} (papers flying about) surface as relative clauses (10e) and (10f).

Among the English speakers foregrounded "background" descriptions and backgrounded story line events are more common than among the Germans. For example, the description of s_{240} in the progressive is not syntactically subordinated in (9j), whereas the protagonist's action *he walks* expressed by simple present appears in a subordinate clause in (9k). Similarly, the description of s_3 *winds blowing* is directly embedded in a matrix clause introducing the narrators perspective *we can hear* in (9d) and is thus part of the foreground given our present assumptions. This suggests that the speaker of (9) is assessing the narration task, roughly speaking, by answering a question like *What do we/you see happening in the film?*, i.e. he introduces the narrator's perspective, cf. (8), and chooses globally for the progressive as the dominant aspectual form. This creates a bias for state descriptions to appear in the foreground, whereas proper story line events which do not comply to the *What is happening?* schema can be put in the background by subordination. As was pointed out before, a global choice of aspectual form has an advantage for the English speakers since aspect marking is obligatory, so if aspect is preset globally, the speaker has less decisions to make on a clause by clause basis.

(12) $[\text{select}(e) \wedge \text{durative}(e) \wedge \text{select}(s_p) \wedge s_p : (perceive(n, e), T_{now})] \Longrightarrow$
     $\text{foreground}(e)$

(12) only says that states, or durative eventualities (and therefore progressive clauses) will be foregrounded, it leaves open whether proper events are backgrounded by subordination. In fact, another way to deal with punctual events that constitute the story line is to introduce a complex eventuality that sums the selected events and thus extends in time from the point of the first till the point of the last event. Obviously, this is the option taken for presenting the events of the $sheet_1$-episode in (9e)–(9h), cf. (13). Since s is a state that extends in time it qualifies for appearing in the foreground in a sequence of main clauses.

---

[9]If the lexicon happens to make a stative expression available for the given concept, that expression can also be used as a backgrounding means. But since there is no *grammatical* aspect, i.e. no productive way of encoding stativity or durativity for any arbitrary concept, one cannot rely on intrinsic aspect strategically in German.

(13) $s : \left(e_3 \oplus s_{15} \oplus e_{10} \oplus s_{37} \oplus ... \oplus e_{12} \oplus ... \oplus e_{17}, \langle t_2^{1:31}, t^{1:34} \rangle\right)$

If an English speaker decides not to introduce the narrator's perspective, which is for instance the case in (14), cf. also (5) in section 4.1, the default solution is to foreground punctual events that propulse the story (15). They will normally surface in a sequence of main clauses in simple present:

(14) a. okay the man arrives in a paper world
     b. and eh everywhere is covered in paper
     c. and he gets hit by a flying piece of paper
     d. and then he walk/ he hustles around
     e. and walks about
     f. and he finds a damp piece of paper
     g. and he pushes the paper
     h. and then he falls through the paper

(15) $[\text{select}(e) \wedge \text{punctual}(e)] \implies \text{foreground}(e)$

The resulting narrative looks very much like a typical German narrative in the sense that story line events are mapped to the foreground clauses. This is a good candidate for a universal principle. However, unlike other universals discussed so far, it is a default and ranked lower than the competing language-specific choices and their consequences like (11) and (12). That is, (15) will take effect in English only when it is not overruled by (12) in combination with the principle that tells us to background all eventualities not compatible with the foreground quaestio (*What is happening?*, asking for durative eventualities), i.e. when the speaker chooses against introducing the narrators perspective. In German (15) will apply to all events with the protagonist as the most likely subject, because otherwise (11) takes precedence and overrules the effect of (15). Notice that the present hierarchy of rules does not imply anything with respect to foregrounding or backgrounding state descriptions like (14b) in English. Neither (15) nor (12) apply in this case (since the narrator's perspective is absent here), so the speaker is free to realize this state description in a main or a subordinate clause, choosing the latter in (14b). In contrast, a German speaker would go for backgrounding and subordination since the state does not involve the protagonist and rule (11) applies.

Finally, it should be noted that German speakers also have an option of introducing the narrator's perspective, as the speaker in (10) does, cf. (10a) and (10g). However, it appears relatively late in the decision hierarchy and does not have any strategic impact (for this reason it is skipped in (7)), in contrast to English where it determines the global choice of the foreground aspectual form.

4.4 Subject selection

The choice of grammatical subject is largely a matter of perspective taking and belongs to the level of microstructural planning. It has very different status in the language-specific decision hierarchies (7) vs. (8). In German, it is the first

decision to make, the value is set globally to the protagonist, which has a range of strategic consequences for narrative planning, in particular for event selection and foregrounding/backgrounding decisions as described in the sections above. In English, the choice of subject is late and potentially influenced by other decisions, and its value is not generally preset. Therefore let's first consider subject selection in English and then briefly return to German.

The data suggest that the subject choice in English depends primarily on the local context and is governed by well-known universal principles, such as (the most prototypical) *agent is subject* (Dowty, 1991) and *topic is subject* (Keenan, 1976), while topichood plays the decisive role, cf. (16). The notion of topic relevant here is the aboutness topic, i.e. an individual referred to again and again over a stretch of discourse. As a consequence, there is a preference for the subject to corefer with some participant, if possible the subject, of the immediately preceding clause, cf. (17). (c_n refers to the current clause, c_{n-1} to the immediately preceding clause.)

(16) $\text{Subj}(c_n) = \text{Topic}(c_n)$

(17) $\text{Topic}(c_n) = \text{Subj}(c_{n-1})$

This generalization also follows from theories of pronoun resolution and pronominalization such as Centering (Brennan et al., 1987; Beaver, 2004). It describes a vast majority of subject choices in the English narratives. For example, if the event of $sheet_3$ hitting the man follows a sequence of clauses describing the protagonist's actions (18d)–(18f), the protagonist remains the subject despite its strong proto-patient status in this event and the verb is passivized (18g).

(18) a. and in the background you see paper flying around like a whirlwind
     b. and it's very windy
     c. and is pieces of paper blowing all over the place
     d. eh so he stands up
     e. and starts look around
     f. eh manages to get out of the way of one piece of paper
     g. and then he's hit by another one
     h. that knocks him down

In contrast, if the same event is embedded in a description of the movements of the sheets of paper, e.g. (19e)–(19g), normally the proto-agent $sheet_3$ becomes the subject and the verb keeps the active voice, cf. (19h).[10]

---

[10]It rarely happens that $sheet_3$ is mentioned as an individual in more than in one clause, so there is usually no strict coreference between that one mention and its antecedent that is topic in the preceding context and licenses $sheet_3$ to appear in the subject position. Typically the antecedent is the set of all papers flying around, cf. *they* in (19f)–(19g), so $sheet_3$'s relation to it is member-set (or mereological part). In most cases, the protagonist is not mentioned in the immediately preceding clause, so $sheet_3$ and m do not compete for topichood. However, it is interesting to note that in (19g) the protagonist is mentioned, but in an oblique position (*him*), so apparently, a weaker relation to the previous subject is more essential for the chances to appear as subject in the current clause than strict coreference with an oblique antecedent.

(19) a. so he gets up
   b. and / and goes towards the sound
   c. and as he does this
   d. you see like
   e. it's a paper flying past him
   f. and they're quite big
   g. they're like the size of him
   h. so one knocks him over

Apparently, global decisions such as the choice to introduce (or not) the narrator's perspective do not play a role in subject selection. Both (18) and (19) contain the narrator's perspective, cf. *you see* in (18a) and (19d), and in both cases the subject of the 'hit/knock-down' event is the topic set up by the local context. Speakers who do not introduce the narrator seem to follow the same principles, cf. (20) and (21).

(20) a. and he hears the dripping again
   b. and he gets knocked down by a piece of paper

(21) a. eh he falls through in / into the next world
   b. which is a / a paper world
   c. and the paper is all flying around him
   d. and knocks him down

It is important to note that what constitutes the relevant local context depends on the main vs. subordinate clause status, and more generally on the foreground vs. background distinction. Roughly, it is only the topics 'on the same plane' that count, i.e. if, for instance, we are choosing a subject for a foreground clause c_n, the relevant c_{n-1} is the last foreground clause, and not literally the last clause if there happen to be intervening background clauses. The clearest case are intervening syntactically subordinated clauses. It is a well-established fact that subjects of subordinate clauses do not have the same power to promote their referent to a topic, as subjects of main clauses (Suri and McCoy, 1994; Cooreman and Sanford, 1996; Miltsakaki, 2003). Thus the choice of $sheet_3$ for subject in (22e) continues the paper topic established in (22b)–(22c), and the fact that the protagonist is subject in (22d) does not matter since this is a subordinate clause.

(22) a. and it's windy
   b. and those papers are blowing
   c. there is there is a little sort of tornado of papers going by him
   d. as he's walking
   e. and a piece of paper flies in his face

In other cases, main clauses seem to be skipped over. E.g. in (14), if (14b) were counted as a foreground clause on a par with (14c), the protagonist would not qualify for topichood in (14c), and hence neither for pronominalization nor for the

(23) a. dann kommt noch so'n   papierhurrikan vorbei
        then   comes   yet   such a paper tornado   past

    'then a paper tornado passes by'

   b. hätt   ihn   auch   fast   erwischt
      would have   him   also   almost   hit

      'almost hits him too'

   c. und   wie   er   nicht   aufpaßt
      and   as   he   not   pays attention

      'and while he is not paying attention'

   d. kommt so   eine   riesen   zeitung
      comes such a huge   newspaper

      'a huge newspaper comes'

   e. und   legt   ihn   um
      and   lays   him   over

      'and knocks him over'

subject position, since he is not mentioned in (14b). This suggests that (14b) is backgrounded (as a stative description) without being syntactically subordinated. Presumably, this disregard for background clauses is one of the consequences of a more general pattern in reference assignment, which is captured by roughly the same principles in a number of theories such as the stack theory (Grosz and Sidner, 1986; Walker, 1996), the veins theory (e.g. Cristea et al., 1998), or the right frontier constraint (Polanyi, 1988; Asher and Lascarides, 2003). The central notion is that of *discourse-structural subordination*, which need not correlate with syntactic subordination, but otherwise has more or less the same consequences for reference assignment as those associated with syntactic subordination, cf. above. The notion of background in narrative studies roughly corresponds to a number of special cases of subordinate discourse material.

   The standard assumption is that the principles regulating subject selection described above are not specific for English and should be at work in German as well. How does this combine with the global $\langle Subj(e) = m \rangle$ choice in the German decision hierarchy (7)? First of all, the $\langle Subj(e) = m \rangle$ constraint has nothing to say about clauses that present eventualities where m does not participate, so (as a null hypothesis) the general principles should apply there. We will not discuss such cases. But what if $\langle Subj(e) = m \rangle$ on the one hand and topichood and proto-agent status on the other suggest distinct referents for the subject? Interestingly, since the $\langle Subj(e) = m \rangle$ applies early on in the decision process and influences other planning decisions at all levels such conflict almost never occurs. First, the speakers do their best in skipping the eventualities where the protagonist is not a proper

(24) a. ja   der   zweite teil is   ähm ne   art   papierwüste
        yes   the   second part is   ehm a   sort   paper desert

        'yes, the second part is a sort of a paper desert'

    b. wo   also   überall /
       where   thus   everywhere

       'where like everywhere...'

    c. der   boden   besteht   aus papierblättern
       the   ground consists of   sheets of paper

       'the ground consists of sheets of paper'

    d. und   wie   sandsturm   fliegen auch papierblätter   durch   die
       and   like sand storm   fly   also sheets of paper   around   the
       gegend
       place

       'and sheets of paper are flying around like in a sand storm'

    e. einmal   auch   mit   so   'ner   art   wirbelsturm
       one time   also   with such a   sort   tornado

       'at one point there is also a sort of a tornado'

    f. und   irgendwann   kriegt   er   auch   so'n   papierblatt   ins
       and   at some point   gets   he   also   such a   sheet of paper   in the
       gesicht
       face

       'and at some point he also gets such a piece of paper blown into his face'

    g. was   ihn   umhaut
       which   him   knocks over

       'which knocks him over'

25

agent. Then they do their best in backgrounding such events if they failed to skip them. The very notion of 'protagonist' implies that it is a global aboutness topic, the individual recurring again and again. As a result, the protagonist is both the agent and the topic in the absolute majority of foreground clauses.

If the protagonist happens to be the patient, as in the episode with $sheet_3$, it is still normally the topic, and since topichood overweighs agentivity, the protagonist normally becomes the subject, while the verb is passivized, as in (10k) in the context of the foreground clauses (10h) and (10j).

Among the very few cases where the $sheet_3$ episode follows a sequence that establishes the sheets of paper as topic we find both instances of $sheet_3$ and of the protagonist as subject, cf. (23e) and (24f). The former suggests that the general principles that select topics for subjects override the language-specific $\langle Subj(e) = m \rangle$ preference. The latter can be interpreted as an indication to the opposite. Alternatively, they can be seen as pop returns from the background to the foreground, i.e. the protagonist becomes subject because it is a foreground clause that continues a sequence of foreground clauses about the protagonist that has been interrupted by the immediately preceding background sequence about the paper.[11] This alternative is consistent with the ranking of the language-independent subject assignment principles over the language-specific ones. Thus the global $\langle Subj(e) = m \rangle$ constraint plays its role in directing the narrative structure as to minimize the disalignment of the protagonist, topic and subject, but if such disalignment takes place, the same reference assignment principles as in English operate on the basis of the local context.

4.5 Predicate selection

Finally, the events selected for narration undergo *packaging* through the choice of lexical concepts, in particular the lexical concept for the main predicate of the preverbal message. For instance, the event e_1 in Fig. 2 can be mapped to *the man's/his eyes open* as shown in (25), where e.g. the predicate $open_1$ corresponds to the lexical concept behind the English intransitive verb *open* or the German verb *aufgehen* (as in *seine Augen gehen auf*).

$$
(25) \quad
\boxed{\begin{array}{l}
e_1, s_1, s_2, t_1^{1:31}, eyes \\
\hline
\text{select}(e_1) \\
e_1 : (s_1, s_2) : t_1^{1:31} : \\
\quad \text{BECOME}(open(eyes)) \\
t_1^{1:31} \subset T_{now}
\end{array}}
\implies
\boxed{\begin{array}{l}
x, y, e, t \\
\hline
man(x), eyes(y) \\
possess(x, y) \\
e : open_1(y) \\
e \subset t, t \subset T_{now}
\end{array}}
$$

---

[11]In the case of (24) this solution presupposes that (a) background eventualities in German do not necessarily have to surface in subordinate clauses (e.g. clauses (24a), (24c) and (24d) that establish the sheets of paper as topic are main clauses); and (b) the retelling of the paper scene in (24) continues the retelling of the preceding scene which *together* are assigned a single discourse structure. Otherwise there is no preceding foreground clause with the protagonist as subject which could serve as antecedent for (24f) after a pop return. This assumption makes sense also because without it the pronoun *er* 'he' in (24f) does not have an antecedent.

Adding $e_{151} : (s_{263}, s_{264}) : t_{130} : \textsc{Intend}(m, \textsc{Cause}(m, open(eyes)))$ and a causal connection between e_{151} and e_1 to the antecedent of (25) licenses the choice of the transitive concept $e : open_2(x, y)$ of the English *open* or the German *öffnen*. Obviously, the global preference for the protagonist as subject in the German decision hierarchy puts pressure on packaging the events the latter way, whereas the intransitive solution in (25) is relatively more likely in English. This often leads to more coarse-grained packaging solutions in German (*er macht ein Loch im Boden* 'he makes a hole in the ground') vs. more fine-grained in English (*he starts digging and a hole appears*).

Here are some more examples of mapping the knowledge base conditions to lexical concepts. The combination of state descriptions s_6, s_7 ($sheet_1$ moves), s_8 (the path of $sheet_1$'s movement is w_2) and s_{16} (m's head is located on that path) licenses the choice of *move/fly/roll towards the man / auf das Männchen zufliegen* as predicate (we ignore the mode of motion) applied to *the sheet of paper / das Papier*. Since the beginning and the end of this movement are not shown in the film and are unspecified in the corresponding DRS, we only find definite temporal locations that are embedded in the maximal movement state s_7, e.g. $\langle t_2^{1:31}, t^{1:34} \rangle \subset s_7$. This licenses the use of the progressive form in English. The event e_3 marks the point when the man starts to see $sheet_1$ moving. Since neither German nor English have a concise way of referring to this event, such as an inchoative form of *see / sehen*, e_3 is verbalized via reference to its poststate *see the sheet of paper flying towards him*. The sheet of paper $sheet_1$ never hits the man in the whole scene ($s_{56} : (\neg\textsc{ApplyForce}(sheet_1, m), \langle t^{1:25}, t^{2:27} \rangle)$). States that never change, and especially those identified by negative propositions, are normally not narrated. However, if there is an expectation or belief that the proposition should have obtained at some point, then negated predicates can be produced. The protagonist's belief that $sheet_1$ might hit him (s_{23}) in the time interval $\langle t_{20}, t_{21} \rangle$ and the fact that it doesn't (s_{57}) licenses lexical choices that lead to the production of sentences like *The sheet of paper does not hit him* or *The sheet of paper misses him*, cf. (1h). Moreover, s_{57} is brought about by m's action originating from his intention to do something so that $sheet_1$ does not hit him (e_{10}). Thus $e_{10} \Rightarrow ... \Rightarrow s_{57}$ licenses the use of such predicates as *avoid the sheet of paper*. Finally, the fact that the intention to prevent collision with $sheet_1$ (e_{10}) gives rise to another intention e_{12} to sit up, i.e. align the spine with the absolute vertical while the legs remain resting on the ground (s_{55}), which is carried out in e_{17}, licenses the production of sentences like *he sits/gets/jumps up (in order) to avoid the sheet of paper*. In sum, these examples are supposed to show how the DRT-based representation of the preverbal knowledge passing through the filter of the event selection procedure gives rise to a variety of expressions for the same content in one language or two different languages.

5 Conclusion

The work presented in this paper is novel insofar it combines methods of psycholinguistic research with methods of formal discourse theory. The perspective on discourse, i.e. on narration production, is also new. The psycholinguistic research objective is to reconstruct the striking differences in the *what* and in the *how* of narration solutions of the same film by speakers of different languages in a formal model. That reconstruction is made operative by determining a ranked set of selection decisions and other operations that must apply to a language-neutral knowledge base as to determine a language dependent 'preverbal message'.

For the first time a methodological tool has been established allowing not only to compare data in one language to that in another but to view them as the result of preparatory operations on that shared basic representation and respective differences therein. Relying on earlier work we make those operations explicit in a grammar-dependent hierarchy of decisions involving event, subject and predicate selection. That hierarchy had allowed us to explain the striking differences in the solutions.

Starting the reconstruction from the language-neutral knowledge base, other constraints on the narration solutions came into focus too, in particular universal rules of discourse coherence, familiar from formal discourse analysis. Taking also these rules into account as operative in the process of preparing the 'preverbal message' allowed a pervasive reconstruction of the data.

Universal principles, such as the coherence rule that causal consequent events cannot be narrated without their antecedents being narrated as well, or the principle of topic-continuation in the selection of subjects are primary constraints that determine the possible solutions already to some extent. But with the knowledge base at hand, the range of decisions and operations that still *must* be taken and so *can* be taken to make narration an easy task in the respective language cannot be overlooked either. So our psycholinguistic hypotheses were finally confirmed in a more subtle view on the data.

Some of these decisions, for instance, foregrounding and backgrounding operations in negotiation between universal constraints on the one hand and language dependent 'strategies' on the other have been formalized in rules. For other operations, e.g. *packaging* in predicate selection, we have precise ideas of how to specify the DRT-based representations to determine lexical selection. Much of giving the 'preverbal message' formal substance as DRT-based structures is still program, but it is a promising one.

References

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Studies in Natural Language Processing. Cambridge University Press.

Beaver, D. and Zeevat, H. (2007). Accommodation. In Ramchand, G. and Reiss,

C., editors, *Oxford Handbook of Linguistic Interfaces*, pages 503–540. Oxford University Press.

Beaver, D. I. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy*, 27(1):3–56.

Brennan, S., Friedman, M., and Pollard, C. (1987). A centering approach to pronouns. In *Proceedings of the 25th annual meeting on Association for Computational Linguistics*, pages 155–162.

Carroll, M., Roßdeutscher, A., Lambert, M., and von Stutterheim, C. (2008). Subordination in narratives and macrostructural planning: A comparative point of view. In Fabricius-Hansen, C. and Ramm, W., editors, *'Subordination' versus 'Coordination' in Sentence and Text: A cross-linguistic perspective*, pages 161–184, Amsterdam. Benjamins.

Chvany, C. V. (1985). Backgrounded perfectives and plot-line imperfectives: Towards a theory of grounding in text. In Flier, M. S. and Timberlake, A., editors, *The Scope of Slavic Aspect*, pages 247–273. Slavica Publishers, Inc., Columbus, Ohio.

Cooreman, A. and Sanford, A. J. (1996). Focus and syntactic subordination in discourse. Research paper 79, HCRC, Edinburgh, Scotland, U.K.

Cristea, D., Ide, N., and Romary, L. (1998). Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of COLING-ACL 1998*, pages 281–285.

Dowty, D. R. (1991). Thematic proto-roles and argument selection. *Language*, 67:547–619.

Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions and the structure of discourse. *Computational Linguistics*, 12(3):175–204.

Jasinskaja, K. and Roßdeutscher, A. (2008). Towards generating narratives from a preverbal message: A DRT-based approach. In Benz, A., Kühnlein, P., and Stede, M., editors, *Constraints in Discourse 3. Proceedings of the Workshop*, pages 45–52, Potsdam, Germany. University of Potsdam.

Kamp, H. and Reyle, U. (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Kluwer Academic Publishers.

Kamp, H. and Roßdeutscher, A. (2005). A part of the conceptual background theory, concerned with time, space and motion. Technical report, IMS, Stuttgart University.

Keenan, E. (1976). Towards a universal definition of 'subject'. In Li, C. N., editor, *Subject and Topic*, pages 303–333. Academic Press, New York.

Lehnert, W. (1981). Plot units and narrative summarization. *Cognitive Science*, 5(4):293–331.

Levelt, W. J. (1999). Producing spoken language: a blueprint of the speaker. In Brown, Collin, M. and Hagoort, P., editors, *The Neurocognition of Language*, pages 83–121. Oxford University Press.

Mann, W. C. and Thompson, S. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. *Text*, 8(3):243–281.

Marcu, D. (1997). From discourse structures to text summaries. In *Proceedings of the ACL Workshop on Intelligent Scalable Text Summarization*, pages 82–88.

Miltsakaki, E. (2003). *The Syntax-Discourse Interface: Effects of the Main-Subordinate Distinction on Attention Structure*. PhD thesis, University of Pennsylvania.

Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of Pragmatics*, 12:601–638.

Rumelhart, D. E. (1975). Notes on a schema for stories. In Bobrow, D. G. and Collins, A., editors, *Representation and Understanding*, pages 211–236. Academic Press, New York.

Rumelhart, D. E. (1977). Understanding and summarizing brief stories. In LaBerge, D. and Samuels, S. J., editors, *Basic processes in reading: Perception and comprehension*, pages 265–303. Lawrence Erlbaum Associates.

Slobin, D. I. (1987). Thinking for speaking. In *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, volume 13, pages 435–444.

Suri, L. Z. and McCoy, K. F. (1994). RAFT/RAPR and Centering: A comparison and discussion of problems related to processing complex sentences. *Computational Linguistics*, 20(2):301–317.

Tomlin, R., editor (1987). *Coherence and Grounding in Discourse*. John Benjamins, Amsterdam.

Vendler, Z. (1967). *Linguistics in Philosophy*. Cornell University Press, Ithaca, New York.

Walker, M. A. (1996). Limited attention and discourse structure. *Computational Linguistics*, 22(2):255–264.

Wilensky, R. (1983). Story grammars versus story points. *Behavioral and Brain Sciences*, 6(4):579–623.