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

Dijk, B.M.A. van

Chapter 2

Modelling Story Characters' Mental Depth

From age 3-4, children are generally capable of telling stories about a topic free of choice. Over the years their stories become more sophisticated in content and structure, reflecting various aspects of cognitive development. Here we focus on children's ability to construe characters with increasing levels of mental depth, arguably reflecting socio-cognitive capacities including Theory of Mind. Within our sample of 51 stories told by children aged 4-10, characters range from flat 'Actors' performing simple actions, to 'Agents' having basic perceptive, emotional, and intentional capacities, to fully-blown 'Persons' with complex inner lives. We argue for the underexplored potential of computationally extracted story-internal features (e.g. lexical/syntactic complexity) in explaining variance in Character Depth, as opposed to story-external features (e.g. age, socioeconomic status) on which existing work has focused. We show that especially lexical complexity explains variance in Character Depth, and this effect is larger than and not moderated by age.

This work was originally published as: Van Dijk, B.M.A. and Van Duijn, M.J. (2021). Modelling Characters' Mental Depth in Stories Told by Children Aged 4-10. In Fitch, T., Lamm, C., and Leber, H., editors, *Proceedings of the Annual Meeting of the Cognitive Science Society*, volume 43, pages 2384-2390. Cognitive Science Society.

2.1 Introduction

From early childhood children tell stories to themselves and others as part of their daily play activities (Cremin et al., 2017; Sutton-Smith, 2012). Such storytelling has been described as a kind of cognitive play that — besides being the source of a lot of fun — forms a natural crossroads of various key areas in child development (Bergen, 2002; Paley, 1990; Smith and Roopnarine, 2018). Telling stories involves language skills at the phonological, lexical, syntactic, and pragmatic levels (Southwood and Russell, 2004). It draws further on cognitive abilities such as memorising, planning, organising knowledge of the world (McKeough and Genereux, 2003), and empathising with others, in particular to work out how characters should behave, speak, feel, and think in ways that are relatable and interesting for an audience (Nicolopoulou, 1993; Van Duijn et al., 2015; Zunshine, 2006).

Here we are interested in the representation of mental activities of characters and the place this has in child development. Existing theoretical work has linked children's ability to render character minds to the mastery of socio-cognitive skills, in particular mindreading or Theory of Mind (ToM).[1] Empirical research has shown that the complexity of characters and their mental activities that children can deal with tends to increase with age (e.g. Nicolopoulou and Richner, 2007; Nicolopoulou and Ünlütabak, 2017).

For this chapter, we recorded and transcribed 51 oral stories elicited from Dutch children of different ages and backgrounds, during storytelling workshops integrated in their daily school and daycare environments. A total of 268 characters were represented in these stories, each of which we assessed in terms of its mental depth. To give two brief opening examples of what we looked at (excerpts translated from stories we collected earlier):

(1) *they sit neatly in a row but the other* [*puppy*] *always enters later* (child 4y;1m)

(2) *they sat down as always until he was not looking* [...] *then they went inside the school director's office and secretly took the key* (child 9y;11m)

Characters presented in excerpts (1) and (2) fall at the lower and higher ends of the scale of mental depth that we will introduce in more detail below respectively. Excerpt (1) introduces characters with arguably different perspectives on the staged set-

[1] For a general overview of literature on mindreading/ToM, i.e. the ability to take others' perspectives and reason about their behaviour in terms of emotional and intentional states, see Apperly (2012). For an overview of theoretical work linking ToM with children's stories see Nicolopoulou (2015) and Zunshine (2019).

ting: some are already inside, while another one enters later. However, there is no fleshing out of mental activity by any of these characters at all, and this is representative of the rest of the story, which revolves around movements (coming in, going out) and actions (eating) only. This is very different in (2), where the implied protagonists' awareness of what the school director does not see, and hence knows, is central in the story's plot.

In line with existing work, we observe an overall increase in mental Character Depth with the age of the children telling the stories in our sample. However, it is our aim to understand in more detail which factors drive children's ability to render more complex characters. To this end, we develop a framework using computational techniques and statistical modelling for mapping out relationships between, on one side, the mental depth of story characters and, on the other side, multiple story-external features (e.g. age, socioeconomics) and automatically parsed story-internal features (e.g. vocabulary, syntax).

Our results show that in particular the lexical complexity a story exhibits can be used as a reliable predictor of Character Depth: it explains a larger proportion of the variance compared to age and is not moderated by age. We discuss the role of lexical complexity and other variables in understanding children's ability to deal with characters of different levels of mental complexity, both within our current sample and in larger, more diverse samples in the future.

2.2 Background

Narrative plays a key role in human communication. On a daily basis adults and children alike use stories to share their perceptions and imaginations with others, typically in causally, temporally, and logically structured ways. Classic definitions of narrative often emphasise criteria such as goal-directedness, causality, or the unfolding of series of actions over time (Duinmeijer et al., 2012). However, in this research we cast the definitional net a bit wider and argue that children's stories could also be descriptions of situations, events, or characters in which goals, causal relations, or a clear temporal development are not immediately present. What we take as our central criterion here to demarcate stories from other speech phenomena is *mediatedness* or *transcendence*, marked by a departure from the actual speaker and its immediate here-and-now (cf. Nicolopoulou and Richner, 2007; Zeman, 2018). For example, children merely describing their situation during the storytelling workshop in which we collected our data would not be telling a story (e.g. *'I am sitting on a chair in the* *group circle...'*), whereas children describing a real or imagined situation set elsewhere would be, even if that situation is not worked out any further with additional characters and events (e.g. *'Yesterday I had a silent disco...'*).

In this chapter we focus on two of the developmental trajectories that naturally intersect in stories that children tell, social cognition and language, against the background of their more general development, which we approximate via age and educational level of the parents/caregivers. Following a large body of research (for an overview see Milligan et al., 2007; Tompkins et al., 2019), we expect these trajectories to be interrelated and it is our longer-term aim to contribute to further understanding of this interrelatedness by studying stories that children tell. Here we develop a framework for mapping out features within such stories that we assume to be manifestations of developmental progression on the linguistic and socio-cognitive levels. Our hypotheses at this stage concern the co-occurrence of and relationships between these features within the stories; testing whether this is indeed indicative of the development of the children who tell them is outside the scope of this chapter.

Social cognition

Firstly, we are interested in socio-cognitive sophistication of the stories, which we operationalise as the mental depth that characters exhibit, in short, Character Depth (CD). Using a slightly adapted version of the typology introduced by Nicolopoulou and Richner (2007) we rate each character's mental activity on a nine-level scale. These levels fall under three main categories: `Actors` undergoing (level I) or performing (level II) simple actions, `Agents` having basic perceptive, expressive, emotional, and intentional capacities (levels III-V), and `Persons` capable of coordinating beliefs, desires, expectations, and so on, with different imagined realities (levels VI-VII) and/or other characters' cognitive states (levels VIII-IX; see Section 2.3 and Table 2.1 below for more details).

Language

Secondly, we are interested in the linguistic qualities of the stories, which we operationalise on two levels: vocabulary and syntax. As a measure of vocabulary sophistication (a.k.a. lexical complexity) we assessed the vocabulary of each story by computing the probability of the occurrence of each lemma that a child used approximated by frequencies in a benchmark lexicon. This metric builds on the idea that the difficulty of words from the perspective of a language learner is strongly negatively

ACTOR	**I - Passive: does nothing or undergoes something passively** *...er was een boerderij met schapen...* (there was a farm with sheep; 060301) *...toen ze vos gingen pakken...* (when they went to catch fox; 060301) **II - Active: coupled to action/description without clear intention/goal-directedness** *...ridders die tegen dino's vechten. en dan gaan ze ook dood...* (knights that fight with dinosaurs. and then they also die; 010701)	
AGENT	**IIIa - Implicit intention in action** *...wij gingen autorijden...* (we went to drive the car; 010501) **IVa - Action in response to/contrast with situation or event** *...er kwam een baby* [...] *en toen ging de vader weg...* (a baby came [...] and then the father went away; 021302) *...ze gingen* [...] *treetje voor treetje maar het ging mis...* (they went [...] step by step but it went wrong; 030303) **Va - Complex or explicit intention/goal-directedness in action** *...en toen kwam onze familie* [ons] *ophalen en toen naar huis brengen...* (and then our family came pick [us] up and bring us home; 010501)	**IIIb - Simple/implicit emotion, perception, attention** *...ik keek naar die...* (I looked at that one; 060701) **IVb - Emotion in response to situation/event or action in response to emotion/perception** *...ik ging* [...] *naar bowling* [...] *en het was heel leuk...* (I went [...] to bowling [...] and it was a lot of fun; 020501) **Vb - Complex/explicit emotion, perception, attention, utterance** *...in de auto was het heel stom, want we moesten lang wachten...* (in the car it was very stupid, since we had to wait long; 010501) *...papa zei er is niets aan de hand...* (daddy said all is fine; 010504)
PERSON	**VI - Explicit intentional state (desire, belief, etc.), implicitly coordinated with actions/events** *...de moedertrol wou de andere trol redden...* (the mother troll wanted to save the other troll; 011301) **VII - Explicit intentional state, explicitly coordinated (equation, contrast, etc.) with non-intentional aspects of the storyworld** *...die wou ik doen maar toen ging zij hem helemaal afmaken...* (that one I wanted to do but then she went to fully complete it; 010503) **VIII - Explicit intentional state, coordinated with the intentional state of another character or past/future self** *...ze wou heel graag buiten spelen, maar het mocht niet van haar moeder...* (she very much wanted to play outside, but her mother did not allow it; 010201) **IX - Multiple-order intentionality** *...haar vader en moeder* [*wouden niet*] *dat ze naar een bos ging.* [*ze*] *wisten dat in het bos een kluizenaar woonde en dat wist Amelie ook dat ze dat wisten...* (her father and mother [did not want] that she went to a forest. [they] knew that in the forest a hermit was living and Amelie knew too that they knew this; 021301)	

Table 2.1: Annotation scheme for CD. All examples are quotes from our dataset, followed by a somewhat liberal/idiomatic English gloss, followed by the unique ID of the story from which it was taken. Underlining indicates character to which the CD level applies in case of multiple characters in an example. Square brackets indicate elements of quotes that were reordered or omitted for purposes of readability.

correlated with how frequently it occurs (Vermeer, 2001). Thus, using less frequent words means using less probable words, and this we take to indicate a more complex vocabulary. The idea is that a more complex vocabulary functions as a communicative and mental toolbox that enables a child to render both the physical and social world better. This toolbox can be especially helpful when engaging in demanding tasks such as telling a story, where there is a sustained pressure for finding the right words to get the desired message across to an audience (Curenton and Justice, 2008).

As a measure of syntactic complexity, we calculated the average distance between syntactically dependent words. It is well-established that language structures which employ longer dependency distances between head words and dependent words are more difficult to process (Gibson, 1998; Gildea and Temperley, 2010). An example of this difference is given by King and Just (1991) in terms of subject-extracted relative clauses (3) and object-extracted relative clauses (4):

(3) The reporter who attacked the senator admitted the error.

(4) The reporter who the senator attacked admitted the error.

In both sentences the verb 'attacked' is dependent on the pronoun 'who'. In (4) these dependents are not adjacent, but have two words in between, which makes that part of the sentence more challenging to process. Average dependency distance seems to capture language skills more generally. For example, it can be used to distinguish English written by natives from that written by L2 learners (Oya, 2011) and speech from individuals with mild cognitive impairments from speech produced by typically developed speakers (Roark et al., 2007). Our idea here is that children capable of handling more complex syntactic structures, as indicated by their stories exhibiting higher average dependency distances, have more powerful formats available for representing events in the social and mental worlds, in discourse as well as in their own strands of reasoning (cf. De Villiers and De Villiers, 2014).

Hypotheses

Firstly, we hypothesise that stories exhibiting a more complex vocabulary contain characters with higher levels of mental depth. Secondly, we hypothesise that stories with larger syntactic dependency lengths contain characters with higher levels of mental depth.

Story-external features

Existing work has shown that the mental depth of characters in stories that children tell increases with their age (Nicolopoulou and Richner, 2007), which is why we include it in the model. Parent education functions as a proxy for socioeconomic status in our model; there is evidence that children from parents with a higher socioeconomic status perform better on ToM tasks (Shatz et al., 2003).

2.3 Methods

Dataset

For our data collection, we offered storytelling sessions to various institutions in the medium-sized Dutch cities Leiden, Tilburg and Utrecht. Three schools (two in Leiden, one in Utrecht), one daycare (Leiden) and one community centre (Tilburg) were willing to cooperate. Around 200 children in total participated in sessions held between September 2019 and June 2020. We were able to include 98 stories told by 54 children ($M_{age}(SD) = 6.81(1.66)$, range = 4.17-10.1; 30 females, 2 unknown) in our database after receiving consent forms from their parents. In order to maximise independence between observations we use only the first story told by each child, and due to missing information on the consent forms an additional 3 stories dropped out, resulting in a subset of 51 stories for this chapter. Our experiment and data management protocols were assessed and approved by the Ethical Committee of the Leiden University Faculty of Science (file no. 2020 – 002).

Our storytelling sessions were held in group circle settings. After briefly exploring some general features of stories interactively (e.g. 'What is a story?', 'What do we find in stories?') and narrating a short standard exemplary fantasy story, we invited children to tell a story about a topic free of choice. Voice recordings were made after informing the children about this. Afterwards, the recordings were pseudonymised and transcribed by the authors and research assistants twice: first orthographically (including 'noise' such as false starts, wrong conjugations, broken-off words, etc.), and second normalised, thus without noisy elements, to enhance compatibility with computational language processing tools. All transcripts were double-checked for consistency with the audio files. In addition to the story data, personal data such as age of the children and parental education levels were collected through consent forms. Transcripts, data, and code are available via `https://osf.io/k52e8/`.

Annotations

We loaded all pseudonymised transcriptions in the open online content annotation tool CATMA (version 6.1.3; (Horstmann, 2020)), where we created a tag set for CD. Tags within this set were based on the typology introduced by Nicolopoulou and Richner (2007). A few adaptations were made, however, in terms of the three main levels (`Actor`, `Agent`, and `Person`) our tag set remained compatible with the original typology. See Table 2.1 for descriptions and examples of the tags we have used to assign a CD level to each character. Our workflow included a first stage in which the authors of this chapter discussed the first 10 stories openly, followed by a second stage in which the remaining 41 stories were annotated by each of the authors independently. In the third stage, all tags that differed were discussed until consensus was reached. Finally, the annotations were considered fixed and downloaded from CATMA in TEI-XML format.

We extracted the maximum CD with a Python script. This feature represented the highest level of CD reached in a story on a scale from 0 to 9, corresponding with the levels in the topology set out in Table 2.1 when discarding subcategories indicated by letters (e.g. IVa and IVb both count as 4), where 0 indicates the theoretical option of no characters being presented in a story (which did not occur in our current dataset), value 1 corresponds with level I in Table 2.1, and so on.

Extracting Linguistic Features

Vocabulary Probability – Our approach was to take the textual vocabulary of a representative reference corpus, which consists of all the lemmas constituting the vocabulary of the corpus (Fengxiang et al., 2016). We use this benchmark to compute the probability of each story vocabulary, treating it as a subset of the textual vocabulary. Lemma probabilities were approximated by relative frequency counts in the reference corpus.

We obtained lemmas for each story by parsing normalised story transcripts with the memory-based Frog parser (Van Den Bosch et al., 2007). We used as reference corpus the 'free text' lexicon (FTL) of the BasiScript corpus (Tellings et al., 2018a), which consists of essays of primary school children with minimal teacher intervention, thus staying close to the free story paradigm. We removed punctuation marks and named entities from the FTL, which yielded a total number of token instances N of 3699822, and a vocabulary V of 46570 lemmas. The estimated probability of some lemma l_i

occurring in story S is given by

$$P(l_i) = \frac{(c_i + 1)\frac{N}{N+V}}{N}, \tag{2.1}$$

with c_i being the count of token instances of l_i in the FTL, adjusted for words not occurring in the FTL. This estimation is based on n-gram smoothing methods as outlined by Jurafsky and Martin (2024); we used Laplacian smoothing since the FTL includes many typical fantasy constructions such as 'trollensnot' (troll snot) with count 1, but not the similar construction 'eenhoornsnot' (unicorn snot) which occurs in our stories. We calculated the probability of the vocabulary of S with

$$L = \frac{1}{S_n} \sum_{i=1}^{n} P(l_i), \tag{2.2}$$

with the fraction being a normalising factor (S_n being the length of S), and converted them to per mille for convenient interpretation. The interpretation of L can be phrased as follows: if one draws a lemma from the FTL, how likely is it that it belongs to the story vocabulary? For complex vocabularies this probability will be lower.

Dependency Distance – We used the Alpino parser (Van Noord, 2006) to extract all dependencies per utterance. The dependency distance of the ith dependency relation DD_i is typically set to 1 for adjacent words, 2 if one extra word occurs in between the dependents, and so on. We follow Wang and Liu (2017) and compute overall dependency distance MD_{sent} for a sentence with n words by

$$MD_{sent} = \frac{1}{n-1} \sum_{i=1}^{n-1} |DD_i| \,. \tag{2.3}$$

Then, for a story consisting of multiple utterances,

$$MD_{story} = \frac{1}{u} \sum_{i=1}^{u} MD_i, \tag{2.4}$$

where u is the total number of utterances in a story.

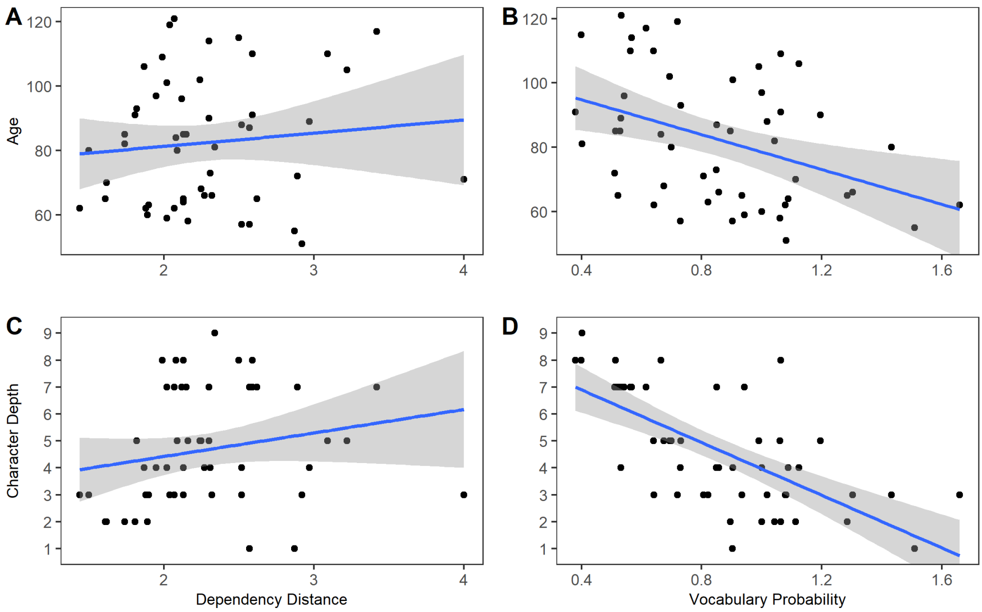

Figure 2.1: Correlation plots with Age in months, average Dependency Distance in number of words, average Vocabulary Probability in per mille, and Character Depth in levels.

2.4 Results

Bivariate explorations

Prior to constructing the linear model that we used for assessing our hypotheses, we explored various correlations between a subset of the features outlined above.

Firstly, it appears that Dependency Distance correlates weakly with Age (Figure 2.1 A, Pearson's $r = 0.104$) and Vocabulary Probability correlates moderately with Age (Figure 2.1B, Pearson's r = -0.403). It makes sense that as children grow older, both their vocabularies and syntax are becoming increasingly complex. Secondly, Dependency Distance correlates weakly with CD (Figure 2.1C, Pearson's r = 0.202) and Vocabulary Probability correlates strongly with CD (Figure 2.1D, Pearson's r = -0.670), indicating that the relationship between the parsed linguistic features and CD are in the expected directions, albeit in quite different gradations. In the next section we scrutinise these bivariate explorations using a linear regression model.

Predictor	β	*SE*	*t*	*p*	95% CI	
Intercept	4.716	.250	18.892	<.001	4.212	5.219
Vocabulary Probability	-1.117⋆	.289	-3.860	<.001	-1.701	-.534
Age	.582⋆	.265	2.193	.036	.047	1.117
Education Parents	.425	.265	1.602	.116	-.110	.961
Vocabulary Probability * Age	.161	.292	.551	.584	-.428	.750
Dependency Distance * Age	.146	.232	.628	.533	-.322	.614
Dependency Distance	-.016	.242	-.067	.947	-.110	.473

Table 2.2: Model estimates sorted on the magnitude of the standardised betas. Stars denote significance at the .05 level, two-tailed.

Hypothesis testing

We consider a linear multiple regression model most appropriate for the analysis; due to the limited number of observations per institution in our dataset, a mixed-effects model did not converge properly. Our model includes Dependency Distance, Vocabulary Probability, Age, Education Parents, and interactions between Vocabulary Probability and Age and between Dependency Distance and Age as predictors of CD. The model accounts for about 53% of the variance in CD $R^2 = .525, F_{6,44} = 8.132, p < .001$, with $M_{CD} = 4.667, SD_{CD} = 2.167$. Standardised coefficients sorted on magnitude are given in Table 2.2.

In line with our first hypothesis, we see that the simple effect of Vocabulary Probability has the largest negative and significant slope. This indicates that as the vocabulary of a story becomes less probable, i.e. the lexical complexity of that story goes up by our measure, characters tend to become more complex in terms of their mental depth, with other effects fixed at mean level. In addition, we observe in Table 2.2 a positive and significant simple effect of Age, which means that as children get older, the characters they use in their stories tend to get more complex in terms of mental depth, with other effects fixed at mean level. However, this effect is only a bit over half the magnitude of that of Vocabulary Probability ($\beta = .528$ versus $\beta = -1.117$).

We learn more about the relationship between Vocabulary Probability and Age by looking at the small non-significant interaction effect Vocabulary Probability * Age in Table 2.2. It indicates that the effect of vocabulary is not moderated by Age, in other words, is not significantly different for children of different ages. This is visible in Figure 2.2, where three lines indicate predictions of CD for various ages, but have similar slopes.

With respect to our second hypothesis, we observe in Table 2.2 that the simple

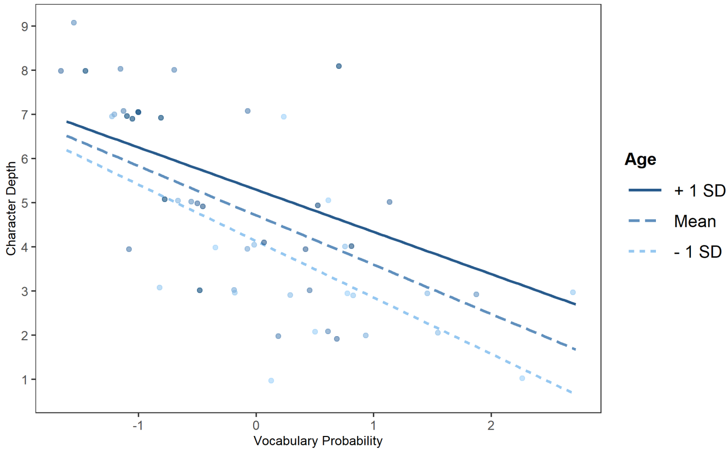

Figure 2.2: Interaction plot of Vocabulary Probability * Age with Age in months, average Vocabulary Probability (z-scored), and Character Depth in levels.

effect Dependency Distance and the interaction Dependency Distance * Age are both small and non-significant. Thus, contrary to our expectation, this model suggests that the distance between syntactically dependent words does not explain the observed variation in the levels of Character Depth, nor can we say that age plays a moderating role here. Finally, we can see in Table 2.2 that the main effect of Education Parents is positive and a bit smaller than age, but non-significant, suggesting that parental years of education do not reliably predict levels of characters' mental depth either.

Although we saw in the bivariate explorations that there is a moderate correlation between Vocabulary Probability and Age, (Pearson's r = -.403), we have no indications that these and other predictors pose multicollinearity issues for the estimates in our model, since all computed Variance Inflation Factors are below 1.44 (with a conservative threshold of 5). We thus find some tentative evidence for the idea that in our model, Vocabulary Probability and Age have independent effects.

2.5 Discussion

Our central finding is that lexical complexity is a key story-internal feature for predicting a story's socio-cognitive sophistication, as manifested in the mental depth of characters. This finding has multiple implications and possible interpretations. Firstly, it seems to follow that rich vocabularies are particularly helpful in organising and describing the storyworld, including its social and mental aspects. In theory, this could be entirely independent of actual socio-cognitive skills possessed by the child telling the story: it could be merely a matter of being able or not to find the right words for fleshing out a character in terms of its emotional and intentional states.

However, with existing research in mind (e.g. De Villiers and De Villiers, 2014; Milligan et al., 2007) it appears more likely that our observed effect extends beyond the realm of the stories as such, and that possessing a more advanced vocabulary not only enhances a child's communication about the social world, but also supports its understanding of and ability to reason about socio-cognitive matters. Here it is particularly salient that the effect is larger than and not moderated by age. This adds a new perspective to the debate about the period in which children start to invoke others' mental states in their language (for an overview see Nicolopoulou, 2015).

Rather than disclosing a 'Rubicon' moment for ToM-language use in children, we propose a methodology that can show what it is about certain aspects of language development, such as having access to a more advanced lexicon, that engenders fleshing out mental activity in more detail, regardless of what age a child has. To substantiate such an interpretation, further research is needed focused on establishing firmer links between patterns observed inside stories and development as it takes place within the children that tell them. Here we see a role for collaborative work involving both (computational) linguists, narratologists and developmental psychologists.

For syntactic complexity the picture is quite different; we see no significant evidence for its contribution to Character Depth in our sample. Although in our bivariate explorations we saw a hint of the relation we hypothesised, in our model it was probably trumped by other effects. A reason for this could be that speech employs overall lower dependency distances compared to written text, which for children may even be stronger the case. If dependency effects are thus generally smaller, we must revisit this prediction with more data and maybe also compare and evaluate different metrics of syntactic complexity, such as clause length and words per finite verb.

A general remark about our methodology is that the use of computational language processing tools makes operationalising 'narrative sophistication', as we have done (and as is also proposed by Nicolopoulou (2016)), a lot easier, more reproducible, and more scalable. With larger datasets we might in the future be able to use story-internal variables to approximate children's narratological and linguistic capacities, as well as related cognitive skills, when no external information about the storytellers is available, or when collection and storage of sensitive data from children or parents is to be minimised.

In addition to (and to provide a more solid foundation for) such computational approaches, we see multiple directions in which research may go that aims to deepen our understanding of the relationships between socio-cognitive development and narrative/linguistic competence. A possibility would be to include stories from a more diverse population, for example by involving atypically developing children, and/or collect additional data about each storyteller's performance on relevant standardised tasks (e.g. those used by Wellman and Liu (2004)). Another exciting possibility would be to compare our sample to story corpora in other languages, ideally differing substantially from Dutch in their syntactic and semantic structuring. Such extensions could help to further bootstrap patterns within the stories on trends in individual development, and shed light on directionality and causality of the interactions.

Finally, insufficient returned consent forms and other factors diminished the number of children per session we could include, which constrained this study to a fixed-effects model. Using more advanced random effect modelling we could most likely make better estimates of the relevant relationships, since such models would be able to take session-bound dependencies between for instance vocabularies into account. With this perspective in mind, we emphasise that a first improvement for our future research will be to focus on more participants per workshop session. Currently, the prospects for our story corpus are looking good: recent data collections in Spring 2021 yielded about 200 additional stories to be analysed. The goal for the rest of this year is to compile a corpus of at least 500 stories, consisting of around 8 hours of high-quality child speech recordings and 50000 tokens, that is open to researchers with all kinds of backgrounds and interests. A huge bonus so far is that children love our storytelling workshop, and are happy to see us come each time.