9 Results Study II

In this section, following the preliminary analyses for differences between participant groups, all results will be displayed in regard to the specific research questions.

9.1 Preliminary Analyses

Summary data describing the characteristics of the children by randomized group assignment are given in Table 11 for tutors and Table 12 for tutees. Tables 13 and 14 provide a comparison of tutees' and tutors' narrative performance.

Tutee Groups

Table 11 provides an overview of group characteristics for the children in the intervention and control conditions. Participants in the PT group (3 boys, 5 girls) had a mean age of 4;7 and a mean exposure to German of 24 months; children in the PP group (5 boys, 3 girls) had a mean age of 4;5 and a mean of 24 months' exposure to German; finally, the children in the CG (3 boys, 5 girls) were on average 4;6 old with an average German exposure of 25 months (*SD*, ranges, and further characteristics appear in Table 11).[66] A Kruskal-Wallis H test yielded no differences between the three groups for age, exposure to German, expressive and receptive language, and nonverbal intelligence.

[66] Although gender was not equally distributed among the groups, the groups were still deemed comparable, as, similar to other studies (e.g., Hipfner-Boucher, 2011), gender analyses conducted in the previously presented study on narrative skills in Turkish-German DLLs did not reveal any significant differences between boys and girls on the measures used in this study (see section 5.1).

U. Licandro, *Narrative Skills of Dual Language Learners*, Diversität in Kommunikation und Sprache / Diversity in Communication and Language,
DOI 10.1007/978-3-658-14673-3_9

Table 11. *Participant Characteristics for the CG, PP, and PT Groups (Tutees)*

Variable	Group	*M*	*SD*	Range	*p*
Age in months	CG	55.63	6.26	48-66	.888
	PP	54.50	7.84	44-66	
	PT	56.50	7.86	47-67	
Exposure to German in months[a]	CG	24.88	12.43	12-48	.997
	PP	23.75	9.39	14-43	
	PT	24.50	11.14	13-46	
Education mother in years	CG	7.75	3.73	0-10	.490
	PP	9.63	3.58	4-13	
	PT	10.00	3.51	4-17	
Education father in years	CG	7.75	3.73	0-10	.435
	PP	9.75	2.82	4-13	
	PT	12.13	4.05	9-17	
Expressive language[b]	CG	20.13	17.47	3-42	.580
	PP	19.88	7.95	8-34	
	PT	25.88	12.79	8-45	
Receptive language[b]	CG	21.13	9.54	8-33	.580
	PP	18.25	4.27	13-24	
	PT	19.50	4.57	10-25	
Nonverbal intelligence[c]	CG	15.13	2.95	12-20	.538
	PP	16.63	2.77	12-19	
	PT	15.50	5.98	8-27	
EINC Frog Story[d]	CG	7.88	4.12	3-13	.749
	PP	9.00	4.21	4-17	
	PT	7.63	4.44	3-17	
EINC Climb Story[d]	CG	6.75	3.41	2-13	.512
	PP	8.50	4.21	4-15	
	PT	8.25	4.86	3-16	

Note. CG = Control; PP = Peer play; PT = Peer Tutoring; each group had 8 participants. Reported *p*-values refer to Mann-Whitney U tests.
[a]Based on parent report.
[b]Raw score sums, LiSe-DaZ expressive and receptive subtests (Schulz & Tracy, 2011).
[c]Raw scores, Raven Coloured Progressive matrices (Raven, 1995).
[d]Measures of narrative complexity based on generations of "Frog, where are you?" (Mayer, 1969); 'Climb Story' was a self-designed picture story. Narrative complexity measured using an adapted and extended version of the INC scoring rubric (Petersen, Gillam, & Gillam, 2008). The maximum score for each story was 26.
None of the above measures were significantly different between the groups.

Narrative performance was assessed based on spontaneous narration of the Frog Story (for procedures, see section 4.3). Microstructural measures (narrative productivity, lexical diversity, and syntactic complexity) as well as narrative complexity (as assessed via EINC) were compared between the groups. Furthermore, narrative com-

plexity was also assessed via a second spontaneous narrative production based on an unfamiliar 7-page-long picture book ("Climb Story") (also see section 8.2).

For the Frog Story's narrative productivity measures, a Kruskal-Wallis H test revealed no significant differences between the three experimental groups, with a mean total number of words (TNW) of 75.00 (SD = 30.03) for PT, 110.50 (SD = 62.52) for PP, and 68.00 (SD = 52.48) for CG, $\chi^2(2) = 1.83$, $p = .400$; similar to the total number of produced C-units, TNCU: PT, M = 43.43, SD = 57.36; PP, M = 23.62, SD = 9.72; CG, M = 17.38, SD = 9.20, $\chi^2(2) = 2.01$, $p = .367$. The measures of lexical diversity, namely number of different words in lemmas (NDW) and the vocabulary diversity statistic (VOCD), respectively, did also not differ significantly between the three groups (NDW: PT M = 30.34 (SD = 13.71), PP M = 29.88 (SD = 13.37), CG M = 22.63, (SD = 14.27), $\chi^2(2) = 1.70$, $p = .427$; VOCD: PT M = 17.28 (SD = 7.13), PP M = 11.16, (SD = 5.32), CG M = 14.33 (SD = 5.70), $\chi^2(2) = 2.99$, $p = .224$). Finally, there was no significant difference for syntactic complexity, as assessed by mean length of C-unit (MLCU), between the PT tutees (PT; M = 3.69, SD = 1.27) children assigned to the PP condition (M = 4.46, SD = 1.27), and the CG (M = 3.49, SD = 1.50), respectively, $\chi^2(2) = 1.93$, $p = .380$.

At pretest, the three groups were also equivalent with respect to narrative complexity scores based on two separate picture book prompted story generations. The EINC score for Frog Story narratives did not differ significantly between children assigned to the PT condition (M = 7.63, SD = 4.44), the PP condition (M = 9.00, SD = 4.21), and the CG condition (M = 7.88, SD = 4.12), $\chi^2(2) = .58$, $p = .749$. Similarly, the EINC score for the self-designed Climb Story did not differ significantly between children assigned to the PT condition (M = 8.25, SD = 4.86), the PP condition (M = 8.50, SD = 4.21), and the CG condition (M = 6.75, SD = 3.41), $\chi^2(2) = 1.34$, $p = .512$.

Tutor Groups

The data for the two tutor groups are presented in Table 12. Mann-Whitney U tests were conducted to detect significant differences between the groups. Tutors in the PT condition (5 boys, 3 girls) had a mean age of 4;11 and tutors in the PP condition (3 boys, 5 girls) had a mean age of 5;2.

Table 12. *Participant Characteristics for the PP and PT Tutors*

Variable	Group	*M*	*SD*	Range	*p*
Age in months	TPP	61.88	6.01	54-69	.561
	TPT	59.38	6.80	50-72	
Exposure to German in months[a]	TPP	42.01	11.20	25-50	.035*
	TPT	45.13	11.29	31-59	
Education mother in years	TPP	10.29	1.72	9-13	> .999
	TPT	10.57	1.25	9-13	
Education father in years	TPP	12.00	3.70	9-17	.718
	TPT	10.25	1.17	9-12	
Expressive language[b]	TPP	37.13	9.94	20-50	.371
	TPT	33.25	7.72	18-43	
Receptive language[b]	TPP	25.75	3.62	18-30	.833
	TPT	24.88	4.70	19-31	
Nonverbal intelligence[c]	TPP	19.00	3.16	9-20	.072
	TPT	15.50	5.66	9-28	
EINC Frog Story[d]	TPP	17.25	3.24	12-22	.494
	TPT	16.25	5.23	10-26	
EINC Climb Story[d]	TPP	14.75	1.49	13-17	.789
	TPT	15.75	3.45	13-24	

Note. TPP = Tutors Peer Play; TPT = Tutors Peer Tutoring; each group had 8 participants. Reported *p*-values refer to Mann-Whitney U tests.
[a]Based on parent report.
[b]Raw score sums, LiSe-DaZ expressive and receptive subtests (Schulz & Tracy, 2011).
[c]Raw scores, Raven Coloured Progressive matrices (Raven, 1995).
[d]Measures of narrative complexity based on generations of "Frog, where are you?" (Mayer, 1969); 'Climb Story' was a self-designed picture story. Narrative complexity measured using an adapted and extended version of the INC scoring rubric (Petersen, Gillam, & Gillam, 2008). The maximum score for each story was 26.
*Statistically significant with $p < .05$.

On average, tutors in the PT condition had a higher previous exposure to German (*Mdn* = 50.50) than tutors in the PP condition (*Mdn* = 27.50), as measured in months, $U = 12.00$, $z = -2.11$, $p = .035$. However, this difference did not translate to significant differences in German expressive and receptive language performance. Also, no significant differences emerged for age in months and nonverbal intelligence.

For the narrative productivity measures, the total number of words (TNW) produced by the PT tutors (PTT; *Mdn* = 117.00) was not significantly different from the amount produced by the PP tutors (PPT; *Mdn* = 178.50), $U = 25.50$, $z = -0.68$, $p = .495$, similar

to the total number of produced C-units, TNCU: PTP *Mdn* = 21.50, PPT *Mdn* = 33.50, *U* = 15.00, *z* = -1.79, *p* = .073. The measures of lexical diversity, namely number of different words in lemmas (NDW) and the vocabulary diversity statistic (VOCD), respectively, did also not differ significantly between the two groups (NDW: PTP *Mdn* = 39.50, PPT *Mdn* = 52.50, *U* = 23.50, *z* = -0.89, *p* = .372; VOCD PTP *Mdn* = 20.82, PPT *Mdn* = 24.48, *U* = 24.00, *z* = -0.46, *p* = .643). Furthermore, there was no significant difference for syntactic complexity, as assessed by mean length of C-unit (MLCU) between the PTT (*Mdn* = 5.31) and the PPT (*Mdn* = 5.18), *U* = 26.00, *z* = -0.63, *p* = .528.

The tutor groups were also equivalent with respect to narrative complexity scores based on two separate picture book prompted story generations. The EINC score for Frog Story narratives did not differ significantly between tutors in the PT (*Mdn* = 15.00) and tutors in the PP condition (*Mdn* = 17.50), *U* = 25.50, *z* = -0.69, *p* = .494. Similarly, the EINC scores assigned for the production of the self-designed Climb Story did not differ between tutors in the PT (*Mdn* = 15.00) and tutors in the PP condition (*Mdn* = 14.50), *U* = 29.50, *z* = -0.27, *p* = .789.

Comparison of Frog Story Narrative Performance of Tutee and Tutor Groups

A tutee-tutor comparison (Mann-Whitney U test) of the Frog Story narratives revealed significant differences in all of the computed microstructural measures and the overall EINC score, such that tutors outperformed the tutees on the group level (see Table 13).

Comparison of Performance on the Self-Designed Story of Tutee and Tutor Groups

The comparison (Mann-Whitney U test) of the narration of the self-designed picture story at pretest revealed significant differences in all measures of narrative microstructure, except for a measure of productivity, namely total number of C-units. There was also a significant difference narrative complexity (EINC), such that tutors outperformed the tutees on the group level (see Table 14). VOCD was not compared, because it could not be computed for half (*n* = 4) of the tutee narratives due to limited story length.

Table 13. *Comparison of Pretest Frog Story Narrative Performance of Tutees and Tutors*

Variable	Group	*M*	*SD*	*Range*	*p*
TNW	Tutees	85.04	52.31	11-199	.004*
	Tutors	163.63	86.99	42-358	
TNCU	Tutees	27.48	32.83	7-172	.034*
	Tutors	31.94	13.94	16-65	
NDW	Tutees	27.52	13.65	4-50	.001*
	Tutors	53.13	24.01	19-104	
VOCD	Tutees[a]	13.86	6.22	1.96-25.46	.014*
	Tutors[b]	23.14	9.99	11.36-46.23	
MLCU	Tutees	3.89	1.37	1.00-6.63	.016*
	Tutors	5.03	1.36	2.21-7.75	
EINC	Tutees	8.17	4.11	3-17	< .001*
	Tutors	16.75	4.23	10-26	

Note. Tutees $n = 24$, tutors $n = 16$.
TNW = total number of words; TNCU = total number of utterances in C-units; NDW = total number of different words in lemmas; VOCD = vocabulary diversity; MLCU = mean length of C-units in words; EINC = Extended index of narrative complexity.
[a]$n = 16$.
[b]$n = 15$.
*Statistically significant with $p < .05$.

Table 14. *Comparison of Pretest Self-Designed Story Performance of Tutees and Tutors*

Variable	Group	*M*	*SD*	*Range*	*p*
TNW	Tutees	38.92	29.56	5-139	.040*
	Tutors	48.99	27.27	19-132	
TNCU	Tutees	8.46	4.01	4-21	.573
	Tutors	8.75	4.30	5-23	
NDW	Tutees	19.13	9.87	2-45	.034*
	Tutors	24.56	8.15	12-41	
MLCU	Tutees	4.22	1.30	1.25-6.63	.003*
	Tutors	5.38	0.92	3.17-7.00	
EINC	Tutees	7.83	4.09	2-16	< .001*
	Tutors	15.25	2.62	13-24	

Note. Tutees $n = 24$, tutors $n = 16$.
TNW = total number of words; TNCU = total number of utterances in C-units; NDW = total number of different words in lemmas; MLCU = mean length of C-units in words; EINC = Extended index of narrative complexity.
*Statistically significant with $p < .05$.

Summary of Preliminary Analysis

In sum, the three groups of tutees in the intervention/control conditions did not differ from each other in any of the computed measures, i.e., all groups were comparable in terms of German language skills, nonverbal intelligence, home environment measures, and narrative performance. The two groups of tutors did also not differ significantly from each other, except for months of German language exposure. Furthermore, as to be expected, marked differences in narrative competence surfaced between children assigned to the tutee and tutor groups.

9.2 Intervention Effects on Tutees – Pre-Posttest Comparisons

To follow up on the first research question, *To what extent does engaging a peer tutor in a narrative-based language intervention improve the tutee's generation of fictional narratives?*, three areas were explored. Firstly, the narrative productions of the familiar Frog Story at posttest were compared across PT tutors, PP tutors, and CG participants.

9.2.1 Narrative Measures

Frog Story productions were compared for differences in microstructure, narrative complexity (EINC), as well as for differences in the use of the individual components of the EINC.

Narrative Microstructure

In the area of narrative microstructure, a Kruskal-Wallis H test revealed no significant differences between the three groups for productivity, with a mean total number of words (TNW) of $M = 213.88$ ($SD = 87.62$) for PT, $M = 154.00$ ($SD = 79.83$) for PP, and $M = 103.75$ ($SD = 69.43$) for CG, $\chi^2(2) = 5.51$, $p = .064$; similar to the total number of produced C-units (TNCU: PT, $M = 36.75$, $SD = 9.22$; PP, $M = 30.14$, $SD = 12.59$; CG, $M = 28.00$, $SD = 15.93$), $\chi^2(2) = 2.68$, $p = .226$, and the measure of syntactic complexity, namely mean length of utterance (MLCU: PT, $M = 5.63$, $SD = 1.08$; PP, $M = 4.01$, $SD = 1.89$; CG, $M = 3.64$, $SD = 1.60$), $\chi^2(2) = 4.91$, $p = .086$.

Meanwhile, there was a statistical difference for a measure of lexical diversity. While group assignment did not significantly affect VOCD with a mean performance of $M = 18.75$ ($SD = 6.11$) for PT, $M = 17.65$ for PP, and $M = 16.25$ ($SD = 10.71$) for CG, $\chi^2(2) = 0.92$, p = .995, the number of different words in lemmas (NDW) differed significantly between groups with a mean rate of $M = 59.50$ ($SD = 20.17$) for PT, $M = 47.86$ ($SD = 18.77$) for PP, and $M = 31.75$ ($SD = 15.51$) for CG, respectively, $\chi^2(2) = 6.97$, $p = .031$.

A subsequent Mann-Whitney U test revealed a statistical difference between participants in the PT ($Mdn = 54.00$) and children in the Control condition ($Mdn = 27.00$) at a bonferroni-corrected significance level of .0167, $U = 8.00$, $z = -2.52$, $p = .012$, $r = -.63$. Neither did the number of different lemmas produced by children in the PP condition ($Mdn = 48.00$) differ from the performance of the PT group ($U = 21.00$, $z = -0.81$, $p = .416$, $r = -.20$), nor from the performance of participants in the Control condition ($U = 13.00$, $z = -1.74$, $p = .082$, $r = -.44$).

Narrative Complexity (EINC)

A Kruskal-Wallis H test computed a statistically significant difference in narrative complexity at posttest between the experimental groups, with a mean Frog Story EINC score of $M = 17.00$ ($SD = 5.13$) for PT, $M = 9.75$ ($SD = 4.03$) for PP, and $M = 8.50$ ($SD = 4.93$) for CG, $\chi^2(2) = 9.36$, $p = .009$. That is, Frog Story narrative complexity was significantly affected by group assignment.

Subsequently, Mann-Whitney U comparisons were conducted to post hoc follow up on the origin of the difference. Bonferroni adjustments were applied, such that all effects are reported at a .0167 level of significance. At posttest, narrative complexity of tutees in the Peer Tutoring group ($Mdn = 17.50$) was significantly higher than narrative complexity of tutees in the Peer Play group ($Mdn = 9.00$), $U = 8.00$, $z = -2.53$, $p = .011$, $r = -.63$, as well as children in the Control group ($Mdn = 8.00$), $U = 7.00$, $z = -2.64$, $p = .008$, $r = -.66$. However, it appeared that narrative complexity was not different between participants in the PP and in the CG, $U = 25.00$, $z = -0.74$, $p = .461$, $r = -.19$.

Figure 16 displays mean narrative complexity scores for all three experimental groups at pre- and posttest.

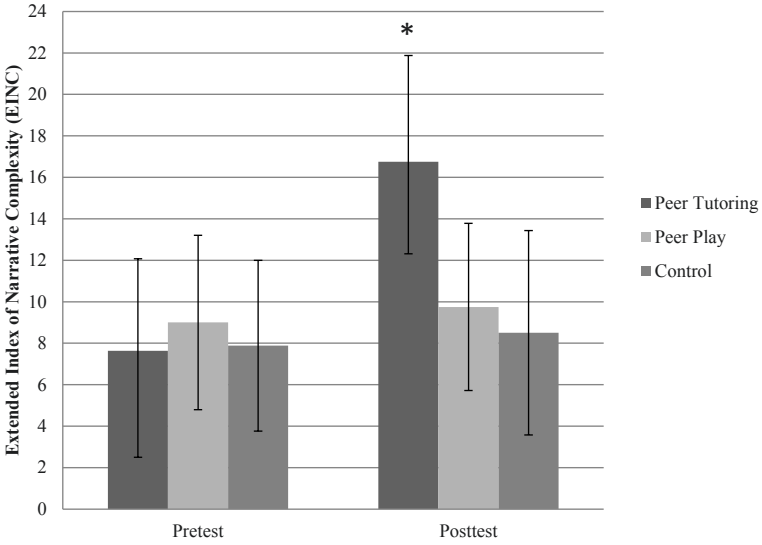

Figure 16. Tutees' Pre-Posttest Changes in Frog Story Narrative Complexity (Means and SD). *Note.* All groups were $n = 8$. Frog Story narrative complexity was based on generations of "Frog, where are you?" (Mayer, 1969). EINC represents composite scores; the maximum score was 26. Error bars represent standard deviation.

In-Depth Analysis of Tutee Differences in Frog Story Narrative Complexity Measures

To further explore the differences in tutees' narrative complexity, cohesive and evaluative language elements of the EINC, as derived from the Frog Story narratives, were analyzed individually to detect specific areas of growth. As the number of points to be reached for these individual EINC components only ranged between 0, 1, and 2, a chi-square-test of independence (with a Bonferroni adjustment applied) was performed to detect statistical differences between groups.

While at pretest, none of the individual EINC items were statistically different between the three experimental conditions (see Table E.1 in Appendix E), a chi-square test performed at posttest found a significant relationship between group assignments

and use of temporal markers, such that children in the PT group used more temporal markers than tutees in the PP group, $X^2(1, N = 16) = 8.77$, $p = .012$. At a bonferroni-corrected significance level of .0167, the difference between PT and the Control group, $X^2(1, N = 16) = 6.57$, $p = .037$, was not significant. Also, the use of metacognitive verbs was higher in the PT group than in the PP group, $X^2(1, N = 16) = 6.56$, $p = .010$, and in the CG, $X^2(1, N = 16) = 7.27$, $p = .007$. While the use of emotional state terms and physical state terms did not differ significantly between PT and PP groups, $X^2(1, N = 16) = 1.76$, $p = .185$, and $X^2(1, N = 16) = 0.71$, $p = .398$, respectively, the use of emotional state terms use was higher in the PT than in the CG condition, $X^2(1, N = 16) = 12.44$, $p < .001$ (physical state terms: $X^2(1, N = 16) = 5.33$, $p = .021$). As for the expression of intent, the performance of the PT group did not differ significantly from either PP, $X^2(1, N = 16) = 0.54$, $p = .464$, or CG condition, $X^2(1, N = 16) = 4.00$, $p = .046$.

9.2.2 Narrative Examples

The mean pre- to posttest scores show an increase in narrative skills (complexity) in tutees. In this section, excerpts from the Frog Story narratives of two tutees offer qualitative insight into individual growth patterns. These two children, Dilara and Fatima[67], were both successive DLLs who were first systematically exposed to German in ECEC and they were both tutees in the Peer Tutoring group. Dilara was 4 years and 5 months old and started the intervention with minimal German abilities after having had 13 months of previous exposure to German in the ECEC setting. Fatima was 5 years and 6 months old and, after an exposure time of 25 months, was relatively fluent and comfortable in speaking German.

At pretest, Dilara mostly names the characters on the pictures. While her minimal German skills certainly contribute to her not creating linguistically rich and detailed narratives, there is no evidence of a story plot or drawn connections between the depicted events. On Westby's (2005) Story Grammar Decision Tree, Dilara's narrative represents the lowest story structure level, a descriptive sequence:

[67] Names were changed to protect participants' identity.

Exerpt[68] 1: Dilara at pretest

> Frosch. Frosch eine Hund. Hund weg hier. Die Mama! Die Mama, die Baby! Die Baby. Der eine Hund. Ja! Die Mama, die Papa, die kriegen Babyfrosch! Oh, die die Baby große. Baby. Hoppa! Die Baby. Die Baby. Hund. Hund. Junge. Zwei Junge zwei Hunde. Was das? Eine Biene. Biene. Biene. Biene. Auch Biene. Biene. Nein! Zwei Junge. Eine Hund. Zwei Junge. Zwei eine Hund. Eine Biene. Eine Hund, eine Junge. Eine Hund, eine Junge. Eine Junge, eine Hund. Zwei Junge, zwei Hunde. Ein Hunde, eine Junge. Ein Frosch, viele Frosch. […]
>
> *Frog. Frog a dog. Dog away here. The mom! The mom, the baby! The baby. He one dog. Yes! The mom, the dad, they have baby frog! Oh, the the baby big. Baby. Oop! The baby. The baby. Dog. Dog. Boy. Two boy, two dogs. What that? A bee. Bee. Bee. Bee as well. Bee. No! Two boy. A dog. Two boys. Two a dog. A bee. A dog, a boy. A dog, a boy. Two boy, two dogs. A dog, a boy. A frog, many frog. [...]*

Meanwhile, at posttest, most children began to include more evaluative words in their stories and to include more crucial story elements. For example, Dilara's narrative is longer overall and clearly more detailed. The disappearance of the frog is identified a clear search pattern is evident. Also, she now frequently uses additive and temporal markers to connect her utterances and includes direct speech to paint a more vivid verbal picture of their narratives and to bring out the characters' perspectives. Overall, her narrative progressed from a descriptive sequence to an action sequence.

Excerpt 2: Dilara at posttest

> Die Hund, die Junge. Und die Hund hat die da ein Frosch hat. Und dann die und die äh die schlaft. Die Hund und die Junge, die noch schlaft. Und dann die Frosch hopp machen. Und dann die hüpft an die Seite. Und dann guckt, weg Frosch. Die äh die Hund guckt hier, auch nicht. Und die äh Junge hier, auch guckt nicht. Hier auch nicht. Hier auch nicht. Hier auch nicht. Und die Junge hat „Was macht du?" sagst. „Und dann mach ich so." sags. Und dann hier guckt und dann da hier da das. Und dann die Biene auch nicht. Die weiß nicht, wo da die Frosch. Guckt hoch. Hier auch nicht. Die weg die Frosch. Da guckt, auch nicht. Und dann hier guckt, auch nicht. […]
>
> *The dog, the boy. And the dog has a frog there has. And then they und they um they sleeps. The dog and the boy, they still sleeps. And then the frog oop make. And then it hops to the side. And then looks, away frog. The um the dog looks here, also not. And the um boy here, also looks not. Also not here. Also not here. Also not here. And the boy has „What are you doing?" says. „And then I do like this", says. And then here look and then there here this one. And then the bee also not. It does not know, where there the frog. Looks up. Also not here. It away the frog. There looks, also not. And then here looks, also not. [...]*

[68] To facilitate narrative cohesion, each example is presented in a narrative format, rather than in C-units, and has been edited for punctuation. Mazes, which were excluded from microstructural measures, are still included here.

Children with greater initial German abilities, such as Fatima, were already able to connect several sentences together to construct a narrative, and their narratives included more events to move the plot forward. However, Fatima loses herself in details when describing the setting, so that it is hard to make out the elements and events central to the story.

Excerpt 3: Fatima at pretest:

> Da sieht ein Mensch und eine Frosch und eine Hund. Eine Bett und eine Lampe. Und Fenster. Und von die Junge Schuhe. Und das ist das T-Shirt. Und das Kiste bei ihrem Bett. Das Junge schlaft. Und ihre Schuhe liegen. Und ihre diese Schuhe liegen. Das Frosch geht von den Glas. Das eine Socke. Ja. Das die Socke liegt da und ihre T-Shirt und ihre Bett. Es ist immer noch Mond. Und das Frosch hat nicht geschlaft. Und das Hund und das Junge geschlaft. Und das äh das Fenster ist da. Und das Lampe ist da. Und das Stuhl ist da. Als das Junge aufgewacht hat, hat die bei den Glas geguckt. Ist das weggegangen. Und ihre Hund. Und ihr das Junge. Und das war noch ihre Socke war da immer noch. Ihre Bett. Und das war noch da. Und das ist nicht abgerutscht. Und und ihre zwei Schuhe die liegen da. Und ihre T-Shirt und das Glas und das. Das Sonne ist irgendwo anders. Und als das Mond da war ist das verschwunden. Äh und dann das liegt noch das Glas. Und die beiden Schuhe immer noch. Und die Fenster und die Lampen und die Glas. Und das Stuhl. Das Hund hat das Glas. Das Hund hat das Glas. Und das Junge zieht sich was irgendwas an. Und da ist das T-Shirt mit das Hose und mit das beide Schuhe. Und mit das Stuhl und mit das Hund. Das Hund hat das Glas und das Lampe sieht noch. Und da sie xx Blatt. Und das Junge hat das Glas. Und das Junge ruft das Frosch. Und das Fenster sieht. […].
>
> *There see a person and a frog and a dog. A bed and a lamp. And window. And from the boy shoes. And this is the t-shirt. And this box by her bed. The boy sleeps. And her shoes lie. And her these shoes lie. The frog goes from the jar. This a sock. Yes. The the sock lies there and her t-shirt and her bed. It is still moon. And the frog has not slept. And the boy and the boy slept. And the um the window is there. And the lamp is there. And the chair is there. When the boy woke up, he looked by the jar. It went away. And her dog. And her the boy. And that was her sock was there still. Her bed. And that was still there. And that did not slide down. And and her two shoes they lie there. And her t-shirt and the jar and that. The sun is somewhere else. And when the moon was there it vanished. Um and then it still lies the jar. And both of the shoes still. And the window and the lamp and the jar. And the chair. The dog has the jar. The dog has the jar. And the boy put something on. And there is the t-shirt with the pants and with both of the shoes. And with the chair and with the dog. The dog has the jar and the lamp still sees. And there she xx leaf. And the boy has the jar. And the boy calls the frog. And the window sees. [...]*

At posttest, many of the advanced DLL children, such as Fatima, provided more events and advanced story structures and frequently used direct speech, painting a more vivid verbal picture of their narrative. According to Westby's (2005) binary de-

cision tree, some of the stories would be classified reactive sequences, or episodes of varying elaborateness, as some children already included causal connectors and a clear goal, attempt, and outcome. Also, Fatima's uses of evaluative words and the emotional state term böse *[upset/angry]* add depth to her characters. Distinct from her pretest narrative, now the characters, setting, and initiating event are clearly identifiable.

Excerpt 4: Fatima at posttest:

> Dann hatte der Hund und ein Junge ein Frosch gefangen. Das war nachts. Dann musste der in Bett. Dann war der in Bett und hat geschlafen. Und das Frosch war weg. Und als die aufgewacht haben, haben die geguckt. „Das Frosch ist weg" hat das Junge gesagt. Dann hatte die zum Schuhe geguckt, aber da war der nicht. Dann hat der geschlafen. Dann waren die im Fenster und haben geruft „Wo bist du, Frosch?". Dann ist der runtergefallen, Hundi. Dann waren die böse, dass der runtergefallen. Dann hat sie gesagt „Hundi, lass mich abzulecken!". Dann waren die da und hatten gesagt „Frosch, Frosch!". Dann hatten die nicht die gehört und nicht gefunden. Dann hatte der „Frosch, Frosch, komm' doch mal raus wo du versteckt hast!". […]
>
> *Then the dog and a boy had caught a frog. That was at night. Then he had to go to bed. Then he was in bed and was sleeping. And the frog was gone. And when they woke up, they looked. „The frog is gone", did the boy say. Then he looked to the shoes, but he was not there. Then this one slept. Then they were in the window and called, „Where are you, frog?". Then he fell down, doggy. Then they were angry, that he fallen down. Then she said, „Doggy, stop licking me!". Then they were there and had said, „Frog, frog, come on out where you are hiding!" [...]*

9.2.3 Generalization Probe

Additional to the Frog Story, children narrated the self-developed "Soccer Story" (see Appendix A) at posttest, which was analyzed for narrative complexity via EINC. Even though differences could be detected descriptively (see posttest, Fig. 17), there was no statistically significant difference between the EINC scores by different group assignment with a mean score of $M = 13.38$ for PT ($SD = 2.26$, $Mdn = 14.50$), $M = 10.38$ for PP ($SD = 4.78$, $Mdn = 10.00$), and $M = 7.50$ for CG ($SD = 4.41$, $Mdn = 6.00$), $\chi^2(2) = 5.74$, $p = .057$. For further interpretation of these results, it should be noted that a Wilcoxon Signed-Ranks Test indicated that narrative productivity (number of words produced) ranks for Frog Story narratives were significantly higher ($M = 157.35$, $SD = 89.02$) than median productivity ranks for the stories produced in response to the

seven-page long unfamiliar picture book in the generalization probe ($M = 44.63$, $SD = 22.97$), $Z = -4.20$, $p < .001$.

9.3 Long-Term Intervention Effects on Tutees

The second research question to be explored was: *To what extent do any improvements in preschoolers' narrative performance maintain following a period of 5 weeks with no intervention?* The maintenance probe assessed narrative complexity and was collected via an unfamiliar wordless picture book that was part of the self-developed materials and occurred after a 5 week no-intervention period following posttest (see Figure 17).

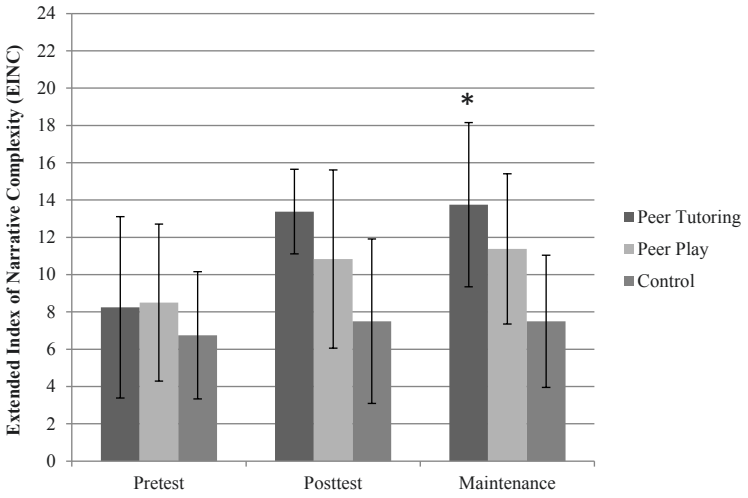

Figure 17. Self-Developed Picture Stories: Tutees' Narrative Complexity Performances at Pretest, Posttest, and Maintenance Probe (Means and SD).
Note. All groups were $n = 8$. Measures were based on generations on self-developed stories. EINC represents composite scores; the maximum score was 26. Error bars represent standard deviation.

For the EINC composite score, a Kruskal-Wallis H test yielded statistically significant group differences ($X^2(2) = 6.90$, $p = .032$), with a mean score of $M = 13.75$ for PT ($SD = 4.40$), $M = 11.38$ for PP ($SD = 4.03$), and $M = 7.50$ for CG ($SD = 3.55$), which

were followed up by group comparisons. It was revealed that narrative complexity measure of the PT group (*Mdn* = 12.50) was significantly higher than narrative complexity of children in the CG condition (*Mdn* = 6.00), *U* = 7.50, *z* = -2.59, *p* = .010, *r* = -.65, while comparisons between PT and PP children (*Mdn* = 12.50) yielded no differences, *U* = 27.00, *z* = -0.53, *p* = .596, *r* = .13. Similarly, results revealed that narrative complexity was not different between participants in the PP and in the CG, *U* = 15.00, *z* = -1.80, *p* = .073, *r* = -.45.

9.4 Intervention Effects on Tutors

The third research question (*Which effect does the intervention have on children serving as the tutors?*) concerned the performance of the tutors, that is, if narrative complexity measures of tutors of the Peer Tutoring (PTT) and Peer Play (PPT) groups would change through the intervention process. To assess this question, narrative performances of PT and PP tutors were compared at posttest and at maintenance probe.

Frog Story Microstructure

At posttest, all microstructural measures derived from the Frog Story narratives were compared between the two tutor groups. For the productivity measures, the total number of words (TNW) produced by the PT tutors (PTT; *Mdn* = 158.50) was not significantly different from the amount produced by the PP tutors (PPT; *Mdn* = 225.00), *U* = 21.00, *z* = -1.16, *p* = .248; similar to the total number of produced C-units (TNCU: PTP *Mdn* = 26.50, PPT *Mdn* = 37.00), *U* = 19.50, *z* = -1.32, *p* = .188. The measures of lexical diversity, namely, the number of different words in lemmas (NDW) and the vocabulary diversity statistic (VOCD), respectively, also did not differ significantly between the two groups (NDW: PTP *Mdn* = 53.50, PPT *Mdn* = 67.00, *U* = 25.00, *z* = -0.74, *p* = .462; VOCD PTP *Mdn* = 25.12, PPT *Mdn* = 21.07, *U* = 25.00, *z* = -0.74, *p* = .462). Finally, there was no significant difference for mean length of C-unit (MLCU) between the PTT (*Mdn* = 6.13) and the PPT (*Mdn* = 6.03), *U* = 32.00, *z* = 0.00, *p* > .99.

Frog Story Narrative Complexity (EINC)

Furthermore, group differences in narrative complexity for the Frog Story productions, as determined by EINC score, were explored. As displayed in Figure 18, the mean performance between pre-and posttest shows an upward trend in the PT tutors. However, a Mann-Whitney U test comparing Frog Story narrative complexity scores at posttest did not reveal a significant difference between the PT tutors (*Mdn* = 18.00) and the PP tutors (*Mdn* = 16.50), $U = 23.50$, $z = -0.90$, $p = .368$.

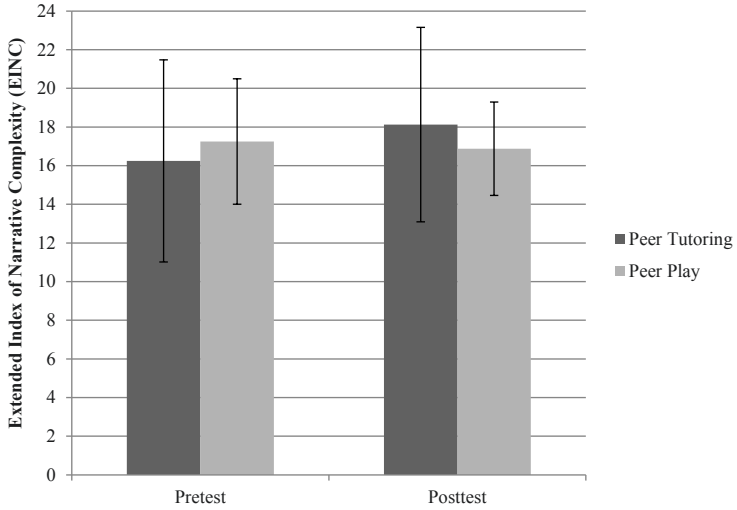

Figure 18. Tutors' Pre-Posttest Changes in Frog Story Narrative Complexity (Means and SD). *Note.* All groups were $n = 8$. Frog Story Narrative complexity was based on generations of "Frog, where are you?" (Mayer, 1969). EINC represents composite scores; the maximum score was 26. Error bars represent standard deviation.

Self-Developed Story Books – Microstructure at Posttest and Maintenance Probe

First, at posttest, all microstructural measures derived from the self-developed story book were compared between the two tutor groups. For the productivity measures, the total number of words (TNW) produced by the Peer Tutoring tutors (PTT *Mdn* = 50.50) was not significantly different from the amount produced by the PP tutors (PPT *Mdn* = 59.50), $U = 23.50$, $z = -0.89$, $p = .371$; similar to the total number of

produced C-units (TNCU: PTP *Mdn* = 9.00, PPT *Mdn* = 9.00), *U* = 30.50, *z* = -0.16, *p* = .871. The computed measure of lexical diversity, namely, the number of different words in lemmas (NDW), also did not differ significantly between the two groups (PTP *Mdn* = 26.00, PPT *Mdn* = 27.50, *U* = 26.50, *z* = -0.58, *p* = .562). Finally, there was no significant difference for mean length of C-unit (MLCU) between the PTT (*Mdn* = 6.11) and the PPT (*Mdn* = 6.14), *U* = 24.50, *z* = -0.79, *p* = .431.

As for the maintenance probe, tutors also maintained their performance. No significant differences emerged between tutors in the PT and tutors in the PP condition. For the productivity measures, the total number of words (TNW) produced by the Peer Tutoring tutors (PTT *Mdn* = 64.00) was not significantly different from the amount produced by the PP tutors (PPT *Mdn* = 57.00), *U* = 25.50, *z* = -0.68, *p* = .495; similar to the total number of produced C-units (TNCU: PTP *Mdn* = 9.50, PPT *Mdn* = 9.50), *U* = 22.50, *z* = -1.01, *p* = .311. The computed measure of lexical diversity, namely, the number of different words in lemmas (NDW), also did not differ significantly between the two groups (PTP *Mdn* = 31.50, PPT *Mdn* = 26.50, *U* = 17.00, *z* = -1.59, *p* = .112). Finally, there was no significant difference for mean length of C-unit (MLCU) between the PTT (*Mdn* = 6.76) and the PPT (*Mdn* = 6.35), *U* = 26.50, *z* = -0.58, *p* = .563.

Self-Developed Story Books – Narrative Complexity at Posttest and Maintenance Probe

Furthermore, group differences in narrative complexity for the narratives collected with the self-developed story books at posttest and maintenance probe, as determined by EINC score, were explored (see Figure 19). Mann-Whitney U tests comparing narrative complexity scores at posttest did not reveal a significant difference between the PT tutors (*Mdn* = 14.00) and the PP tutors (*Mdn* = 15.00), *U* = 25.00, *z* = -0.75, *p* = .450, similar to the results at maintenance probe, PT tutors (*Mdn* = 15.50), PP tutors (*Mdn* = 14.00), *U* = 18.00, *z* = -1.50, *p* = .133.

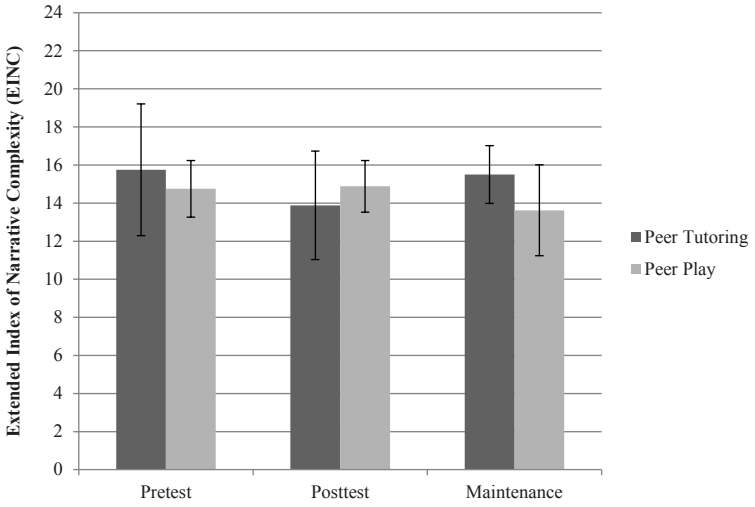

Figure 19. Self-Developed Story Books: Tutors' Narrative Complexity Performances at Pretest, Posttest, and Maintenance Probe (Means and SD).
Note. All groups were $n = 8$. Measures were based on generations on self-developed stories. EINC represents composite scores; the maximum score was 26. Error bars represent standard deviation.

In summation, study participation did not significantly affect tutors' narrative performance.