

Guidelines for the Annotation of Parameters of Narration

Version 1.0

Nico Lehmann^a, Dina Serova^a, Julia Lukassek^a, Sophia Döring^a, Frank Goymann^a, Anke Lüdeling^a & Roodabeh Akbari^a

^aHumboldt-Universität zu Berlin

The present guidelines describe the annotation of narrative phenomena on the clause level, using a combination of ideas and methods from linguistics and literary studies. The main categories marking the discourse strategy “narration” in stretches of text have been narrowed down to mediacy, i. e. involving a narrator, and sequentiality of events. This document specifies how to define mediacy, and in turn determine whether a narrator is present, as well as how to identify events and their sequential ordering. Lastly, a functional layer annotation is proposed which allows researchers to compare different types of narrative instances. This offers a basis for investigating a potential narrative register which is said to be important for many kinds of register studies.

Keywords: narration, register, annotation, corpora, guidelines

1 Introduction

These guidelines describe the multi-layer annotation of different aspects of narration developed and tested in the CRC 1412 “Register”¹. Narration – loosely described for our purposes as the act of representing events or telling stories – is found everywhere in human communication and has also played a large role in publications on register (cf. Mehler 2008; Neumann 2014; Biber & Conrad 2019; Egbert & Mahlberg 2020; Biber et al. 2021). We define register as *intra*-individual variation, that is, the conventionalized and recurrent linguistic patterns of language use in a speech community (cf. Lüdeling et al. 2022). Contrary

¹<https://sfb1412.hu-berlin.de>



to the common practice in register and corpus research where *complete* texts are simply labeled “narrative” (in contrast to, for example, argumentative; see for an overview Lee 2002), we agree with linguistic and literary theory that narration is a conflated category that really encompasses a set of aspects. We aim to answer questions about common traits of narrative passages across different contexts, times, modalities and languages. This has two consequences: (a) we view narration not as a category that necessarily pertains to a whole text – rather, a text can contain any number of narrative passages (and argumentative passages, expository passages etc.); (b) different aspects of a narration need to be considered separately, e. g. with respect to register analyses. We therefore set out to decompose the complex notion of narration in order to identify narrative passages and annotate different aspects.

1.1 Annotation overview

The following is a complete list of the annotation values (tags) and their respective tiers with examples as used in these guidelines. These represent the main parameters identified for the discourse strategy “narration”. There is a minimal and an extended version for the annotation. The minimal version includes the following main tiers:

- MEDIACY (3),
- NARRATOR (3.3),
- EVENT (4.1),
- SEQ (4.2),
- NAR.ID (5),
- four functional layers (6)
 - F.ENT,
 - F.FCT,
 - F.ILL,
 - F.SOC.

The (optional) extended version additionally adds the tier MOTIVE (3.2), in which indicators for mediacy may be specified, and tiers that detail characteristics of the narrator such as *Part* (3.3), in which can be determined whether a narrator is part of the narrated world.

Guidelines for the Annotation of Parameters of Narration

| Tag | Tier | Description | Example |
|-----------------|---------|---|---|
| mediate | MEDIACY | multi-layered structure, e. g. via indicators of mediacy | <i>While cooking dinner, they heard a strange noise.</i> |
| non-mediate | MEDIACY | lack of multi-layered structure | <i>I like playing the piano.</i> |
| FID | MOTIVE | free indirect discourse: merging of perspectives (narrator ↔ protagonist) | <i>Perhaps it was finally his big day today.</i> |
| FoF | MOTIVE | future of fate: events reported in the past for narrator, but future for protagonist (protagonist > event < narrator) | <i>She was going to find out soon.</i> |
| HP | MOTIVE | historical present: present tense used for past events (present tense event < narrator) | <i>I told him to read the book and so he reads it the very same evening.</i> |
| Past | MOTIVE | past tense: multi-layered structured via past tense (event < narrator) | <i>She grabbed a coffee and sat down in the park.</i> |
| direct speech | MOTIVE | mediacy via representing a third party's speech in an unfiltered way | <i>He said: "I don't like it."</i> |
| indirect speech | MOTIVE | mediacy via representing a third party's speech in a filtered way | <i>He said that he didn't like it.</i> |
| narrated speech | MOTIVE | mediacy via narrating a third party's speech | <i>She decided to try it.</i> |

| | | | |
|-------------------|----------|---|--|
| narrator-001 | NARRATOR | ID of the mediating authority whenever mediacy is determined; number increases when narrators change; numbering added for subordinate narrators | <i>Then he said: (narrator-001) "I had to go." (narrator-001-01)</i> |
| none | NARRATOR | no mediacy and therefore no narrator | <i>I like playing the piano.</i> |
| part | PART | narrator is part of the narrated world | <i>I went to him and we discussed it.</i> |
| non-part | PART | narrator is not part of the narrated world | <i>Once there was a princess who dreamed of the sea.</i> |
| e-001 | EVENT | first event in a text | <i>While cooking dinner, they heard a strange noise.</i> |
| e-002 | EVENT | second event in a text | <i>While cooking dinner, they heard a strange noise.</i> |
| e-001 c e-002 | SEQ | sequence of events: overlap of event 1 and event 2 | <i>While cooking dinner, they heard a strange noise.</i> |
| e-001 > e-002 | SEQ | sequence of events: succession; precedence of event 1 before event 2 | <i>After cooking dinner, they heard a strange noise.</i> |
| e-001 <> e-002 | SEQ | sequence of events: underspecified relation of event 1 and event 2 | <i>They cooked dinner and they heard a strange noise.</i> |
| nar-001 | NAR.ID | identifier for each narration, consecutively numbered | <i>A: I slept all day before coming here. (nar-001) B: I had ice cream before coming here. (nar-002)</i> |

| | | | |
|------------|-------|---|--|
| entertain | F.ENT | entertainment function of a narration | <i>A man enters a doctor's office and ...</i> |
| fact | F.FCT | factual transfer of information function of a narration | <i>I broke the leg while walking down the steps.</i> |
| illustrate | F.ILL | illustrative function of a narration | <i>They put the wrong screw in and that's why the whole thing broke apart.</i> |
| social | F.SOC | social management function of a narration | <i>I had such a wonderful day yesterday. First I slept late ...</i> |

1.2 Annotation procedure

This section presents a schematic overview of the annotation process proposed by these guidelines including a workflow in Figure 1 and an exemplary annotation in Table 2 that illustrates all the annotation layers on a schematic word and clause level segmentation with dummy annotation values.

Figure 1 displays a decision tree for the annotation procedure. Starting from the text level, annotators decide for each clause whether it is mediated (Section 3), i. e. whether it depicts a transmission or representation via a narrating authority. An optional layer (Motive) may be used to specify the reason for analysing a passage as mediate, i. e. an indicator of mediacy as delineated in Section 3.2.² Whenever a passage is determined as mediated, this points towards the agency of a narrator³. Each clause therefore receives a narrator ID, consecutively numbered (Section 3.3). This is particularly vital for differentiating separate narrations in the same text by different narrators. Further characteristics of the narrator may be annotated optionally. Mediate text passages then receive an event annotation

²Some corpora may have language-specific annotations resulting from semi-automatic procedures that can help to find such indicators, such as TAM categories or representations of speech.

³The “pan-narrator-hypothesis” has been opposed by some scholars (cf. Köppe & Stühling 2011; Eckardt 2015), but we believe that mediacy and in turn a mediating authority, irrespective of whether being overt or covert, is crucial to successfully *operationalize* parameters of narration – in particular when looking at the clause level. We therefore opt to include the mediating agent, here called “narrator”, in our annotation.

for each eventive token, consecutively numbered (Section 4.1). In the next step, the sequential relation between the events is annotated (Section 4.2). All passages that fulfil the criteria of mediacy and contain a sequence of events receive a narration ID (Section 5). Each identified narration is lastly annotated for its main function.

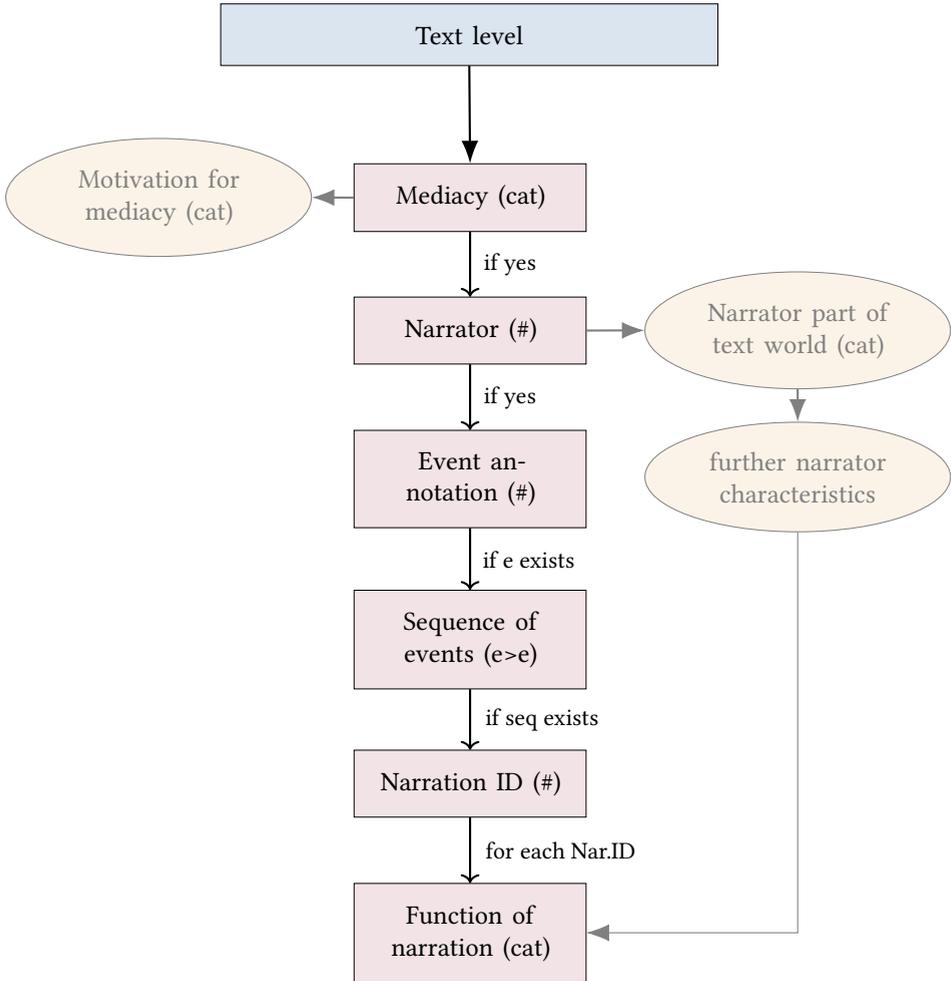


Figure 1: Annotation workflow: yellow ellipses are optional annotations, red square boxes move towards the next layer only when positive annotations exist. Abbreviations in brackets define annotation types: # = consecutively numbered categories, e = consecutively numbered events, cat = categorical variants.

This procedure results in an annotation as exemplified in Table 2 (see also ex-

ample annotations applied for different text types in the Appendix). The table displays the annotation for the example in (1), which is an excerpt from a conversation between friends where the narrator starts to tell a joke but interrupts their mediation to remark on the need to drink something before continuing.

- (1) Then entered a little girl, and she said: someone broke my new flute.
w1 w2 w3 w4 w5 w6 w7 w8 w9 w10 w11 w12 w13
[S1] [S2] [S3]
I need to drink something.
w14 w15 w16 w17 w18
[S4]

The indicators past tense and direct speech (Section 3.2), annotated in the tier MOTIVE, identify the clauses with mediacy, i. e. clauses S1-S3 in Table 2. S4 is not mediate. This is annotated in the tier MEDIACY. The person telling the story is designated as *narrator-001*. Even though we have no further information about the narrator from the excerpt, the fact that mediacy is involved tells the annotator that this is the first narrator here. When the girl starts speaking, she becomes the second narrator, but as she is a protagonist in narrator-001's story, she receives the ID *narrator-001-01*. On the tier PART, it is annotated that narrator-001 is not part of the narrated world, while narrator-001-01 is a part of that world. For each eventive verb, an event ID is provided in the tier EVENT. The resulting events are assigned one of three relations on the tier SEQ (Section 4), i. e. succession (<), overlap (c), and underspecified (<>). Here, e-001 precedes e-002 and e-002 precedes e-003 while e-003 may further be related to events in the continuation of the story. All narrative passages that belong to a single sequence of events are specified as narrations and receive a narration ID on the tier NAR.ID.⁴

The multi-layered structure may involve embedded narrations inside of narrations with different functions, which is why narrations embedded in another narration receive a subordinate narration ID, marking them as part of the superordinate narration as well as constituting a new narration. As a last step, each narration (including each embedded narration) identified thus far is analysed for its functional properties – annotators select maximally one major and one minor (in round brackets) function from a list of four types of functions: entertain, factual transfer, illustrate, social management; each has its own tier for each level of narration.

⁴Note that events are only annotated where mediacy has already been established.

Table 2: Annotation setup for an exemplified series of tokens (w1-w18) as in Example 1, segmented into clauses (S). Abbreviations for layers: TOK = token layer, S = clause layer, MEDIACY = layer for mediacy, MOTIVE = optional layer for motivation of mediacy, NARRATOR = layer for ID of mediating authority, PART = optional layer for feature of narrator as being part of narrated or text world, EVENT = event layer, SEQ = layer for sequential relation of events (possible as pointing relations), NAR.ID = layer for identification of narration, F = layer for functional analysis, .ENT = function of entertainment, .FCT = function of factual transfer, .ILL = function of illustration, .SOC = function of social management. Abbreviations for values: non-part = story-external narrator, part = story-internal narrator, e = event, nar = narration.

| | | | | | | | | | | | | | | | | | | |
|-----------|--------------|----|----|----|--------------|----|------|-----|---------------------|-----|-----|-----|-------------|-----|-----|-----|-----|-----|
| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 | w11 | w12 | w13 | w14 | w15 | w16 | w17 | w18 |
| S: | S1 | | | | S2 | | | | S3 | | | | S4 | | | | | |
| MEDIACY: | mediate | | | | mediate | | | | mediate | | | | non-mediate | | | | | |
| MOTIVE: | Past | | | | Past | | | | direct speech, Past | | | | | | | | | |
| NARRATOR: | narrator-001 | | | | narrator-001 | | | | narrator-001-01 | | | | none | | | | | |
| PART: | non-part | | | | non-part | | | | part | | | | | | | | | |
| EVENT: | e- | | | | | | e- | e- | | | | | | | | | | |
| | 001 | | | | | | 002 | 003 | | | | | | | | | | |
| SEQ: | e- | | | | | | e- | e- | | | | | | | | | | |
| | 001 | | | | | | 002 | 003 | | | | | | | | | | |
| | > e- | | | | | | > e- | <> | | | | | | | | | | |
| | 002 | | | | | | 003 | e- | | | | | | | | | | |
| | | | | | | | | ??? | | | | | | | | | | |
| NAR.ID: | nar-001 | | | | nar-001 | | | | nar-001-01 | | | | | | | | | |
| F.ENT_L1: | entertain | | | | entertain | | | | entertain | | | | | | | | | |
| F.FCT_L1: | | | | | | | | | | | | | | | | | | |
| F.ILL_L1: | | | | | | | | | | | | | | | | | | |
| F.SOC_L1: | | | | | | | | | | | | | | | | | | |
| F.ENT_L2: | | | | | | | | | fact | | | | | | | | | |
| F.FCT_L2: | | | | | | | | | | | | | | | | | | |
| F.ILL_L2: | | | | | | | | | | | | | | | | | | |
| F.SOC_L2: | | | | | | | | | | | | | | | | | | |

Throughout these guidelines, we will focus on German as a use case, yet these guidelines are not language-specific. This means that annotators will at times have to determine language-specific categories prior to annotation, as explained in the respective sections.

2 General motivation

2.1 Research questions

Our ultimate goal is to tease apart the notions register variation and narration. The presented annotation scheme allows us to compare various types of texts

that contain narrative passages by defining and annotating a number of parameters that have been suggested in the literature (from linguistics as well as narratology) as constitutive for narration. With this, we aim to answer the following research questions:

- RQ 1:** Is there a “universal” narrative register? More specifically, can we find properties of narration across languages, times, contexts, and text types?
- RQ 2:** How can we recognize and operationalize narrative passages in narrative and non-narrative texts?
- RQ 3:** Are the notions register variation and narration independent or conceptually intertwined? In other words, is narration itself a parameter in the situational-functional context of a register or do different constellations of narrative aspects relate to different registers?

2.2 Narration in its essence

According to recent research in linguistics and narratology, narrative texts are characterized by a number of specific, partly obligatory and partly facultative features (Genette 2010; Scheffel et al. 2014; Lahn & Meister 2016; Zeman 2016; 2018; 2020a,b). It is common to distinguish between narrated events (the abstract substance of the story; *fabula*; *histoire*; *story*; *What?*) and the way they are represented (representation within discourse; *sjuzet*; *discours*; *discourse*; *How?*).

One main characteristic of the discourse strategy narration on the clause-level is the “iconic sequence of events”, for which we use an event-based definition (see Labov & Waletzki 1967; Asher & Lascarides 2003; Zeman 2018). This is accompanied by the idea that at least two events are placed in a temporal or causal relation within the framework of the text. We want to look at whether the sequence in which events are presented in the text corresponds to the sequentiality in the narrated world. As a second characteristic, narratives are said to have a double- or multi-layered structure on the discourse level (thus, on the *how?*-level). This means that events do not simply relate to one another, but are represented or narrated or depicted according to the principles of mediacy and perspectivization. Each of these parameters alone may not be obligatory or exclusively characteristic for narration as they can also occur in passages with other discourse strategies, e. g. instruction, persuasion, description, etc. However, it is the combination of these features which has the potential to indicate narrativity.

Therefore, we present a method to annotate for mediacy, i. e. the fact that something is transmitted by a narrating authority (as coined by Stanzel 1955), in Section 3 and sequentiality in Section 4, outlining indicators for each level that help

annotators to identify mediacy as well as events together with their relation to one another. In addition, we present a categorization of narrations according to their overall discourse functions in Section 6 – these narrative subtypes may be used to different extents in different registers, pointing towards narration as not being a universal category.

Our approach is thus to look at properties on the clause level of texts to annotate narrative passages, with mediacy and sequentiality of events being the minimum requirements for the identification of a narrative text passage. Properties that relate to the text level, i. e. the whole text, are not used to inform the annotation as they are not subject to the goals of these guidelines; such properties include e. g. markers indicating each speaker in a transcription or a play, which would otherwise hint at a narrative passage. The *Parlamentsreden-Corpus* ([Deutscher Bundestag 2023](#)), for example, which contains transcripts of German parliamentary debates, entails narrative characteristics on the overall text level and therefore pertains to the text as a whole: firstly, it is a written record of parliamentary speeches and thereby mediated by the stenographer; and secondly, it includes markers for each speaker, e. g. *Dr. Angela Merkel: [...]*, which resemble indirect speech marking. However, we do not treat this sort of text-level marking as an indicator for mediacy because we focus on the clause level; we treat this as text-external metadata when marked in the text directly.

3 Mediacy

3.1 Background

Mediacy is the starting point for our annotation as it has been determined as one of the key factors for defining a text as being narrated (cf. [Zeman 2016](#); [Genette 2010](#)). Every narration is presented from a specific point of view ([Zeman 2016](#))⁵, i. e. events, words, and impressions of the inner and outer world are *re-presented* or *mediated* by an authority⁶, which we label the narrator ([Zeman 2016](#); [2018](#);

⁵In narratological discourse, mediacy is also seen as a scalable parameter that is used to describe the distance between the represented information and the recipient – this relation is mediated by the narrator ([Genette 2010](#); [Martínez & Scheffel 2009](#)), yet we reduce this notion to a categorical value to increase the consistency across annotators, deciding between either mediate or not mediate.

⁶This can be referred to as “perspectivation” induced by the telling, narration or transmission of events etc. Perspectives can also change within a text, so that perspectivation is not to be considered a fixed or bound category.

Martínez 2011; Genette 2010; Lahn & Meister 2016; Martínez & Scheffel 2009).⁷ The main annotation thus concerns mediacy and in consequence also the narrator (see Section 3.3). We differentiate for each clause whether a narrator is present as well as the number of narrators in each text. This allows for the differentiation of narrations in a single text, e. g. when several interlocutors are telling different stories in a conversation. Further optional annotations about the characteristics of the narrator are possible, e. g. whether a narrator appears in the narrated world themselves.

The narrator as the mediating authority is taken here as an obligatory and integral part of narrative passages, though there are passages where traces of the narrator are harder to determine, making their annotation particularly challenging (Lahn & Meister 2016: 75-76). In order to determine whether mediacy applies in any given passage or not, it is helpful to look out for certain indicators of mediacy in the text: most prominently multi-layered structures. Section 3.2 provides details about such indicators of mediacy and how they inform the annotation decision; however, these indicators are not necessary criteria. They are taken here as sufficient criteria that support a decision about mediacy.

3.2 Indicators of mediacy

Mediacy, through the mediating authority, implies a change in perspective. We can view these perspectives as different layers in the structure of a text, which is why the concept of multi-layered structure is strongly connected to mediacy. A multi-layered structure is generated by the representation of events (by establishing an order of representation, a perspective on the events etc.). Following Zeman (2018), we assume that the layered structure of narration, in its essence minimally a double-layered structure, amounts to the distinction between the level of protagonists and events and the level of the mediating authority or narrator. Due to the fact that narrations can include embedded narrations where protagonists become narrators themselves, there can be more than two layers, i. e. iterations of double-layered structures, and we thus speak of a multi-layered structure. Such a structure is not unique to narrative discourse, as it is also observable in contexts of propositional attitude⁸ expressions. However, we can assume

⁷The focus of the annotation is on the level of the narrating authority, which must be strictly separated from the role of the text producer, as these are not necessarily identical: it is the text producer who shapes and models the narrator, i. e. a specific image of the narrator, who can come forth implicitly or explicitly within the text.

⁸Propositional attitudes can be indicators of multiple layers. However, this class is so heterogeneous that we cannot operationalize it appropriately for our annotation.

that narration necessarily features several levels, i. e. at least two. In this way, identifying multi-layered structures supports the annotator with the detection of mediacy.

A series of linguistic means points towards the existence of multiple layers. These means include *free indirect discourse*, *future of fate*, *historical present* and also *past tense* as well as *representations of speech*. We take these indicators as sufficient criteria because they all establish a multi-layered structure. As a result, these guidelines include various categories such as reports under the umbrella of narrative discourse next to fictional narrations because they display the main characteristics mediacy and sequences of events. We use the functional layer (see Section 6) to then distinguish different types of narrative strategies in order to be able to examine their differences and similarities. This list of indicators is not exhaustive and annotators will have to determine language-specific triggers for multi-layered structures, potentially deriving them from existing or semi-automatic annotations of tokens, lemma or verb morphology.

Free indirect discourse (FID; see for an overview McHale 2019)

In free indirect discourse, the narrator and the protagonist's perspective appear to be merged by using linguistic elements to invoke the level of the protagonist at the same time as using elements to evoke the level of the narrator, e. g. there may be mismatches between tense (past) and temporal adverbials as in (2), where the protagonist's thoughts are reported. While the temporal adverb *heute* 'today' is interpreted from the protagonist's perspective, tense and pronouns (*sein* 'his') are from the narrator's perspective.

- (2) vielleicht weil heute sein dreißigster Geburtstag war
maybe because today his thirtieth birthday was
'maybe because today was his thirtieth birthday' (NoSta-D-Kafka⁹)

Future of fate (FoF; see for an overview Zeman 2020c)

Future of fate describes text passages where the narrator reports on events that are in the past relative to the narrator but in the future relative to the protagonist, thus generating two layers. FoF in German, for instance, is usually characterized by modal verbs which do not have a modal but a future interpretation (in most cases the preterite form of *sollen* 'shall' – see (3a), but also *würde* 'would' – see (3b)), as may be determined by temporal adverbials with a future reading.

⁹Dipper et al. (2013): <https://korpling.german.hu-berlin.de/annis3/?id=c638705b-668f-4748-ae60-3c42870f4466>

- (3) a. doch er **sollte** bald eines Besseren belehrt werden.
yet he should soon a better taught become
'Yet he was soon to be taught better.' (Berner Zeitung 2009)
- b. Sie **würde** schon morgen das Ausmaß des Problems
she would already tomorrow the extent the.GEN problem
erkennen.
realize
'Tomorrow already she were to realize the full extent of the problem.'

Historical Present (HP; see for an overview Fludernik 1991)

With the historical present, the present tense itself indicates mediacy when used within narratives for events that are situated in the past relative to the narrator, e. g. marked by adverbial modifiers as in (4).

- (4) 1918 beendet Ludwig Wittgenstein seinen Tractatus
1918 finished Ludwig Wittgenstein his Tractatus
Logico-Philosophicus.
Logico-Philosophicus
'In 1918, Ludwig Wittgenstein finished his Tractatus Logico-Philosophicus.'
(Leibniz-Institut für Deutsche Sprache 2018)

Past Tense (Past)

Past tense (or similar linguistic devices used to refer to past events) generally marks a distance between the speaker and the reported events, as in (5), thereby separating two layers between which is mediated.

- (5) Als ich diesen Sommer mein Bachelor beendet habe, wollte ich
as I this summer my Bachelor finished have, wanted I
eigentlich versuchen eine arbeit zu finden
actually try a job to find
'As I finished my bachelor this summer, I wanted to try to find a job'
(Falko¹⁰)

Representation of speech (see also Brunner et al. 2020)

A further major overt indicator of mediacy is the representation of speech¹¹, in

¹⁰The example is from Falko (Lüdeling et al. 2008), a learner corpus. The learner utterances are represented in the corpus as the learners wrote them. See corpus online here: <https://huberlin.de/falko>

¹¹When texts as a whole are representations of speech, such as transcripts, these are not annotated on the clause level, as this is part of the metadata for the text.

which utterances are reproduced, thereby marking a switch of perspective and consequently an increase in mediacy. Texts that involve narration tend to include representations of e. g. dialogues and monologues as simulations of discourse.¹² This can be rendered, for example, by the distinction between *direct*, *indirect* and *narrated* speech¹³ (Genette 2010; Lahn & Meister 2016). Brunner et al. (2020) have developed detailed guidelines for the representation of speech in German, which may alternatively be used to mark these indicators of mediacy in the tier MOTIVE; while their guidelines differentiate oral speech, thought and writing, we subsume these categories under the overall term “speech”.

Direct speech refers to a statement or thought of a protagonist that is generally represented by the use of quotation marks in modern written language contexts¹⁴ or equivalent graphic or oral devices. Optionally, one may find introductory elements such as *inquit*-formula, e. g. ‘he said’ or ‘she asked’, (see (6a)) or independent quotations (see (6b)).¹⁵ Direct speech is a “classical” device to represent speech in narrative texts and invokes mediacy by giving the impression that an “unfiltered” statement by a third party is reproduced.

- (6) a. “Also Timo, was genau ist gestern mit deinem Bruder Julio so Timo what exactly is yesterday with your brother Julio passiert”, fragte der Polizist
happened asked the policeman
‘So, Timo, what exactly happened with your brother Julio yesterday, asked the policeman’ (Falko¹⁶)

¹²We conceptually separate the notion of transtextuality, i. e. links between texts, from representation of speech at this point.

¹³Narrated speech here corresponds to their reported speech. With this, we avoid the possible confusion from different uses of indirect and reported speech.

¹⁴However, in e. g. historical contexts other graphic devices (such as changes in size or color of the writing, markings by specific layout choices, etc.) may occur in order to introduce this kind of speech representation. In oral contexts, statements of third parties are often indicated by an intentional change of intonation and of voice (cf. *mimesis*) being accompanied by gestures and facial expressions, etc. The reproduced statement in oral communication may additionally involve specific linguistic features imitating e. g. the sociolect, idiolect, dialect, personal style etc. of the third party.

¹⁵In other language contexts, independent direct speech can be recognized when turn-taking without direct introduction of the respective speakers can be detected. In these cases, only contextual/co-textual and media-dependent information can help to distinguish between the protagonists.

¹⁶<https://korpling.german.hu-berlin.de/annis3/?id=a05cc945-46da-4894-ba5e-ae5e8c5c33db>

- b. “Die meisten Universitätsabschlüsse sind nicht praxisorientiert
the most university.degree are not practice.oriented
und bereiten die Studenten nicht auf die wirkliche Welt vor.”
and prepare the student not on the real world for
‘Most of the university degrees are not practically oriented and do not
prepare students for the real world.’ (Falko¹⁷)

Indirect speech (also called “reported or transposed speech”) means that essential features of the protagonist’s verbatim statement or thought are included but may have received modification by the narrator. Thus, the resulting text may contain individual idiolectic or stylistic features of the narrator. The tendency is to not render indirect speech within quotation marks in written media, but there may be co-occurring changes within the reproduced statement, including transpositions of first person pronouns to the 2nd or 3rd person, of demonstratives as well as of spatial and temporal indexicals, changes in mood etc. (on typological characteristics, see Spronck & Nikitina 2019). The statement may be introduced by an *inquit*-formula (see (7a)) or be independent (see (7b)).

- (7) a. Der Arzt sagte mir, dass ich eine Magen-Infektion hatte
the doctor told me that I one stomach-infection had
‘The doctor told me that I had a stomach infection’ (Falko¹⁸)
b. Tran sei hier noch nicht diskriminiert worden
Tran be.SUBJ here yet not discriminated been
‘Tran was presumably not yet discriminated here.’ (Falko¹⁹)

Lastly, *narrated speech* stands for a statement or thought of a protagonist that underlies the complete control of the narrator (see (8)). The narrator decides which parts are to be reproduced in which way, e. g. by shortening, abstracting and also occasionally omitting the content of the statement, for instance by summarizing statements and emotions of protagonists via *verba dicendi* or *verba sentendi*. This type of speech reproduction shows the strongest degree of mediacy and thus distance between the original statement and its representation. The verbatim statement of the protagonist cannot be retraced on the basis of the text.

¹⁷The writer is quoting the prompt in their essay: <https://korpling.german.hu-berlin.de/annis3/?id=5b484492-23fc-4f24-99f1-35898d9612b2>

¹⁸<https://korpling.german.hu-berlin.de/annis3/?id=834aa620-7ba7-45a3-90ef-333a823f09d8>

¹⁹<https://korpling.german.hu-berlin.de/annis3/?id=50ac4a30-3ade-4707-938f-456ff19ab95f>

- (8) Eines Tages beschloss seine liebenswerte Frau, den reichen Mann zu
one day decided his lovely woman the rich man to
verlassen
leave

‘One day, his lovely wife decided to leave the rich man.’ (Falko²⁰)

3.3 Annotating mediacy

When mediacy is detected, the clause is annotated on the MEDIACY tier with the value *mediated*.

Tier characteristics for mediacy:

tier name: MEDIACY
values: mediate, non-mediate
exponent: clause

The motivation behind this decision may be optionally annotated on the tier MOTIVE using the indicators of mediacy (see Section 3.2) as annotation values. Other motivations for assuming mediacy than presented in Section 3.2 may be annotated here as well.

Tier characteristics for mediacy motivation:

tier name: MOTIVE
values: FID, FoF, HP, Past, direct speech, indirect speech, narrated speech, etc.
exponent: clause

Furthermore, an ID is provided for each narrator on the tier NARRATOR, i. e. the value *narrator* and an indexical number consisting of three digits. The number indicates which narrator is present in the respective clause; the number increases by 1 when a new narrator is introduced. Narrators that have a superordinate narrator receive the ID of the superordinate narrator plus a two digit identification

²⁰<https://korpling.german.hu-berlin.de/annis3/?id=4a838f30-bd77-4082-aabf-3fcc0b500139>

number. This reflects how a new layer is added in the multi-layered structure by the embedded narration, e. g. when a character starts telling a story. When previous narrators reappear, the annotation uses the same indexical number as before for this particular narrator so that unique narrators can be identified throughout a text.

Tier characteristics for the narrator:

tier name: NARRATOR

values: narrator-001, narrator-002, narrator-002-01, none

exponent: clause

Whenever a narrator is present, more characteristics about the narrator may be specified. We suggest to annotate whether the narrator is part of the narrated world as this will influence the type of narration.²¹ This point is about knowledge and perspective: Who has access to or represents information from what perspective? This is also referred to as the distinction between *homodiegetic* and *heterodiegetic* (Genette 2010; Lahn & Meister 2016).

- the narrator is part of the narrated world as a protagonist (*homodiegetic*)

TAG: *part*

(9) Als ich anfing gab es leider keine relevante -
when I begin exist it unfortunately no relevant
d.h. deutschorientierte Arbeitsaufgaben
that.means german.oriented exercises

‘When I started, there were unfortunately no relevant German oriented exercises.’
(Falko²²)

- the narrator is not part of the narrated world as a protagonist (*heterodiegetic*)

TAG: *non-part*

²¹Further features about the narrator could include the overtness of the narrator, which will be reflected in the mediacy of the narration (see for example Lahn & Meister 2016), yet we reserve such detailed annotations for respective research questions.

²²<https://korpling.german.hu-berlin.de/annis3/?id=1606709d-2653-4e17-a4e2-37fc197418af>

- (10) Eines Nachts, waehrend er im Bett liegt, faellt ihm die Idee auf,
one night while he in bed lie fall him the idea on
seine Waffe von seinem Kleiderschrank rauszunehmen
his weapon from his closet take.out
'One night, while lying in bed, he had the idea to take out the gun
from his closet' (Falko²³)

Note: The use of specific personal pronouns (e. g. 1SG, 3SG.M, 3SG.F) is not a grammatical indicator for homodiegetic or heterodiegetic instances (Genette 2010).

Tier characteristics for the narrator's involvement in the narrated world:

tier name: PART

values: part, non-part

exponent: clause

4 Sequentiality of events

Sequentiality is generally understood as the “linear, unidirectional succession of elements” (Grabes 2014: § 1), where elements can be any type of entities (numbers, physical objects, events, abstract objects). In narratological terms, sequentiality is defined as the sequence in which the narrated events are presented, cf. Grabes (2014) and Zeman (2018), irrespective of whether event boundaries overlap partly, completely or not at all. It is seen as a necessary but not sufficient condition for narration.

The definition of *events* in narratology includes two different readings:

The term ‘event’ refers to a change of state as one of the constitutive features of narrativity. We can distinguish between event I, a general type of event that has no special requirements, and event II, a type of event that satisfies certain additional conditions. A type I event is any change of state explicitly or implicitly represented in a text. A change of state qualifies as a type II event if it is accredited in an interpretive, context-dependent decision – with certain features such as relevance, unexpectedness, and unusualness (Hühn 2013: 1).

²³<https://korpling.german.hu-berlin.de/annis3/?id=9919a9b8-c66d-44cf-b757-907bf436fd58>

The interpretation of the term *change of state* differs in linguistics and narratology. Whereas in narratological terms, a change of state is any sort of change in the established diegetic world, in linguistics, a change of state is determined via its aspectual properties, i. e. a change of state is a transition of one state to another either in a durative or a punctual manner, cf. Vendler (1967).

Because the conceptualization of events in literary studies differs from the usual understanding of events in linguistics²⁴, we describe in detail our functional application of the notion *event* in the following. Generally, we follow the first reading of *event* from Hühn (2013, see previous paragraph) for our annotation of sequentiality. In linguistic terms, we adopt a wide conceptualization of *events* that includes not only changes of state, but also *non-culminating events* and *states* (cf. Davidson 1967; Parsons 1990). *Events* in this wider sense can participate in sequences, even though they do not refer to a change of state. They are conceptualized as having or being plausibly able to have at least one temporal boundary, i. e. excluding generic descriptions. They can form background information for culminating events (see e. g. (17)) or come to an end without explicit lexical encoding of an ending (see (14)).

4.1 Event encoding

Events are typically encoded in verbs or VPs, respectively (cf. (11a) and (11b)).²⁵ However, these are not the only linguistic means to describe events. Nouns (11c) and combinations of copula and a predicative (11d) can also denote events, as described by Stede (2018). Nouns, for instance, may denote happenings and actions that include temporal boundaries as with *Sprechen* ‘speaking’ in (11c-i), which in the context of a clause such as *Sie hat beim Sprechen das Atmen vergessen*. ‘While speaking she forgot to breathe.’ refers to a time span where the subject is speaking and in that time frame the subject performed another action, here namely the forgetting to breathe. For the identification of sequentiality, all types of event denoting expressions have to be annotated. With this point of view, we follow ideas from TimeML in Saurí et al. (2006) and Stede (2018).

- (11) a. verbs
- i. *sprechen* ‘speak’,
 - ii. *leuchten* ‘shine’,

²⁴Note that in linguistics, there is also a variety of theories of events and/or eventualities, see Maienborn (2019) for an overview.

²⁵Here again, the presented list is partially language-specific and will have to be extended or modified for other languages.

- iii. *verschwinden* ‘disappear’
- b. **verb phrases (VP)**
 - i. *ein Gedicht aufsagen* ‘to recite a poem’,
 - ii. *Apfelkuchen essen* ‘to eat apple pie’
- c. **nouns**
 - i. *beim Sprechen das Atmen vergessen* ‘to forget breathing while speaking’,
 - ii. *Besprechung* ‘discussion’,
 - iii. *Lauf* ‘(a) run’
- d. **predicatives**
 - i. *reich werden* ‘to become rich’,
 - ii. *am Strand sein* ‘to be at the beach’,
 - iii. *zu Hause bleiben* ‘to stay at home’

4.1.1 Event annotation

Any single event constitutes a basic event unit. Basic event units are annotated on the **event** layer at the level of the token, i. e. the exponent of the variable **EVENT** is a single token, e. g. a finite verb like *kommt* ‘came’ in (12). If the event is expressed by more than one token like in the analytic tensed verb form *habe geholfen* ‘HAVE help.PTCP’, the (semantic) head is taken as the event exponent, i. e. *geholfen*.

- (12) Interessant genug habe ich dieser Semester einen Student **geholfen**,
interesting enough have I this semester a student help.PTCP
der Englisch nicht gut sprechen und schreiben konnte. Er kommt
who English not good speak and write could he come
aus einem Township in Bloemfontein, und **hatte** wirklich Problemen
from a township in Bloemfontein and had really problems
mit seine Studium wegen seine Unfähigkeiten mit Englisch.
with his studies because.of his inability with english
‘Interestingly I helped a student this semester who couldn’t speak or write
English well enough. He is from a township in Bloemfontein and really
had problems with his studies because of his inability with English.’
(Falko²⁶)

²⁶<https://korpling.org/annis3/?id=a846b722-2a11-407a-b67c-558e2e98df98>

Guidelines for the Annotation of Parameters of Narration

Every event exponent is tagged as an event with the tag = *e* and receives a unique id-number per text, e. g. the event exponent *geholffen* has the tag *e-001*.

| | | | | | | | | | |
|------|----|-----|----|----|-----------|-----|-----|-------|-----|
| TOK: | w1 | ... | w9 | w8 | geholffen | ... | w18 | hatte | w20 |
| SEQ: | | | | | e-001 | | | e-002 | |

The following types of tokens are tagged as events:

1. Verbs:

- a) finite or infinite forms of full verbs denoting an individual event that occur(red) or happen(ed)

(13) Als ich joggen war, ist mir **eingefallen**, dass ich
 when I jogging was is 1P.SG.DAT remember.PTCP that I
 eigentlich beim Zahnarzt sein sollte.
 actually at.the dentist be should
 ‘When I was jogging, I realised that I should have been at the dentist.’

- b) finite or infinite forms of full verbs denoting an individual state that holds or pertains for a specifiable amount of time.

(14) Als ich **joggen** war, ist mir **eingefallen**, dass ich
 when I jogging was is 1P.SG.DAT remember.PTCP that I
 eigentlich beim Zahnarzt sein sollte.
 actually at.the dentist be should
 ‘When I was jogging, I realised that I should have been at the dentist.’

2. Nouns:

- a) nominalizations denoting individual events that occur(red) or happen(ed)

(15) Nach der **Abdankung** König Peters wurde direkt
 after the abdication king Peter.DAT became directly
 Königin Petra gekrönt.
 queen Petra crowned
 ‘Immediately after the abdication of king Peter, queen Petra was crowned.’

b) nominalizations denoting individual states that pertain or hold

- (16) Bei der **Ankunft** des Zuges ereignete sich etwas
at the arrival the train.DAT happened REFL something
Unerhörtes.
unheard-of
'With the arrival of the train, something unheard-of happened.'
- (17) Er hat sich beim **Sitzen** einen Bandscheibenvorfall
he has himself during.the sitting a disc.herniation
zugezogen.
contracted
'He contracted a disc herniation while sitting.'

3. **Predicatives:** nouns, adjectives and prepositions functioning as the head of a predicative that holds for a specifiable amount of time

- (18) Die Armen wurden immer **ärmer**, die Reichen immer **reicher**.
the poor became always poorer the rich always richer
'The poor became increasingly poorer, the rich increasingly richer.'

4.1.2 Non-Events

Not all verbs or VPs denote events. Non-eventive verbs occur in the following linguistic contexts and are not counted nor tagged as events:

- embedded events in complements²⁷

- (19) Complement clause:
Ida behauptete, dass Emil das letzte Eis **gegessen** hat.
Ida claimed that Emil the last ice.cream eaten has
'Ida claimed that Emil ate the last ice cream.'

²⁷In event literature, direct or indirect speech reproduction is often also not counted as event, but we do count it, because it's crucial for us, see mediacy (3). Also, dialogues can include embedded narrations mediated by the respective speaker and constructed from events.

- events embedded under negation and modality²⁸

(20) a. Negation:

Emil hat das letzte Eis nicht **gegessen**.

Emil has the last ice.cream not eaten

‘Emil did not eat the last ice cream.’

b. Modality:

Ich könnte das **lernen**.

I could this learn

‘I could learn this.’

- pseudo events: full verbs only seeming to denote events but which cannot be anchored temporally with the actual events in the story, i. e. usually iterative or non-specific events

(21) a. Iterative:

Emil **isst** gewöhnlich das letzte Eis.

Emil eats usually the last ice.cream

‘Emil usually eats the last ice cream.’

b. Non-specific:

Wenn die Frau auf Arbeit ist, **kümmert** der Mann sich um

when the woman at work is, take.care the man REFL of

den Haushalt.

the household

‘When the woman is at work, the man takes care of the house.’

²⁸There are, however, exceptional contexts, where full verbs embedded under modal verbs can denote events. For instance, the German modal verb *wollen* can have a future auxiliary function in narrative contexts. In relation to the narrator perspective, events embedded under *wollen* are then in the past and part of an event sequence, whereas in relation to the protagonist they are situated in the future. In that case, the embedded full verb would be tagged as an event.

(i) das Bürle gieng in die Stadt und **wollte** das Fell dort verkaufen

the boy went in the city and wanted the fur there sell

‘the boy went to the city and wanted to sell the fur there’

(Märchenkorpus

(Walter 2013): <https://korpling.german.hu-berlin.de/annis3/?id=7e988edf-0fda-4723-abd7-d4fba322c3e3>)

- non-events: full verbs not denoting any event, e. g. generic states²⁹

- (22) a. Das letzte Eis **schmeckt immer** am besten.
the last ice.cream tastes always best
'The last ice cream always tastes best.'
- b. Hunde **sind** schnelle Lerner.
dogs are fast learner
'Dogs are quick learners.'

- imperatives, as they are not temporally anchored³⁰

- (23) a. **Lass** uns ins Museum gehen.
let us in.the museum go
'Let us go to the museum.'
- b. **Reich** mir bitte mein Telefon.
give me please my phone
'Hand me my phone, please.'

Tier characteristics for the annotation of events:

tier name: EVENT

values: e-001

exponent: token

²⁹For a detailed overview over genericity and different types of non-eventive generic clauses, see Gerstner-Link & Krifka (1993: 967f.).

³⁰This includes linguistic forms that are not imperative but perform the same speech act. For instance, in (i) we do not assume an event sequence, as the events are not actually performed.

- (i) Du ziehst jetzt deine Schuhe aus und wäschst dir die Hände.
you pull now your shoes off and wash 2P.SG.DAT the hands
'You will take off your shoes and wash your hands now.'

4.2 Temporal ordering

On the temporal ordering layer, we annotate the temporal ordering of events in the form of two-place relations. In order to do so, all events of a text have to be marked up with unique identifiers (as described in the previous section). We annotate three possible relations between events. These relations can be expressed with connectors and temporal expressions (see below) but they also occur unmarked.

Stede (2018) uses three basic temporal relationships: (1) sequences of events in any direction with and without time in between them; (2) inclusion of events within other events irregardless of how long they are in relation to one another and when the included event starts and ends within the containing event, as long as it is completely contained within the containing event; (3) overlapping of events: the beginning or start of event A is within another event B while the end of event A is not in event B. For pragmatic reasons, we slightly modify this approach in merging the categories (2) “inclusion” and (3) “overlapping”. We further add the relation “underspecified” for cases where the relation is not explicitly encoded.

1. Succession:

An event is followed or precedes another event.

TEMPLATE: event X and then event Y

POSSIBLE MARKERS: *bevor* ‘before’, *danach* ‘after’, *dann* ‘then’, *am nächsten Tag* ‘the next day’, *morgen* ‘tomorrow’

TAG: e-001 > e-002

- (24) Bevor sie am nächsten Tag zum Bewerbungsgespräch **aufbrach**,
 before she on next day to job.interview started
überprüfte sie nochmal ihre Unterlagen.
 checked she again her documents
 ‘Before she left for her interview the next day, she double-checked her documents.’

| | | | | | | | | | | |
|-------|----|----|----|----|---------------------|-------|----|----|----|-----|
| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
| EVENT | | | | | e-001 | e-002 | | | | |
| SEQ. | | | | | e-002 > e-001 | | | | | |

2. Overlap:

An event overlaps in part or completely with another event. This also contains total inclusion.

TEMPLATE: event Y intersects with event X

POSSIBLE MARKERS: *als* ‘when’, *gerade* ‘just’, *während* ‘while’

TAG: e-001 c e-002

- (25) Als ich gerade dabei war mein Essen zu **kochen**, **klang** das
 when I just about was my food to cook rang the
 Telefon.
 phone
 ‘When I was cooking my food, the phone rang.’

| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
|-------|----|----|----|----|---------------------|-------|----|----|----|-----|
| EVENT | | | | | e-001 | e-002 | | | | |
| SEQ. | | | | | e-001 c e-002 | | | | | |

3. Underspecified:

An event is temporally ordered to another event, but the exact sequential relation cannot be determined.

TEMPLATE: event X temporally ordered wrt event Y

POSSIBLE MARKERS: *und* ‘and’

TAG: e-001 <> e-002

- (26) Ich war **schwimmen** und **besichtigte** die Stadt.
 I was swimming and visited the city
 ‘I went swimming and visited the city.’

| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
|-------|----|----|----|----|----------------------|----|-------|----|----|-----|
| EVENT | | | | | e-001 | | e-002 | | | |
| SEQ. | | | | | e-001 <> e-002 | | | | | |

To identify the ordering relation on the word level we use tense, aspect, and aktionsart of the verb. On the clausal and phrasal levels, the relations may be marked by connectors and temporal expressions (cf. [Stede 2018](#)). The latter can either be fully specified (e. g. *am 23.08.1957* ‘on August 1957, 23rd’, *Viertel vor zwölf* ‘quarter to twelve’), specified relative to the time of utterance (e. g. *letzten Dienstag* ‘last Tuesday’, *morgen* ‘tomorrow’) or specified relative to another time expression (e. g. *am nächsten Tag* ‘on the next day’). Relative time expressions usually refer to intervals (e. g. *im Sommer* ‘in summer’) instead of specific points in time. The German examples are again merely illustrative, for this step requires the development of language-specific lists of indicators.

We mark the relations at the exponent of the event, that is a token. The relation is always marked at the first argument of the relation. That is, a before-relation is marked at the first event. An in-relation is marked at the including event, not at the included event. This is exemplified in (27).

- (27) Event annotation: event tier (EVENT) in round brackets, sequentiality tier (SEQ) in square brackets

Emil sitzt (e-001, [e-001 c e-002 e-003 e-004 e-005, e-001 > e006]) in
Emil sits in

seinem Sessel. Währenddessen liest (e-002, [e-002 <> e-003]) er
his seat meanwhile read he

Zeitung und trinkt (e-003, [e-003 > e-004]) seinen Kaffee. Dann
newspaper and drink his coffee then

kommt (e-004, [e-004 > e-005]) eine Fliege und
came a fly and

setzt (e-005, [e-005 > e-006]) sich auf seine Nase. Er
sits itself on his nose he

steht (e-006, [e-006 > e-007]) auf und holt (e-007) die Fliegenklatsche.
stands up and gets the swatter

‘Emil is sitting in his chair. Meanwhile, he is reading the newspaper and drinking his coffee. Then, a fly appears and sits down on his nose. He stands up and gets the swatter.’

Due to the fact that sequences of events can occur on different layers of a text (narrator or protagonist level), we connect sequences of events only if they are situated on the same level, i. e. only events on the protagonist level can be in a sequence and only events on the narrator level can be in a sequence. In some cases, the sequences on the protagonist and narrator level can be intertwined.

These cases can be reconstructed from the overlap between the temp-ordering and the narrator level.

Tier characteristics for the sequence of events:

| |
|--|
| <p><i>tier name:</i> SEQ</p> <p><i>values:</i> > (before), c (overlap), <> (underspecified)</p> <p><i>exponent:</i> token</p> |
|--|

5 Identification of narration

To identify a stretch of text as narration, we have determined two basic requirements: a) there is a mediacy at play, indicating the presence of a narrator (see Section 3) and b) at least two events depicted in associated passages stand in a sequential relation to another, being either in temporal succession, temporally overlapping or otherwise temporally related, e. g. temporally underspecified (see Section 4). The narration will now span from the first clause with an event in the sequence to the last adjacent clause with an event in the sequence, whereby each narration is consecutively numbered (nar-001).

| | | | | | | | | | | |
|--------|---------|----|----|----|----|----|------|----|----|-----|
| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
| S | S1 | | | S2 | | | S3 | | | |
| NAR.ID | nar-001 | | | | | | none | | | |

Table 3: Simple span of clauses in a narration

There are a few cases where it will not be straightforward to create a span for one narration:

1. Split narration:

In case clauses belonging to a single narration are not adjacent, for example when clauses with only non-events are in between or when a different, independent narration intervenes, the annotation span for the narration cannot include the parts not belonging to the narration. The solution is to label all adjacent parts with the same label and repeating the label for the next stretch of text belonging to this narration. The fairy tale excerpt in (28) demonstrates the clause by clause differentiation, for the first clause (*nar-001*) fulfils the requirement for narration, involving mediacy and an

event, while the second clause has no event due to its modal structure.³¹ Therefore it does not receive an annotation on the NAR.ID level, but the narration *nar-001* is continued afterwards where the woman fires up the stove more.

| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
|--------|---------|----|----|----|----|----|----|---------|----|-----|
| S | S1 | | | S2 | | | | S3 | | |
| NAR.ID | nar-001 | | | | | | | nar-001 | | |

Table 4: Split spans of clauses in a single narration: the non-narration in S2 intervenes between the parts of narration *nar-001* that encompasses S1 and S3.

- (28) Sie machte also auf ihrem Herd ein Feuer zurecht, und
she make so on her stove a fire ready and
damit es desto schneller brennen sollte, zündete sie es
in.order.of it the.more faster burn should ignite she it
mit einer Hand voll Stroh an.
with one hand full straw on
‘She started a fire on her stove and in order that it should burn
faster, she fired it up with a handful of straw.’ (Maerchenkorpus³²)

2. Embedded narration:

When a narration is embedded in a larger narration, it will receive the same label as the one it is part of but with the addition of another consecutively numbered ID attached to the label. Example: larger narration has the label = *nar-004*, embedded narration receives the label = *nar-004-01*. Any additional narrations on the first embedding level increase the numbering while any additional layer of embedding adds a number (-01) as in *nar-004-01-01* for embedding level 2. The excerpt of a fairy tale in (29) exemplifies this, starting with the main narrator describing the coal’s speaking event (*nar-001*) before adding what the coal said in direct speech, where the coal acts as a narrator telling their story of escaping the woman’s fire (*nar-001-01*) – which is still part of the main narrator’s storytelling –, and

³¹Although the clause does receive an annotation for mediacy – as we have an instance of FID where the narrator portrays the inner state of the protagonist –, the clause only denotes an intention and not an event according to our definition, so that the passage does not receive a narration annotation. It is exactly this decomposition of narrative aspects that will help us to look into and also compare the linguistic characteristics of these different levels.

³²Walter 2013: <https://korpling.german.hu-berlin.de/annis3/?id=3c6ba29a-0200-4277-a109-89024f9a63de>

finally moving back to the main narrators description of the reaction by the coal's addressees (nar-001).

| TOK: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
|--------|---------|----|----|------------|----|----|----|---------|----|-----|
| S | S1 | | | S2 | | | | S3 | | |
| NAR.ID | nar-001 | | | nar-001-01 | | | | nar-001 | | |

Table 5: Embedded narration in a larger narration: S2 is part of the narration nar-001, which encompasses S1, S2 and S3.

- (29) Die Kohle antwortete „ich bin zu gutem Glück dem Feuer
the coal answered I am to good luck the fire
entsprungen [...] wollen wir [...] auswandern und in ein fremdes
out.jump want we emigrate and in a foreign
Land ziehen.“ Der Vorschlag gefiel den beiden andern
land move the suggestion like the both others
‘The coal answered: “Luckily, I escaped the fire by jumping out ...
do we want to move out of the country?” The other two liked this
suggestion.’ (Maerchenkorpus³³)

Tier characteristics for narration identification:

tier name: NAR.ID

values: nar-001, nar-002, nar-003, nar-003-01, none

exponent: clause

6 Function of narration

While the previous annotations have focused on the internal setup of narration, the “function” of narration is concerned with what purpose narration has in the larger context of the text. As Biber et al. (2021) has shown, coherent stretches of discourse usually have multiple communicative purposes such as description, commentary, figuring things out etc. These are potential register-related parameters that are presumed to have a strong influence on register-related phenomena,

³³Walter 2013: <https://korpling.german.hu-berlin.de/annis3/?id=dadc130b-9579-4629-9c4a-0c1284d6a981>

for “the register perspective focuses on the ways in which linguistic features occur frequently and pervasively in texts because they serve the communicative functions required by the situational contexts of the texts” (Biber et al. 2021: 22). Narrative characteristics are pervasive in many of these communicative purposes and have previously also been used as their own category in communicative goal taxonomies (cf. Quasthoff et al. 2017; Neumann 2014 among others). However, Biber et al. (2021) argue that these taxonomies do not capture what they find in corpora, in particular spoken data, as goals of single communication units, i. e. the smallest coherent stretch of text, were often hard to define, and therefore have unsatisfactory inter-annotator agreements; furthermore, communication units were found to usually serve more than one single purpose.

For us, the big question is whether narrative passages serve a universal function in themselves and therefore also lead to the selection of particular linguistic features in the sense of register features, i. e. whether narration is a universal register. Seeing that narration co-occurs with many other proposed communicative purposes, however, it might also be the case that different functions of narration lead to the use of different linguistic phenomena, therefore belonging to separate registers.

To answer this research question, we thus need to look at narration in its specific context and answer the question why text producers decided to make use of narrative strategies, meaning what function do they serve in each context. As we understand it, narration serves various specific communicative goals in human interaction, be it illustrating a point to win an argument or making a joke. In the following, we create annotation categories for these functions based on the context in which narrative passages occur.

6.1 What to annotate?

The function of narration is to be annotated for a complete narrative passage, which we will take as one narration unit. Each narration unit consists of only one continuous sequence of events. It may consist of more than one clause, but must consist of more than one event (see criteria for narration in Section 5). Narration units are therefore minimal stretches of narrative text that create a coherent whole. Any such unit must have the characteristics described in Sections 3-4.

The annotation for function focuses on the following main questions: why did the narrator choose to tell a narration in this particular context and what is the central goal of the narration unit? Due to the fact that annotators cannot ask the narrator directly, our annotation is based on the author’s assumed goals which are deduced from the way a narration unit is embedded and what it is

made of. Take for instance a doctor's appointment where the doctor asks: "How did the injury occur?" The patient will most likely give an account that includes narrations of how the injury occurred. These provide factual information about past events, as in (30). This will be one main function of the narration. A narration unit may also serve different purposes or functions at the same time. Although in theory all functions may be at work in one narration to equal extents, we propose to limit the main functions to two in each narrative passage for pragmatic reasons similar to the procedure in [Biber et al. \(2021\)](#). For this reason, we choose to assign at most two functions to any narration unit. If one function is assumed to be secondary to the other, it may be annotated in rounded brackets.

- (30) **factual transfer:** doctor asked patient to report an accident
Ich wollte aus der Badewanne aussteigen und bin dabei
I wanted from the bath.tub step.out and am in.doing.so
ausgerutscht. Ich bin voll mit dem Knie aufgekommen und seitdem
slipped I am fully with the knee land and since
knirscht es, wenn ich das Bein anwinkle.
crunch it when I the leg bend
'I wanted to step out of the bath tub when I slipped. I landed on my knee
and since then there is a crunching sound when I bend the leg.'

Important for determining the function(s) of a narration is the addressee. It is not necessarily an actual individual that plays a role here, as it could be an imagined addressee by the narrator ([Lahn & Meister 2016](#)). Annotators will have to make assumptions about the intended addressee to answer the question why the narrator narrates something at all. The following primary functions (exemplified with sub-categories) have been determined for narrations (see similar parameters for communication purposes in [Biber 1994: 44](#)).

1. entertain
2. factual transfer
 - e. g. reporting
 - e. g. claiming
3. illustrate
 - e. g. persuading
 - e. g. explaining

- e. g. enabling
4. social management
- e. g. social bonding / group strengthening
 - e. g. conveying morals

The *entertaining* function is reserved for narrations that are primarily used for the sake of the narration itself, to bring joy to the listeners, such as jokes and dreams. *Factual transfer* is the main purpose of narrations that are used to report events or at least present events as if they are reported (claiming) – the narrator states (assumed or believed) facts and wishes to transfer the information in order to inform the addressee. With narrations for *illustration* purposes, there are various contexts which the illustration may serve, including for example argumentative texts where the narration supports an argument and is employed to persuade, or texts in which the narrator is in the process of explaining something to the addressee (e. g. instructions) and similarly when the narrator tries to enable the reader to do something and uses narrations to demonstrate. The last category with respect to a narration's function is *social management*, which includes all situations in which narrators find themselves in a social circle and they narrate to navigate this social space, for instance in order to bond with circle members or to strengthen existing bonds with peers through the sharing of experiences and evoking emotions; furthermore, social management may also include using narrations to convey morals and important social categories, for instance to children, thus establishing a sense of social unity in a society. This last sub-category, conveying morals, is a good example for the intersection between some of these functions, for narrations that convey morals also illustrate to some extent, although in this case it is left implicit what is being illustrated, so that social management should receive the priority as a function.

6.2 Annotation

For each narration level, we create one annotation layer for each function, resulting in four functional annotation layers for the top-level narration that is not embedded in another narration, namely a level for entertain (F.ENT), factual transfer (F.FCT), illustrate (F.ILL) and social management (F.SOC). For any narration that is embedded in the top-layer, there will be another set of these layers. Each functional layer begins with the label "F", followed by a label for the specific primary function; a level code indicates to which narration level this function annotation applies (L1, L2).

At most two functions may be picked at any given level, indicated by the respective function value (*entertain, fact, illustrate, social*). In case one function is secondary, the value may be put into round brackets: (*illustrate*). Otherwise, there is no annotation on the particular function level. When narrations are broken up, e. g. into clauses or by non-narrative text segments in between, each individual part of the narration unit receives the (same) functional annotation on the respective level and function layer. As any narration includes all its embedded narrations, the embedded narrations are also annotated at the superordinate level, receiving the same functional annotation as the superordinate narration. The embedded narration receives its independent function annotation on the respective level, e. g. when a wife tells her husband how their daughter convinced her to let her stay out late with a story about a kid without friends who never went out, this would result in the whole narration from the wife being annotated as social management on level L1, including the parts addressing the daughter’s story, and the embedded narration originally told by the daughter receives an annotation for the illustration function on level L2 (see also Table 6). Any embedded narration thus inherits the function from the superordinate level, as it is part of this narration and must inevitably also serve the same goal as the whole. In its own domain, however, the embedded narration can serve a different function with respect to its context inside of the other narration, e. g. the daughter uses narration to convince the mother – *inside* of the mother’s story, which at this sub-level is an illustration, whereas that daughter’s narration as *part* of the mother’s narration serves to strengthen the bond with the husband together with the rest of the top-level narration.

| tok: | w1 | w2 | w3 | w4 | w5 | w6 | w7 | w8 | w9 | w10 |
|-----------|-----------|----|----|--------------|----|----|-----------|----|----|-----|
| S | S1 | | | S2 | | | S3 | | | |
| Nar.ID | nar-01 | | | nar-01-01 | | | nar-01 | | | |
| F.Ent_L1: | entertain | | | entertain | | | entertain | | | |
| F.Fct_L1: | | | | | | | | | | |
| F.Ill_L1: | | | | | | | | | | |
| F.Soc_L1: | | | | | | | | | | |
| F.Ent_L2: | | | | | | | | | | |
| F.Fct_L2: | | | | fact | | | | | | |
| F.Ill_L2: | | | | (illustrate) | | | | | | |
| F.Soc_L2: | | | | | | | | | | |

Table 6: Functional annotation of an embedded narration in a larger narration: narration nar-01 has the function to entertain, the embedded narration nar-01-01 has the function to transfer facts and also secondarily to illustrate.

Tier characteristics for the narration function:

tier name: F.ENT, F.FCT, F.ILL, F.SOC

values: entertain, fact, social, illustrate

exponent: all narration units, repeated for each clause in one unit

Funding information

These guidelines have been developed in the course of the research undertaken within the CRC 1412 “Register: Language Users’ Knowledge of Situational-Functional Variation” at Humboldt-Universität zu Berlin. The research is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – CRC 1412, 416591334.

References

- Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation* (Studies in Natural Language Processing). Cambridge: Cambridge University Press.
- Berner Zeitung. 2009. *Unterwegs ins Weisse Haus*. [website]. <https://www.bernerzeitung.ch/unterwegs-ins-weisse-haus-387231656707> (20 March, 2023).
- Biber, Douglas. 1994. An analytical framework for register studies. In Douglas Biber & Edward Finegan (eds.), *Sociolinguistic Perspectives on Register*, 31–56. Oxford: Oxford University Press.
- Biber, Douglas & Susan Conrad. 2019. *Register, Genre, and Style*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/9781108686136>.
- Biber, Douglas, Jesse Egbert, Daniel Keller & Stacey Wizner. 2021. Towards a taxonomy of conversational discourse types: An empirical corpus-based analysis. *Journal of Pragmatics* 171. 20–35. <https://doi.org/10.1016/j.pragma.2020.09.018>.
- Brunner, Annelen, Lukas Weimer, Stefan Engelberg, Fotis Jannidis & Ngoc Duyen Tanja Tu. 2020. *Annotationsrichtlinien des Projekts “Redewiedergabe. Eine literatur- und sprachwissenschaftliche Korpusanalyse”*. Version 1.2. <https://doi.org/10.5281/zenodo.3759617>.
- Davidson, Donald. 1967. The logical form of action sentences. In *Essays on Actions and Events*, 105–122. Oxford: Clarendon Press.
- Deutscher Bundestag. 2023. *Parlamentscorpus aus Berichten des Deutschen Bundestages der 16. und 17. Legislaturperiode*. [online resource]. <https://www.bundestag.de/protokolle> (4 February, 2023).

- Dipper, Stefanie, Anke Lüdeling & Marc Reznicek. 2013. NoSta-D: A corpus of German non-standard varieties. In Marcos Zampieri & Sascha Diwersy (eds.), *Non-standard Data Sources in Corpus-based Research*, 69–76. Aachen: Shaker Verlag.
- Eckardt, Regine. 2015. Speaker and Narrator. In Dorothee Birke & Tilmann Köppe (eds.), *Author and Narrator: Transdisciplinary Contributions to a Narratological Debate*, 153–186. Berlin, München, Boston: De Gruyter. <https://doi.org/10.1515/9783110348552.153>.
- Egbert, Jesse & Michaela Mahlberg. 2020. Fiction – one register or two? *Register Studies* 2(1). 72–101. <https://doi.org/10.1075/rs.19006.egb>.
- Fludernik, Monika. 1991. The historical present tense yet again: tense switching and narrative dynamics in oral and quasi-oral storytelling. *Text* (11). 365–398.
- Genette, Gérard. 2010. *Die Erzählung*. Paderborn: Fink. <https://doi.org/10.36198/9783838580838>.
- Gerstner-Link, Claudia & Manfred Krifka. 1993. Genericity. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann (eds.), *An International Handbook of Contemporary Research, 1. Halbband*, 966–978. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110095869.1.14.966>.
- Grabes, Herbert. 2014. Sequentiality. In Jan Christoph Meister (ed.), *The Living Handbook of Narratology*, 1–12. Hamburg: Universität Hamburg. <https://www.lhn.uni-hamburg.de/node/91.html> (23 April, 2023).
- Hühn, Peter. 2013. Event and eventfulness. In Jan Christoph Meister (ed.), *The Living Handbook of Narratology*, 1–36. Hamburg: Universität Hamburg. <https://www.lhn.uni-hamburg.de/node/39.html> (23 April, 2023).
- Köppe, Tilmann & Jan Stühling. 2011. Against pan-narrator theories. *jlse* 40(1). 59–80. <https://doi.org/10.1515/jlse.2011.004>.
- Labov, William & Joshua Waletzki. 1967. Narrative analysis. In June Helm (ed.), *Essays on the Verbal and Visual Arts*, 12–44. Seattle, London: University of Washington Press.
- Lahn, Silke & Jan Christoph Meister. 2016. *Einführung in die Erzähltextanalyse*. Berlin: Springer-Verlag. <https://doi.org/10.1007/978-3-476-05415-9>.
- Lee, David. 2002. Genres, Registers, Text Types, Domains and Styles: Clarifying the Concepts and Navigating a Path through the BNC Jungle. 245–292. https://doi.org/10.1163/9789004334236_021.
- Leibniz-Institut für Deutsche Sprache. 2018. Das Präsens. [dynamic online resource]. In “*Propädeutische Grammatik*”. *Grammatisches Informationssystem grammis*. Mannheim: Leibniz-Institut für Deutsche Sprache. <https://doi.org/10.14618/programm>.

- Lüdeling, Anke, Artemis Alexiadou, Aria Adli, Karin Donhauser, Malte Dreyer, Markus Egg, Anna Helene Feulner, Natalia Gagarina, Wolfgang Hock, Stefanie Jannedy, Frank Kammerzell, Pia Knoeferle, Thomas Krause, Manfred Krifka, Silvia Kutscher, Beate Lütke, Thomas McFadden, Roland Meyer, Christine Mooshammer, Stefan Müller, Katja Maquate, Muriel Norde, Uli Sauerland, Stephanie Solt, Luka Szucsich, Elisabeth Verhoeven, Richard Waltereit, Anne Wolfgruber & Lars Erik Zeige. 2022. Register: Language Users' Knowledge of Situational-Functional Variation: Frame text of the First Phase Proposal for the CRC 1412. *Register Aspects of Language in Situation* 1(1). 1–59. <https://doi.org/10.18452/24901>.
- Lüdeling, Anke, Seanna Doolittle, Hagen Hirschmann, Karin Schmidt & Maik Walter. 2008. Das Lernerkorpus Falko. In *Deutsch als Fremdsprache*, vol. 2, 67–73. Berlin: Erich Schmidt Verlag. <https://doi.org/10.37307/j.2198-2430.2008.02>.
- Maienborn, Claudia. 2019. Events and states. In Robert Truswell (ed.), *Oxford Handbook of Event Structure*, 50–89. Oxford: Oxford University Press.
- Martínez, Matias & Michael Scheffel. 2009. *Einführung in die Erzähltheorie*. München: C.H. Beck.
- Martínez, Matías. 2011. Theorie der erzählenden Literatur, I. Grundbestimmungen, 1. Erzählen. In Matías Martínez (ed.), *Handbuch Erzählliteratur*, 1–12. Berlin: Springer.
- McHale, Brian. 2019. Speech Representation. In John Pier Hühn, Wolf Schmid & Jörg Schönert (eds.), *The living handbook of narratology*. Hamburg: Hamburg University. <http://www.lhn.uni-hamburg.de/article/speech-representation> (24 April, 2023).
- Mehler, Alexander. 2008. Corpus compilation and corpus types: Collection strategies and design decisions. In Anke Lüdeling & Merja Kytö (eds.), *Corpus Linguistics*, vol. 1, 154–382. Berlin, New York: De Gruyter Mouton. <https://doi.org/10.1515/9783110211429.1.154>.
- Neumann, Stella. 2014. *Contrastive Register Variation: A Quantitative Approach to the Comparison of English and German*. Berlin: De Gruyter Mouton.
- Parsons, Terence. 1990. *Events in the Semantics of English*. Cambridge: MIT Press.
- Quasthoff, Uta, Vivien Heller & Miriam Morek. 2017. On the sequential organization and genre-orientation of discourse units in interaction: An analytic framework. *Discourse Studies* 19(1). 84–110. <https://doi.org/10.1177/1461445616683596>.
- Saurí, Roser, Jessica Littman, Bob Knippen, Robert Gaizauskas, Andrea Setzer & James Pustejovsky. 2006. TimeML annotation guidelines. *Version* 1(1). 31.
- Scheffel, Michael, Antonius Weixler & Lukas Werner. 2014. Time. In Jan Christoph Meister (ed.), *The Living Handbook of Narratology*, 1–20. Hamburg:

- Universität Hamburg. <https://www.lhn.uni-hamburg.de/node/106.html> (15 February, 2023).
- Spronck, Stef & Tatiana Nikitina. 2019. Reported speech forms a dedicated syntactic domain. *Linguistic Typology* 23(1). 119–159. <https://doi.org/10.1515/lingty-2019-0005>.
- Stanzel, Franz Karl. 1955. *Die typischen erzählsituationen im roman: dargestellt an tom jones, moby-dick, the ambassadors, ulysses, u. a.* Vol. 63 (Wiener Beiträge zur englischen Philologie). Wien, Stuttgart: W. Braumüller.
- Stede, Manfred. 2018. *Korpusgestützte Textanalyse*. 2nd edn. Tübingen: Narr Francke Attempto Verlag.
- Vendler, Zeno. 1967. Verbs and times. In *Linguistics in Philosophy*, 97–121. Ithaca, New York: Cornell University Press.
- Walter, Maik. 2013. *Maerchenkorpus*. [online resource]. https://korpling.german.hu-berlin.de/annis3/#_c=TWFlcmNoZW5rb3JwdXM (20 March, 2023).
- Zeman, Sonja. 2016. Introduction: Perspectives on narrativity and narrative perspectivization. In Natalia Igl & Sonja Zeman (eds.), *Perspectives on narrativity and narrative perspectivization* (Linguistic Approaches to Literature), 17–42. Amsterdam: John Benjamins.
- Zeman, Sonja. 2018. What is a narration – and why does it matter. In Markus Steinbach & Annika Hübl (eds.), *Linguistic Foundations of Narration in Spoken and Sign Language* (Linguistics Today 247), 173–206. Amsterdam, Philadelphia: John Benjamins.
- Zeman, Sonja. 2020a. Grammatik der Narration. *Zeitschrift für germanistische Linguistik* 48(3). 457–494.
- Zeman, Sonja. 2020b. Narrativität als linguistische Kategorie. *Zeitschrift für germanistische Linguistik* 48(3). 447–456.
- Zeman, Sonja. 2020c. Parameters of Narrative Perspectivization: The Narrator. *Open Library of Humanities* 6(2). <https://doi.org/10.16995/olh.502>.

Appendix

Table 7: Annotation Example: Politics³⁴

| tok | S | MEDIACY | MOTIVE | NARRATOR | PART | EVENT | SEQ | NAR.ID | F.ENT-L1 | F.FCT-L1 | F.ILL-L1 | F.SOC-L1 |
|--------------------|----|-------------|--------|--------------|----------|-------|-----|--------|----------|----------|----------|----------|
| Deswegen | s1 | non-mediate | none | none | | | | | | | | |
| dieses | | | | | | | | | | | | |
| zweistufige | | | | | | | | | | | | |
| Verfahren. | s2 | non-mediate | none | none | | | | | | | | |
| Wir | | | | | | | | | | | | |
| sind | | | | | | | | | | | | |
| ja | | | | | | | | | | | | |
| in | | | | | | | | | | | | |
| Deutschland | | | | | | | | | | | | |
| noch | s3 | non-mediate | none | none | | | | | | | | |
| Lernende, | | | | | | | | | | | | |
| was | | | | | | | | | | | | |
| die | s4 | non-mediate | none | none | | | | | | | | |
| Privatfinanzierung | | | | | | | | | | | | |
| von | | | | | | | | | | | | |
| Verkehrsprojekten | s5 | mediate | Past | narrator-001 | non-part | e-001 | | | | | | |
| anbetrifft. | | | | | | | | | | | | |
| Die | | | | | | | | | | | | |
| Geschichte | | | | | | | | | | | | |
| dieses | | | | | | | | | | | | |
| Tunnelprojektes | | | | | | | | | | | | |
| reicht | s6 | non-mediate | none | none | | | | | | | | |
| bis | | | | | | | | | | | | |
| 1987 | | | | | | | | | | | | |
| zurück; | s7 | non-mediate | none | none | | | | | | | | |
| damals | | | | | | | | | | | | |
| hat | | | | | | | | | | | | |
| man | s8 | non-mediate | none | none | | | | | | | | |
| mit | | | | | | | | | | | | |
| der | | | | | | | | | | | | |
| Planung | s9 | non-mediate | none | none | | | | | | | | |
| begonnen. | | | | | | | | | | | | |
| Deswegen | | | | | | | | | | | | |
| ist | s6 | non-mediate | none | none | | | | | | | | |
| der | | | | | | | | | | | | |
| Weg, | | | | | | | | | | | | |
| den | s7 | non-mediate | none | none | | | | | | | | |
| wir | | | | | | | | | | | | |
| jetzt | | | | | | | | | | | | |
| vorschlagen, | s6 | non-mediate | none | none | | | | | | | | |
| vernünftig; | | | | | | | | | | | | |
| Wir | | | | | | | | | | | | |
| wollen | s8 | non-mediate | none | none | | | | | | | | |
| schauen, | | | | | | | | | | | | |
| ob | | | | | | | | | | | | |
| das | s9 | non-mediate | none | none | | | | | | | | |
| Projekt | | | | | | | | | | | | |
| als | | | | | | | | | | | | |
| F-Modell | | | | | | | | | | | | |
| geeignet | | | | | | | | | | | | |
| ist. | | | | | | | | | | | | |

³⁴Parlamentsredenkorpus: <https://korpling.german.hu-berlin.de/annis3/?id=b65eb73a-0cb7-4077-a82d-68df88387154>

³⁵Maerchenkorpus (Walter 2013): <https://korpling.german.hu-berlin.de/annis3/?id=dadc130b-9579-4629-9c4a-0c1284d6a981>

³⁶Biber et al. 2021: 29 (DU 3.2)

³⁷<https://www.chefkoch.de/rezepte/1736871282723243/Salat-aus-roten-Linsen.html>

Table 8: Annotation Example: Fairy tale³⁵

| tok | S | MEDIACY | MOTIVE | NARRATOR | PART | EVENT | SEQ | NAR.ID | F.ENT- L1 | F.FCT- L1 | F.FILL- L1 | F.SOC- L1 | F.ENT- L2 | F.FCT- L2 | F.FILL- L2 | F.SOC- L2 |
|---|----|---------|------------------|---------------------|--------------|-------|---------------------------------------|----------------|----------------|--------------|---------------|--------------|--------------|--------------|---------------|--------------|
| Als sie die Bohnen in den Topf schüttete , | s1 | mediate | Past | narrator- 001 | non- part | | | nar- 001 | enter- tain | | | | | | | |
| entfiel ihr unbemerkt eine , | s2 | mediate | Past | narrator- 001 | non- part | e-001 | e-001 > e-002 | nar- 001 | enter- tain | | | | | | | |
| die auf dem Boden neben einen Strohalm zu liegen kam; | s3 | mediate | Past | narrator- 001 | non- part | | | nar- 001 | enter- tain | | | | | | | |
| bald danach sprang auch eine glühende Kohle vom Herd zu den beiden herab . | s4 | mediate | Past | narrator- 001 | non- part | e-003 | e-003 > e-004 | nar- 001 | enter- tain | | | | | | | |
| Da fieng der Strohalm an und sprach | s5 | mediate | Past | narrator- 001 | non- part | e-004 | e-004 > e-005, e-004 c e-008 | nar- 001 | enter- tain | | | | | | | |
| Die Freunde , von wannen kommt ihr her ? | s6 | mediate | Past | narrator- 001 | non- part | e-005 | e-005 c e-006 | nar- 001 | enter- tain | | | | | | | |
| Die Kohle antwortete „ich bin zu gutem Glück dem Feuer entsprungen | s7 | mediate | direct speech | narrator- 001-01 | part | e-006 | e-006 > e-007 | nar- 001 | enter- tain | | | | | | | |
| Die Kohle antwortete „ich bin zu gutem Glück dem Feuer entsprungen | s8 | mediate | Past | narrator- 001 | non- part | | | nar- 001 | enter- tain | | | | | | | |
| Die Freunde , von wannen kommt ihr her ? | s9 | mediate | direct speech | narrator- 001-02 | part | e-007 | | nar- 001-01 | enter- tain | | | | | | | so- cial |

Guidelines for the Annotation of Parameters of Narration

Table 9: Annotation example: Casual conversation³⁶

| tok | S | MEDIACY | MOTIVE | NARRATOR | PART | EVENT | SEQ | NAR.ID | F.ENT-L1 | F.FCT-L1 | F.ILL-L1 | F.SOC-L1 |
|---|-----|-------------|--------|--------------|------|-------|---------------------------------|---------|----------|----------|----------|----------|
| A: I used to enjoy playing that | s1 | mediate | Past | narrator-001 | part | e-001 | e-001 c e-002, e-001 c e-004 | nar-001 | | | | social |
| B: did you win? | s2 | mediate | Past | narrator-002 | part | | | | | | | |
| A: yeah | s3 | non-mediate | none | none | | | | | | | | |
| I was like first in the world | s4 | mediate | Past | narrator-001 | part | e-002 | e-002 >e-003 | nar-001 | | | | social |
| B: were you? | s5 | mediate | Past | narrator-002 | part | | | | | | | |
| A: but probably nobody was playing like | s6 | mediate | Past | narrator-001 | part | | | | | | | |
| B: oh somebody 's took your crown | s7 | mediate | Past | narrator-002 | part | e-003 | | nar-001 | | | | social |
| now you 'd better go back on | s8 | non-mediate | none | none | | | | | | | | |
| A: probably yeah | s9 | non-mediate | none | none | | | | | | | | |
| I was like first in the world for like that month or so | s10 | mediate | Past | narrator-001 | part | e-004 | | nar-001 | | | | social |
| I don't know if I was ever first in the world world | s11 | non-mediate | none | none | | | | | | | | |

Table 10: Annotation example: Recipe³⁷

| tok | S | MEDIACY | MOTIVE | NARRATOR | PART | EVENT | SEQ | NAR.ID | F.ENT-L1 | F.FCT-L1 | F.ILL-L1 | F.SOC-L1 |
|-------------|-----|-------------|--------|--------------|------|-------|---------------|---------|----------|----------|----------|----------|
| Die | s1 | non-mediate | none | none | | | | | | | | |
| Linsen | | | | | | | | | | | | |
| in | | | | | | | | | | | | |
| der | | | | | | | | | | | | |
| Gemüsebrühe | | | | | | | | | | | | |
| ca. | | | | | | | | | | | | |
| 8 | | | | | | | | | | | | |
| Minuten | | | | | | | | | | | | |
| kochen | s2 | non-mediate | none | none | | | | | | | | |
| und | | | | | | | | | | | | |
| dann | s3 | non-mediate | none | none | | | | | | | | |
| abgießen. | | | | | | | | | | | | |
| Auf | | | | | | | | | | | | |
| der | | | | | | | | | | | | |
| Packung | | | | | | | | | | | | |
| steht | | | | | | | | | | | | |
| bei | | | | | | | | | | | | |
| mir | | | | | | | | | | | | |
| 12 | | | | | | | | | | | | |
| Minuten | | | | | | | | | | | | |
| - | s4 | non-mediate | none | none | | | | | | | | |
| sie | | | | | | | | | | | | |
| sollten | | | | | | | | | | | | |
| aber | | | | | | | | | | | | |
| schon | | | | | | | | | | | | |
| noch | s5 | non-mediate | none | none | | | | | | | | |
| Biss | | | | | | | | | | | | |
| haben! | s6 | non-mediate | none | none | | | | | | | | |
| Einfach | | | | | | | | | | | | |
| probieren. | | | | | | | | | | | | |
| Außerdem | | | | | | | | | | | | |
| steht | | | | | | | | | | | | |
| auf | | | | | | | | | | | | |
| meiner | | | | | | | | | | | | |
| Packung. | | | | | | | | | | | | |
| man | | | | | | | | | | | | |
| solle | | | | | | | | | | | | |
| die | | | | | | | | | | | | |
| Linsen | | | | | | | | | | | | |
| vorher | | | | | | | | | | | | |
| kalt | | | | | | | | | | | | |
| abspülen. | | | | | | | | | | | | |
| Dadurch | s7 | mediate | Past | narrator-001 | part | e-001 | e-001 > e-002 | nar-001 | | fact | | |
| wurden | | | | | | | | | | | | |
| sie | | | | | | | | | | | | |
| aber | | | | | | | | | | | | |
| beim | | | | | | | | | | | | |
| ersten | | | | | | | | | | | | |
| Mal | | | | | | | | | | | | |
| sehr | | | | | | | | | | | | |
| matschig. | | | | | | | | | | | | |
| ich | s8 | mediate | Past | narrator-001 | part | | | nar-001 | | fact | | |
| hab | | | | | | | | | | | | |
| es | | | | | | | | | | | | |
| dann | | | | | | | | | | | | |
| nicht | | | | | | | | | | | | |
| mehr | | | | | | | | | | | | |
| gemacht | | | | | | | | | | | | |
| und | s9 | non-mediate | none | none | | e-002 | | | | | | |
| und | | | | | | | | | | | | |
| fand | s10 | non-mediate | none | none | | | | | | | | |
| das | | | | | | | | | | | | |
| besser. | s11 | non-mediate | none | none | | | | | | | | |
| den | | | | | | | | | | | | |
| knolauch | | | | | | | | | | | | |
| schälen | | | | | | | | | | | | |
| und | | | | | | | | | | | | |
| fein | | | | | | | | | | | | |
| hacken. | | | | | | | | | | | | |