Theory of mind in language, minds, and machines: a multidisciplinary approach

Dijk, B.M.A. van

Chapter 5

Character Perspective Representation in Freely Told Stories

Story characters not only perform actions, they typically also perceive, feel, think, and communicate. Here we are interested in how children render characters' perspectives when freely telling a fantasy story. Drawing on a sample of 150 narratives elicited from Dutch children aged 4-12y, we provide an inventory of 750 instances of Character Perspective Representation (CPR), distinguishing fourteen different types. We observe first of all that character perspectives are ubiquitous in freely told children's stories and take more varied forms than traditional frameworks can accommodate. Second, we discuss variation in the use of different types of CPR across age groups, finding that character perspectives are being fleshed out in more advanced and diverse ways as children grow older. Thirdly, we explore whether such variation can be meaningfully linked to automatically extracted linguistic features, thereby probing the potential for using automated tools from Natural Language Processing to extract and classify character perspectives in children's stories.

This work was originally published as: Van Duijn, M.J., Van Dijk, B.M.A., and Spruit, M.R. (2022). Looking from the Inside: How Children Render Characters' Perspectives in Freely-told Fantasy Stories. In Clark, E., Brahman F., and Iyyer, M., editors, *Proceedings of the 4th Workshop on Narrative Understanding*, pages 66-76. Association for Computational Linguistics.

5.1 Introduction

Story characters are everywhere around us: we meet them in the books we read, the TV series we get caught up in, or in a gossipy tale we tell each other during everyday social gatherings. Some characters may be modelled on real people, whereas others exist only in the imagined worlds of fantasy and fiction.

In its most basic form, a story character is an entity involved in some kind of action or description. Yet typically we also get to share in some of its *perspectives* on the story world and the objects, events, and other characters within it. There are long-standing traditions in linguistics and literary studies, especially within the subfields of stylistics and narratology, studying the ways in which such character perspectives can be rendered (e.g. Banfield, 1973; Leech and Short, 2007; Vandelanotte, 2009). Three main types commonly distinguished in studies of speech and thought representation are *direct*, *indirect*, and *free (in)direct* speech or thought (see Table 5.2 for examples). While most attention has been paid to literary texts, scholars have also identified such types in cinema (Verstraten, 2009), theatre (McConachie and Hart, 2006), and other domains such as news articles (Sanders, 2010), everyday conversations between parents and young infants (Köder, 2016), or speech from individuals with psycho-pathological conditions (Van Schuppen et al., 2020).

It is largely an open question as of yet how children render characters' intentions, perceptions, emotions, speech, and thought when asked to freely tell a fantasy story. This is worthwhile exploring for a variety of reasons. It has been widely argued that representing different perspectives reflects a central function of language usage (e.g. Dancygier et al., 2016): human interaction is characterised by 'polyphony', meaning that we rarely only express our own perspective. Instead, the default is that we use language to *orchestrate* multiple perspectives.

Even though this pervades all speech domains, stories are a key finding place for linguistic and narratological patterns supporting this function (Fludernik, 1996), and arguably also the 'sandbox' where both children and adults test and refine their perspective orchestration skills (Vermeule, 2009). Mapping how children of different ages render character perspectives is as such of interest to language acquisition research, but also to cognitive psychology as it provides insight into how children learn to understand the social world and others' minds, and the role narratives can play herein. Tools from Natural Language Processing (NLP) can fuel all such research, for example by automatically identifying contextual information associated with different character perspectives. NLP researchers, in turn, can learn about phe-

nomena relevant for embarking on tasks involving more complex classification or extraction of perspectivised content.

In the current contribution we draw on a sample of 150 stories, told by children aged 4-12y as part of storytelling workshops we offered across The Netherlands. Our sample features 750 instances of Character Perspective Representation (CPR), which we categorise into fourteen different types based on manual qualitative analysis. As discussed below, the type categories and analytical framework we adopt are primarily inspired on 'classic' speech and thought representation literature (mainly Leech and Short, 2007). However, we complement our framework with additional types based on research into children's development as storytellers and relevant insights from cognitive linguistics, allowing for a more refined and inclusive way of mapping character perspectives.

The best way to introduce our approach in concrete terms is to discuss the analysis of an example story (given in Table 5.1). Doing so will also make clear how we position this chapter: as an effort to build a bridge between *qualitative* analysis of narrative material as traditionally done in the humanities, and *quantitative* analysis, driven by the automatic extraction of linguistic information, as customary in computational approaches. However, in Section 5.3 below, we will first provide more details on our language sample and annotations, introduce two automatically extracted linguistic features, and then discuss an example story (given in Table 5.1) and our CPR typology. Thereafter we discuss in Section 5.4 the relation between the occurrence of different CPR types and the age of the storytellers, as well as the relation with lexical and syntactic characteristics of the utterances in which the CPR types occur. We end with a reflection on our findings in Section 5.5.

5.2 Background

Children tell stories to themselves and others as part of their daily play activities (Cremin et al., 2017; Sutton-Smith, 1986). While being the source of a lot of fun in the first place, such storytelling has been analysed as a form of *cognitive play* that is essential for child development in various key areas, including the acquisition and refinement of communicative skills (Southwood and Russell, 2004), organising knowledge of the (social) world (McKeough and Genereux, 2003), and empathising with others and understanding their motives and intentions (Gallagher and Hutto, 2008; Nicolopoulou, 2018; Zunshine, 2019). Phenomena of CPR are situated at a natural crossroads of these key developmental areas: their absence or presence in freely

told stories arguably reflects children's communicative abilities, but also their understanding of the social world and capacity to imagine others' inner workings. Here we explore the occurrence of different patterns of CPR across different age groups, and we believe that our contribution can ultimately fuel research in developmental psychology and language acquisition research. However, it is important to note that claims about whether the patterns we find in our stories are indicative of a specific child's development are outside the scope of this chapter.

5.3 Methods

Data

The storytelling workshops for the creation of our database were held between 2019-2021 at seven elementary schools, a daycare, and a community centre located in various areas across in The Netherlands. Each session was held in a classroom setting involving 5-30 children at a time, of varying ages between 4-12. Sessions started by discussing some general characteristics of stories (e.g. 'Where can you find stories?', 'What kind of stories do you like?') and interactively narrating an exemplary fantasy story with the participating children. Next, we invited children to take the floor and tell a fantasy story about a topic free of choice. After informing children about this, voice recordings were made, which were pseudonymised and transcribed afterwards by the authors and research assistants. Transcripts were double-checked for consistency with the audio files.

As of now, we have collected over 600 stories in our database called ChiSCor (Children's Story Corpus).[1] Our data collection and data management protocols were assessed and approved by the Leiden University Faculty of Science Ethics Committee (file no. 2020 – 002).

For the current research we drew a sample from ChiSCor according to the following steps:

1. We included only the first story told by each child (many children told multiple stories), which reduces dependence between stories. This yielded a subset of 350 stories.

2. We selected stories with a length (in number of words, $\bar{x} = 108.64, \sigma = 99.62$) falling in the interquartile range (IQR), i.e. 50% around the median ($min =$

[1] The stories used in this chapter, along with our current annotations and scripts, are available via the Open Science Framework (OSF): `https://osf.io/9q32v/`

$4, Q1 = 35, Med = 75, Q3 = 151, max = 626$), to prevent over- or underrepresentation of data from children with exceptionally long or short stories.

3. We then defined three age categories, 'Young', 'Middle' and 'Older', in line with the division common in Dutch primary education into 'Onderbouw', 'Middenbouw', and 'Bovenbouw'. Young corresponds to Onderbouw which involves ages 4-6; Middle corresponds to Middenbouw which involves ages 6-9; Older corresponds to Bovenbouw which involves ages 9-12.

4. We included 150 stories in total (12879 words), 50 for each group. For the Young and Middle groups these were randomly drawn out of 60 and 78 stories falling within the IQR, respectively. The older group had only 39 stories within the IQR; here we added 11 stories closest to $Q1$ and $Q3$ to balance the age groups.

Annotations

The 150 stories were put into a large table in random order and without showing additional information to avoid (unconscious) interference with decisions in the annotation process.[2] Existing line breaks, introduced during transcription of the audio recordings according to a standardised protocol, were used to chunk each story into smaller units, henceforth referred to as 'utterances'. We identified 568 unique characters that in total made 1472 appearances (the same character can obviously appear in multiple utterances within the same story), 722 of which involved only descriptions or simple actions without insight being offered into the character's perspective. The remaining 750 appearances were given one of fourteen different labels representing our types of CPR. In rare cases where multiple types applied, the most 'advanced' label was chosen in terms of the stages introduced below.

One author of this chapter who has a background in grammar and narratological theory, took the lead in the annotation process, while regularly discussing categorical distinctions as well as individual utterances with the second author. In some specific cases, expertise was gathered from external experts. While we can see how this procedure may be problematic from the perspective of current standards in NLP, two considerations should be added with regard to our approach in this chapter. Firstly, we point out that we base our annotations on long-standing traditions of textual analysis within cognitive linguistics, narratology, and stylistics, known to support high degrees of intersubjective agreement and reproducibility between researchers within

[2] E.g. the age or school of the storyteller. Note that such interference could only be avoided to a certain degree; after all, we were ourselves involved in recording the stories.

these fields (for a broader discussion of a 'grounded theory' approach, see Charmaz, 2006). Secondly, it is important to note that the statistical analyses in Section 5.4 are based on *merged* categories. While discussion is sometimes possible about the most appropriate type of label for specific utterances (e.g. deciding between direct and indirect speech on grammatical grounds; see also Köder, 2016), such discussions would rarely affect the overarching merged category under which this utterance falls.[3] Nevertheless, we consider it an important next step within our larger project to gather CPR annotations from at least one additional, independent annotator.

We discuss our full typology of CPR further below, along with the example story and inventory of the occurrence of each type in our sample. However, it is important to single out one type beforehand: `ego-narration`. We see this as a 'preliminary stage' of the fuller mastery of CPR that is characteristic of the other thirteen types. We marked cases as `ego-narration` if there was no (or an unclear) distinction between the child narrating the story and a referent indicated with first-person pronouns ('I', 'me', 'we', 'us') within a story.

Consider the following example from story with ID 022501 in ChiSCor:

(1) [...] *and I do a lot of horse riding / and ride a lot of horses / and we have a lot of very sweet horses in the stables* [...]

Example (1) counts as `ego-narration`, since the 'I' who regularly does a lot of horse riding refers to the child in the immediate situation of telling the story. This is different in the following example from story 082601:

(2) [...] *and then came well myself in fact who came with a gun / and I said why are you fighting Batman and Superman* [...]

In example (2), the 'I' is making an appearance in a story world clearly detached from the here-and-now of telling the story.[4]

The rationale for singling out `ego-narration` as a preliminary phenomenon is that it evidences a lack of 'transcendence' (Zeman, 2020), marking a departure from the actual speaker and its immediate here-and-now, which we consider a key feature of storytelling. Such transcendence is warranted by a distinction between the child telling the story (ego), the narrator seen as a theoretical entity or 'role', and characters

[3] An exception is found in line 7 of the example story presented in Table 5.1.

[4] The full Dutch stories can be found in our OSF repository (footnote 1). Utterances are separated with forward slashes. English translations are our own and were made only for the purpose of discussing them here; CPR annotations within this chapter are based on the Dutch stories.

within the story.[5]

What the remaining thirteen types of character representation have in common is that they exhibit storytelling in this sense, i.e. a specific form of communication in which a narrator-entity provides all kinds of linguistic cues inviting listeners (or readers) to imagine a story world including objects, characters, actions, events, etc. (Dancygier, 2011). In this way it is possible for narrators to tell a story entirely from the 'outside perspective', without directly cuing listeners to imagine what the story world would look like from any character's point of view; this is what we observed in utterances containing only character appearances consisting of descriptions or simple actions, plus in utterances containing no character appearances at all. In each of the remaining utterances we found essentially a mix of narrator and character perspectives. The way in which, and degree to which these character perspectives were explicitly fleshed out and/or separated from that of the narrator, determine which of the thirteen CPR types applies.

Linguistic features

There is evidence that socio-cognitive skills, in particular the capacity to understand and reason about others' mental states known as Theory of Mind (Apperly, 2012), are positively correlated to lexical and syntactic proficiency in children. For example, children possessing a larger vocabulary, or mastering clausal complementation, perform better in reasoning about others' mental states in standardised clinical tasks (for an overview see Milligan et al., 2007).

We see overlap between children's development of socio-cognitive capacities and their ability to flesh out characters' perspectives in a narrative. Therefore we include lexical and syntactic complexity here as two theoretically motivated features, that can potentially provide us with information about the linguistic context in which different CPR types occur, and connect this to age groups of the storytellers in our sample. Doing so, we might also anticipate linguistic information encoded in (the middle layers of deeper) neural networks, that could be helpful for automatically extracting and/or classifying perspectivised information in children's narratives in the future (Jawahar et al., 2019).

To calculate Lexical Complexity (LC), we approximated for each word in utterance U featuring CPR its lemma probability $P(l)$ by its relative frequency count in the BasiScript lexicon, a large benchmark corpus of written child output (Tellings et al.,

[5]We refrain from going into the widely debated narratological concept of the narrator here and refer to Zeman (2020) for a to-the-point overview.

	Utterance
1.	*a girl went to the zoo and she saw a huge lot of tigers and other animals* [...]
2.	*and she went home all alone*
3.	*but her little brother was left behind he was sitting on the monkey*
4.	*then said the sister of the little boy where is my little brother now*
5.	*she went back again to the zoo*
6.	*then she saw that the little brother was sitting on the monkey*
7.	*oh little brother where are you now*
8.	*the end*

Table 5.1: Example story with ID 072201.

2018a). The perplexity of the utterance $PP(U)$ is then given by the set of lemmas $U = \{l_1, l_2...l_N\}$ and its probabilities with

$$PP(U) = \sqrt[N]{\frac{1}{P(l_1, l_2, ...l_N)}}. \tag{5.1}$$

Utterances with more infrequent lemmas show higher perplexity with respect to the lexicon. Lemma frequency has been argued to be a good measure of lemma complexity given that infrequent lemmas are overall harder to learn (Vermeer, 2001).

To calculate Syntactic Complexity (SC), for each utterance U featuring CPR we extracted a dependency tree, a directed graph $G = (V, A)$ with V as the set of words and A as the set of arcs indicating dependency relations between words. We extracted the maximum number of arcs between the root node and a leaf node in G. This measure is also known as tree depth and is a common measure of syntactic complexity: utterances employing longer paths are syntactically more complex (Dell'Orletta et al., 2011).

CPR types in an example story

In order to illustrate our approach in more detail, we will now discuss the analysis of a story excerpt given in Table 5.1, featuring five types of CPR found throughout our sample. Afterwards, the remaining types will be briefly introduced along with a complete overview of examples and counts in Table 5.2.

First of all, we can observe that this is a story narrated in third person, past tense. For a large part it consists of narrator descriptions of actions and situations ('went to the zoo', utterance 1; 'went home all alone', utterance 2; 'her little brother was left behind' and 'sitting on a monkey', utterance 3; etc.); however, as listeners/readers

we also get a few glimpses into the perspective of one character: the 'girl'.

In utterance 1 we learn about the animals she 'saw'. It could be defended that this is still entirely the narrator's voice telling us 'from the outside' what the girl would have been seeing at the zoo. Yet, as discussed in Section 5.2 above, and in line with what cognitive linguists have argued in recent years (e.g. Van Duijn and Verhagen, 2018), we suggest that perspectivisation of content in narratives can be seen on a cline, ranging from pure narrator view on the one extreme, to full character view with minimal narrator mediation on the other extreme. Following this approach, the report of what the girl 'saw' in utterance 1 implies a modest but certain invitation for listeners or readers of the story to imagine the girl's perspective on objects within the story world: 'a huge lot of tigers and other animals'. This is a case of Perception (PER) in our system of types. Another instance is found in utterance 6.

What is more, we note a difference between how the situation of the 'little brother' is described ('was left behind', 'sitting on a monkey', utterance 3) and some of the descriptions of actions performed by the 'girl' (e.g. 'went home all alone', utterance 2; 'went back again to the zoo', utterance 5). Following developmental psychologists and children's story researchers Nicolopoulou and Richner (2007) we classify the latter as cases of Intention-in-action (IIA), i.e. actions coupled to a clear goal or result within the immediate story context. As these authors argue, such actions are not yet fully explicit intentional states that would disclose characters' perspectives for the audience, but hint at them implicitly. So, compared to PER and other forms of CPR discussed below, IIA represents the lowest degree of inviting a shift from the narrator's to a character's perspective. Yet mere descriptions of a character's situation, appearance, attributes, or actions without an immediately specified result or goal do not invite such a shift at all, or to an even lesser degree. This is why we see IIA as the most basic type in our staging of CPR.

In utterance 4 we find a case of Direct Speech (DS) with an inquit formula ('said the sister of the little boy') and a reported clause ('where is my little brother now'.[6]) The reported clause has three features supporting our classification as DS. Firstly, a shift to the present tense can be observed ('is' as opposed to 'said' in the inquit formula). Secondly, there is a shift from the third to the first person as expressed by the pronoun 'my'. And thirdly, the addition of 'now' ('nou' in the original Dutch story) can be seen as an idiomatic exclamation, expressing a degree of wonder or confusion (which is not satisfactorily covered by the English translation 'now'). This wonder or confusion

[6]The absence of a question mark after the reported clause is due to standardised transcription of the recorded oral stories.

is clearly to be interpreted as part of the 'girl'-character's experience, and not of the narrator's, just as 'my little brother' from the character's point of view indicates the same referent as 'the little boy' from the narrator's point of view. The present tense is congruent with the girl-character's experience at the moment of speaking within the story plot.

Finally, utterance 7 features Free Direct Speech (FDS). Here we see the same shift to present tense ('are') and the same exclamation ('nou' in the Dutch original), complemented with another exclamation at the beginning of the sentence ('oh'). The absence of an inquit formula makes it a case of FDS rather than DS. Or, a different possible interpretation of utterance 7 is that we are looking at a form of 'monologue intérieur' in which the girl-character produces this utterance for herself, rendering it a case of Free Direct Thought (FDT) rather than speech. The context does not resolve this ambiguity. One can argue that she is addressing the boy, given that she has just found him in the preceding utterance, but one can equally well argue that utterance 7 should be read as an internal expression of her surprise, given that he is sitting on a monkey.

CPR types in the rest of the stories

In Table 5.2 it can be seen that ego-narration (EGO-NARR), the preliminary stage of CPR we distinguished earlier, occurs 47 times in our sample. IIA, which we consider to be CPR in its most basic form, is with 350 occurrences by far the most frequently observed type. Usage of IIA entails that the narrator reports what a character is doing, and to what end. Similarly, with PER, of which we recorded 53 instances, it is the narrator who reports what a character is perceiving. Both happen without the narrator intruding into the character's mental world: rather, a description is given that invites the listener to imagine what a character intends or perceives, thereby effectively getting to share in the character's perspective on the story world to some degree. Narrative Reports of Speech Acts (NRSA) and cases of (Free) Indirect Speech (*(F)IS*), relate what a character says or said primarily in the words of the narrator, while (Free) Direct Speech (*(F)DS*) is to be taken as the literal rendition of a character's words.

Still, what all these forms of speech reporting have in common is that they do not imply that the narrator has direct insight into characters' minds. Here too it is strictly speaking the listener who is cued to draw conclusions about a character's perspective based on the report of what they say or said. This contrasts with thought

Type	Example	Counts	ID
Ego-narration (EGO−NARR)	*'I love music'*	47	061401
Intention-in-action (IIA)	*'she went back again to the zoo'*	350	072201
Perception (PER)	*'she saw a huge lot of tigers and other animals'*	53	072201
Narrative Report of Speech Act (NRSA)	*'she did not ask the teacher about it'*	15	033401
Direct Speech (DS)	*'then said the sister* [...] *where is my little brother now'*	74	072201
Free Direct Speech (FDS)	*'oh little brother where are you now'*	14	072201
Indirect Speech (IS)	*'she said that they had to stop swimming'*	5	114201
Free Indirect Speech (FIS)	NA	NA	NA
Narrative Report of Mental State (NRMS)	*'he did not like that'*	98	061401
Viewpoint Package (VP)	*'because he entered secretly'*	44	101901
Direct Thought (DT)	*'then he thought I want to protect her'*	17	052901
Free Direct Thought (FDT)	*'shall I make some invitations for her friends'*	1	052901
Indirect Thought (IT)	*'the family thought that they were safe'*	17	112301
Free Indirect Thought (FIT)	*'he could wish for everything that he now wants'*	15	014901

Table 5.2: Fourteen types of CPR and their frequencies as found in our sample.

representation in its different forms, where access to a character's mind is relied on by default.[7] This goes for Direct Thought (DT) and Indirect Thought (IT) alike, even though in the latter case the contents of the character's thoughts are rendered in the narrator's words (see also the examples in Table 5.2). Narrative Report of Mental State (*NRMS*) is an ambiguous type in this respect; it can sometimes imply access to a character's mind, but in other cases reflect the narrator's reading of a mental state 'from the outside' (viz. characterising someone as 'happy' can be based on their behaviour as well as on narratorial access to their inner life).

[7] In classic narrative theory this is referred to as narrator omniscience; cf. Margolin (2014). Furthermore, for an extensive discussion of FIS and FIT as forms mixing elements of direct and indirect representation, see Vandelanotte (2009).

Looking at frequencies in the representation of speech and thought, it is apparent that DS is the most used type of speech representation (74 occurrences), whereas the much more indirect NRMS is most frequent (98) in representing thought. Finally, the type Viewpoint Package (VP), recorded 44 times, is introduced by us based on recent work by Van Duijn and Verhagen (2018) that we found useful in our children's story context. In short, Viewpoint Packages are single words implying a mental state contrasting with a state of affairs or with another mental state. For example, if a character does something 'secretly', this implies that there is a perspective from which this is *not noticed* and a perspective from it is indeed *desired* that it remains unnoticed.

We follow Nicolopoulou and Richner (2007) in their analysis suggesting that, for a storyteller, IIA and PER require less advanced efforts on a cognitive level, compared to handling character speech representation. Dealing with character thought, in turn, is argued to be more advanced on a cognitive level than handling speech, for exactly the reason discussed in the preceding paragraph: thought representation requires the narrator to intrude into character minds, whereas speech representation does not. Following this analysis, plus our own analysis of ego-narration, the order in which we present the fourteen types in Table 5.2 can be seen as indicating different stages, ranging from preliminary (EGO−NARR), to basic (IIA, PER), to intermediate (NRSA, (F)DS, (F)IS), to advanced (NRMS, VP, (F)DT, (F)IT).

Hypotheses

First it is our aim to explore variation in the use of CPR types within our sample as a whole. Second, we hypothesise that the occurrence of these types is not uniformly distributed over age groups. From the idea that some CPR types can be seen as more advanced than others, as we discussed in the previous section, we predict that preliminary and basic types of CPR occur more often at younger ages, while intermediate and advanced types are more often found in older children. Third, we aim to explore links between CPR types and linguistic information extracted using NLP tools. We predict that more advanced types of CPR are more likely to co-occur with utterances exhibiting higher lexical and syntactic complexity.

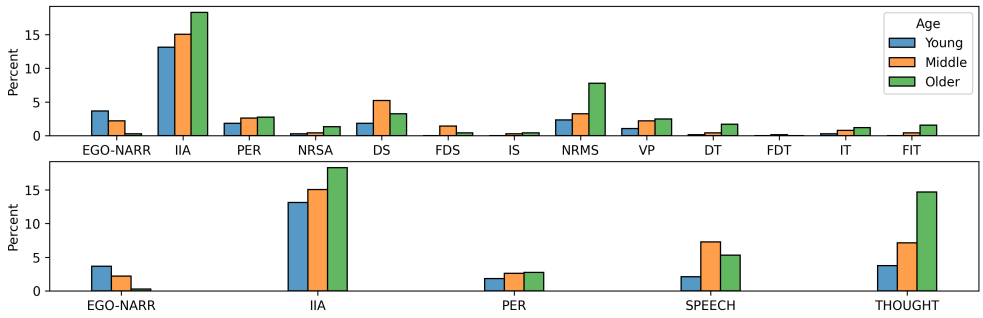

Figure 5.1: Occurrence of the original (top) and merged (bottom) CPR types in stories by children from three age groups, in percentages. Young: 4-6y, Middle: 6-9y, and Old:9-12y. Bars stack to 100%.

Type	O_{young}	O_{middle}	O_{old}	E	χ^2	p
EGO−NARR	28	17	2	15.67	21.74	<.001*
IIA	96	114	140	116.67	8.39	.015
PER	12	20	21	17.67	2.75	.252
SPEECH	12	55	41	36	26.72	<.001*
THOUGHT	28	53	111	64	56.66	<.001*

Table 5.3: Observed frequencies (O), expected values (E), and χ^2 statistics with $df = 2$ for all merged CPR types. Since we run 5 separate tests on the same variable, α was set to $.05/5 = .01$. Asterisks indicate $p < \alpha$.

5.4 Results

Development

For statistical analyses of the observed counts we merged CPR types that are theoretically closely related. In line with the stages discussed above, NRSA, DS, FDS, IS, and FIS were grouped as SPEECH, and NRMS, VP, DT, FDT, IT, and FIT as THOUGHT. CPR as found in our sample is plotted for both the five merged and thirteen original types in Figure 5.1. We conducted several χ^2 goodness-of-fit tests to probe whether observed frequencies for a given CPR type differed significantly from a uniform distribution among the three age groups. Test statistics and p-values are given in Table 5.3, with

$$\chi^2 = \sum_{k=1}^{n} \frac{(O_k - E_k)^2}{E_k}, df = k - 1. \tag{5.2}$$

We see that younger children use a lot more EGO−NARR, but older children a lot less compared to the expected value E; the distribution is significantly different from

Lexical Complexity				Syntactic Complexity			
Age	$\bar{x}$	σ	N	Age	$\bar{x}$	σ	N
Young	5.72	.63	176	Young	2.69	.85	176
Middle	5.96	.65	259	Middle	2.70	.97	259
Older	5.99	.63	315	Older	2.75	.85	315

Table 5.4: Descriptive statistics for lexical complexity and syntactic complexity, for a total of 750 utterances featuring CPR from 150 stories, 50 stories per age group.

uniform. This suggests children 'outgrow' ego-narration, which we argued is a preliminary stage of CPR, and as hypothesised it seems to disappear from children's storytelling as they get older. For both IIA and PER, which we called basic types of CPR, the distributions do not differ significantly from uniform. Thus, there are no age-specific preferences among children for either IIA or PER, contra our hypothesis that these basic types occur mainly at young age.

With regard to SPEECH, the distribution among age groups is significantly different from uniform. We see little use among young children compared to the expected value E, but a peak in use at middle age and then a slight decrease in use for the older group. This supports our hypothesis that SPEECH, which we argued is an intermediate type of CPR, is increasingly used at a later age. THOUGHT is significantly different from uniformly distributed and seems to take off rather late. The younger and middle groups use less THOUGHT compared to the expected value E, whereas the older group uses it a lot more. This pattern observed regarding THOUGHT offers clearest support for our prediction that advanced types of CPR are increasingly employed at a later age.

In summary, children of all ages in our sample tell stories in which character perspectives are represented. As children grow older, perspectives of their characters tend to be fleshed out in more diverse and advanced ways. For the middle group we observe that characters more often speak and have various kinds of thoughts and other mental states. The older group relies even more often on forms of thought representation, and slightly less on character speech; possibly using the first partly instead of the latter.

Linguistic contexts

Next we examine whether more complex types of CPR co-occur with utterances that are lexically and syntactically more complex. We automatically extracted Lexical Complexity (LC) and Syntactic Complexity (SC) for each utterance.

		Lexical Complexity		Syntactic Complexity	
	Effect	Estimate	*SE*	Estimate	*SE*
Simple	(Intercept)	5.759*	0.138	2.363*	.182
	IIA	.211	.151	.327	.204
	PER	-.028	.228	.733*	.318
	SPEECH	.331	.238	.002	.329
	THOUGHT	.344	.186	.419	.255
	Middle	.055	.239	-.051	.306
	Older	1.114*	.470	.602	.673
Interaction	IIA x Middle	.091	.254	.198	.331
	PER x Middle	.216	.328	-.201	.448
	SPEECH x Middle	.172	.323	.248	.431
	THOUGHT x Middle	.049	.285	-.014	.377
	IIA x Older	-.896	.474	-.433	.682
	PER x Older	-.575	.521	-.689	.748
	SPEECH x Older	-1.022*	.515	-.313	.739
	THOUGHT x Older	-1.136*	.488	-.770	.700
Random	Child (intercept)	.103	.012	.076	.010
	Residual	.329	.021	.744	.031

Table 5.5: Terms for two linear mixed models with by-child varying intercepts. EGO−NARR and Young age group are the reference classes, i.e. the intercept is the average perplexity/tree depth for an utterance of a young child with the Ego-narrator type. Asterisks indicate $p < .05$.

For LC, we first lemmatised utterances with the spaCy parser (Honnibal and Johnson, 2015), and calculated the lexical perplexity; for SC, we also used the spaCy parser to extract the maximum depth of the parsed tree, as described in Section 5.3. Means and standard deviations of the extracted features are given in Table 5.4. As can be seen, average differences for both lexical and syntactic complexity are small across the three age groups. Next, we employed LC and SC as dependent variables in two linear mixed models. We included our five merged types of CPR as categorical predictors and included interactions with our three age groups, to find out whether different CPR types have significantly different mean LC and SC values, while taking potential age differences into account. Coefficients are given in Table 5.5. Our overall finding is that the link between lexical and syntactic complexity and specific types of CPR is not as we anticipated.

We first discuss the results for LC. Here we see that the only significant simple effect is Older, which means that with respect to the young EGO−NARR reference class, older children use ego-narration in utterances that are lexically more complex than young children do. Further, we see two significant negative interactions with SPEECH and THOUGHT, indicating that as we ascend from our reference class to older children

that use these intermediate and advanced forms of CPR, the lexical complexity of the utterances decreases, which is contrary to what we hypothesised with respect to LC. We do not see evidence for our hypothesis that average LC for more complex types of CPR is higher compared to ego-narration, while taking age differences into account.

Next we elaborate on our results for SC. Here we see no evidence for our hypothesis that more complex forms of CPR co-occur in utterances that have higher average syntactic complexity, while taking age differences into account. Main and interaction effects are all insignificant, except PER as simple effect, which implies that with respect to our young EGO−NARR reference class, average SC is higher when young children employ PER as type of CPR. This is contrary to what we hypothesised, as PER is a basic CPR type which we expected to co-occur with less complex syntax.

Our results are not in line with earlier work suggesting that children's more advanced lexical and syntactical skills co-occur with better socio-cognitive skills (as reviewed in e.g. Milligan et al., 2007). One possibility is that the way we looked at lexical and syntactic information in utterances here, provides a too limited view on the contexts in which different types of CPR occur. Given that other studies demonstrate that lexical complexity on the level of the entire stories children tell, predicts the occurrence of more sophisticated story characters (see Chapter 2 and Chapter 4) we suggest that automatically extracted information on the story level (as opposed to the utterance level only) could be more helpful for modelling CPR occurrence in the future.

5.5 Discussion

Our inventory shows that CPR is ubiquitous in freely told children's stories and that it takes many different forms. We discussed that classification of perspective phenomena into a system of CPR types requires knowledge of linguistic and narratological theory, and that it is regularly dependent on thorough analysis of utterance context within a story. Reliance on a single annotator is a weakness of this study; however, we believe to have met the goal of building a meaningful (foundation of a) bridge between long-standing research traditions in the humanities and current approaches in the computational sciences.

Regarding ego-narration we have identified cases exhibiting a problematic mixing between children's own perspective and the narrator's or characters' perspectives in the story, and argued for seeing these as a preliminary stage of CPR. Also, building on existing work from developmental psychology and cognitive linguistics, we have

introduced the types IIA, PER, and VP in our analysis, covering perspectives implied in actions, perceptions, and single lexical units such as 'secretly'. This was particularly useful for getting a grasp on the more basic stages of perspective coordination as present in our sample of children's stories. Although we did not see occurrence of these basic stages peak at younger ages, as we expected, we presented evidence that indeed more complex types are implemented more frequently at later ages.

Furthermore, our aim was to link automatically extracted linguistic information to the occurrence of different types of CPR, while also taking age differences into account. The picture that emerged for lexical and syntactic complexity was more complicated than we anticipated. By taking into account dependencies between utterances coming from the same speaker by using random intercepts, and by including interactions with age in our statistical models, we tried to describe as much variation as possible in the language children use when rendering character perspectives. As we saw, overall average differences in lexical and syntactic complexity between ages were small at the outset, and we were not able to link higher linguistic complexity to advanced types of CPR. Here the overall sparse occurrence of several of the individual types likely calls for exploiting a larger part of our story database in the future. We also learned that using perplexity and tree depth to describe the immediate (utterance-level) contexts in which CPR types occur, is challenging, suggesting that additional types of linguistic information from wider (story-level) contexts could be needed.

All in all, these findings and lessons encourage us to pursue the line of inquiry set out in this chapter. This will also require refining our framework, models, and automatically extracted information in interaction with linguistic and narratological theory, for which additional interdisciplinary cooperation is indispensable.