

The Macrostructure of elicited narratives by bilingual children. An analysis of the factor Age of Onset

Die Makrostruktur elizitierter Narrative bilingualer Kinder: Eine Analyse des Faktors Age of Onset

Kristin Haake

Humboldt-Universität zu Berlin

In this paper, data from a current study on bilingual language acquisition and language promotion of children is presented. 96 narratives from 32 Turkish-German and Russian-German bilingual children were examined with regard to the acquisition of narrative ability in three rounds of tests. The macrostructure of each narrative was evaluated based on the theories of Westby (2005), Stein and Glenn (1977) and Gagarina et al. (2012). In the quantitative analysis, the factor age of onset (AoO) was considered and therefore, two hypotheses were introduced: 1) There is an influence of AoO on the narrative ability of *L2* German bilingual children. And 2) The narrative ability will converge over time and after three years there will be no difference between the groups. Neither of those hypotheses could be confirmed by the examined narrative data. Hence, other influences on narrative ability were discussed in the last chapter and prospects for further research were given. In sum, the article shows that more narrative data of these children should be collected to make a comprehensive conclusion about the influence of AoO on narrative ability.

1 Introduction

In the last years, the *second language acquisition* (*L2A*) research has developed different assessment instruments to measure the *narrative ability* (*NAY*) of bilingual children. Many factors that can influence *NAY* are not examined yet. Therefore, a quantitative analysis of the factor *Age of Onset* (*AoO*) on successful *NAY acquisition* (*NAYA*) will be performed in this article. In research, the factor *AoO* is mostly considered in *L2A* of adults. But here, the influence of *AoO* on simultaneously and sequentially bilingual children in case of *NAY* will be

examined. The data was collected during a study of the *Berliner Interdisziplinärer Verbund für Mehrsprachigkeit* (BIVEM) and provides narratives of Russian-German and Turkish-German bilingual children. Those narratives were recorded in preschools in multilingual neighborhoods in Berlin for three years. The goal of *BIVEM* is to find language promotion tasks for bilingual children that can assist the teachers in preschools. Moreover, the children will be prepared for the upcoming daily tasks in school.

The analysis of narrative data is adequate for the examination of *narrative* as well as *grammatical ability*. Therefore, narrative data is a valid object of study. In this article, the *NAy* and the influence of factor *AoO* will be examined concerning the research question:

- Is there an influence of *AoO* on *NAy* of bilingual children in German narratives?

In order to statistically analyze the question, two hypotheses were formulated:

1. There is an influence of *AoO* on *NAy* of *L2* German bilingual children.
2. *NAy* will converge over time and after three years there will be no difference between the simultaneously and sequentially bilingual *NAy*.

To examine the two hypotheses, the narrative data was divided into two groups. The groups represent simultaneously and sequentially bilingual children. In group 1, the children's *AoO* was between 0;0 and 1;11 years, while in group 2 it was between 2;0 and 2;11 years. This division is performed due to Rubergs (2013) theory on language acquisition (LA). Each one of the children produced three narratives in three years. There was a one-year break between each round of tests.

Both hypotheses are based on L2A theories and studies on that topic. Due to Singleton and Ryans (2005: 61) rule: "The younger = the better in the long run" the first hypothesis was postulated. Moreover, Dimroth (2007) claimed that for a successful LA the factor *age* is relevant. Since no difference between simultaneously and sequentially *L2A* was found in studies on *L2A* of children and adults (Abrahamsson & Hyltenstam, 2009), the second hypothesis was postulated. In this case there should be quantitatively no difference between the performances of both groups in the last round of tests. The article's goal is to examine both hypotheses. Therefore, the theoretical background on narratives and *LA* will be presented in chapter 2 as well as the terms which are used in the article. Firstly, the term narrative and the narrative structure will be introduced. Different assessment tests for narrative data collection and analysis will be presented as well as the MAIN tool, which has been used in the current data collection.

Secondly, theories about N_{Ay} acquisition, bilingual L_A as well as the factor AoO will be introduced. In chapter 3, the method for data collection and statistically analyses will be presented and the methodological choice will be illustrated. The main part of this article is the presentation of the results and their discussion regarding the two hypotheses in chapter 4 and chapter 5. Finally, ideas for further analyses will be suggested in chapter 5.3.

2 Theoretical Background

In this chapter the theoretical background for the analysis of children's N_{Ay} will be given. For this purpose, the usage of the term narrative in the context of this work will be discussed in relation to the possible types of stories and the underlying structure model. Further, the actual state of the research on L_{2A} of bilingual children will be presented and the factor AoO will be illuminated.

2.1 Narratives

The term narrative will be used for a special kind of story in this article. Engel (1995) claimed that a narrative should be seen as a report of experiences and events that a protagonist has undergone including his inner perception of this situation. In Engel's assumption the protagonist of a story does not have to be humanoid and real, but can also be animalistic and fictional. The form of the narrative can differ between written and oral narration (Hughes et al., 1997). In this article verbal narratives will be used as basis to examine the N_{Ay} of bilingual children in their L₂. Since storytelling is a frequently used task in German preschools, even very young children might have a stable foundation in the N_{Ay} (Uchikoshi, 2005). The story telling part was elicited by pictorial stimuli. These were designed such that they are comparable due to their underlying *story structure* (Gagarina et al., 2012). One is the *macrostructure*, which is the shaping structure of a narrative where the plot of the story is embedded. It is based on the *story grammar model* (SGM) by Stein and Glenn (1977) which contains the setting (S) of the narrative and the *episode structure* (ES). According to Trabasso and Nickels (1992) these elements are language independent and therefore, it should be adequate to assess the general N_{Ay} of bilingual children in this task (Paradis et al., 2010).

Setting (S) in Stein and Glenn's (1977) *story structure model* (SSM) introduces the protagonist, the time and the location of the narrative (Fiesta, 2008). In the version which is used in this study, Gagarina et al. (2012) decided to count only time and location. The episodic structure (ES) is based on Westby's (2005) binary decision tree – see Appendix 1 for the detailed scheme – which differentiates between complete and incomplete episodes. Complete

episodes (GAO) contain the goal (G) of the protagonist, the attempt (A) and the outcome (O). The incomplete episode contains only one or two of them. In different studies, further *ES* structure elements are complemented (Fiestas, 2008; Uchikoshi, 2005 indicates Chang, 2004; Labov, 1972; Peterson & McCabe, 1983). For example, Gagarina et al. (2012) added *Internal State Terms (IST)* to their *ES* to refer to the episode's initial event and the reaction of the protagonist. *IST* are verbal expressions (e.g. happy, shout, cry, angry) to refer to inner and mental states of the protagonist (Grazzani & Ornaghi, 2012; Symons, 2004). Further, they show the developmental status of the children's Theory of Mind (Lorusso et al., 2007; Tomasello, 2003) which is relevant for the development of *NAY* to express the inner state of the protagonist (Gagarina et al., 2012; Schneider et al., 2006).

To collect narrative data there are various tests developed. One of the first was Renfrew (1969), who introduced the pictorial stimuli the *Bus Story* to examine British children. The method was extended by Glasgow and Cowley (1994) to test US-American children between 3;0 and 6;11 years. This material is not normalized for any other cultural and language background and therefore not usable for bilingual children. In the last ten years further tests were generated (Gillam & Pearson, 2004; Gagarina et al., 2012). For example, the *Test of Narrative Language (TNL)* by Gillam and Pearson (2004) contains pictorial stimuli to gather narrative data of bilingual children between 5;0 and 12;0 years. Equally, the *Edmonton Narrative Norms Instrument (ENNI)* was developed for children between 5;0 till 9;0 years with specific language impairment (Schneider et al., 2005). This test provides two pictorial stimuli with simple or complex structures. Moreover, Gagarina et al. (2012) provide the *Multilingual Assessment Instrument for Narratives (MAIN)* which is used for the elicitation of the narrative data in the present article. *NAY* will be comparable through a score that measures every episode structural element. *MAIN* provides two sequential pictorial stimuli with six pictures each. Both sequences introduce *S* and three *GAOs* to give the children more than one possibility to produce the elements of the *SSM* (Stein & Glenn, 1977; Berman & Slobin, 1994). The underlying *ES* by Westby (2005) was reduced by Gagarina et al. (2012). In the end a *GAO* contains the following:

- (1) a) *ISTe*
 - b) *G*
 - c) *A*
 - d) *O*
 - e) *ISTr*

The pictorial stimuli were created to test children in different languages and from different cultural- and socioeconomic background (Gagarina et al., 2012). Altogether children can achieve 17 credits in each narrative. The score is divided into two credits for *S*, and one for every episode element. In the quantitative analysis a GAO will contain only G, A and O. In this article the total number of IST tokens per narrative will be examined. MAIN was chosen in the BIVEM study as instrument to examine NAY of bilingual children due to the comparability of NAY by quantitative analysis and the consideration of heterogeneous background.

2.2 Language Acquisition

The assumption that NAY is acquired equally by mono- and bilingual children (Peterson & McCabe, 1983) could be confirmed by previous research on the universality of the concept of macrostructure (Berman & Slobin, 1994; Fiestas, 2008; Strömquist & Verhoeven, 2004; Paradis et al., 2010). This assumption will be checked regarding the factor AoO.

NAY contains the understanding of SGM, the causal correlation, to take into account different perspectives, the planning of narratives, and even the assessment of protagonist's behaviour (Trabasso & Nickels, 1992; Trabasso & Rodkin, 1994). Studies on bilingual children's NAY have shown that they are grammatically able to produce narratives with 4;0 years (Slobin, 1985; Radford, 1995; Botting, 2002). Nevertheless, NAY will be acquired sequentially (Fiestas, 2008; Gillam & Johnston, 1992; Leadholm & Miller, 1992; Miller, 1991; Peterson & McCabe, 1983; Shapiro & Hudson, 1991). Children at age 3 are capable to label objects on pictures (e.g. objects, protagonists or actions) but without any order. Further, the inner states of the protagonists and causal correlations are not verbalized (Fiestas, 2008; Peterson & McCabe, 1983). Even though some children use *IST* in their narratives, they mainly do not refer on the inner state of the protagonist but their own wishes (e.g. 'I don't want to tell the story') (de Villiers, 2007). The narratives at this stage are short (Leadholm & Miller, 1992) and the story structure is less complex as well as incomplete (Peterson & McCabe, 1983; Shapiro & Hudson, 1991). Between 3;10 and 5;0 years the children usually achieve the cognitive maturation to interpret and name the inner state of other people (Miller, 2006). Furthermore, between 6;0 and 9;0 years the narratives of children are called classic narratives by Peterson and McCabe (1983: 36-41). Then, the children are able to describe sequentially the action and the inner state of the protagonist as well as they refer on causal links.

Even though NAY should be language independent, research on NAY on Spanish-English bilingual children has shown that the production of the *ES* elements was language dependent (Fiestas, 2008; Peña et al., 2006). Since the

factor AoO in *L2A* has not been considered in research on NAYA until now, this present article will investigate if there is an influence by AoO on the NAYA of Turkish-German and Russian-German bilingual children.

To analyse the influence of AoO on NAY, the data was divided into two groups. Group 1 had its first contact with German in the first 1;11 years of their life, whereas the children in group 2 between 2;0 and 2;11 years. The basis for the division at 2;0 underlies the theory on LA of Ruberg (2013). In both cases, the literature calls the children bilingual. Though, Ruberg (2013) distinguishes between simultaneously and sequentially L2A. Simultaneously L2A occurs between 0;0 and 1;11 years while sequentially L2A is proceeded between 2;0 and 3;11.

Table 1: Types of *LA*. Ruberg (2013: 182). Translated by the author.

Types of multilingual <i>LA</i>	<i>AoO</i> in years	<i>AoO</i> in years;months
Simultaneously	0–2	0;0–1;11
Sequentially	3–4	2;0–3;11
<i>L2A</i>	5–10	4;0–9;11

The differences are drawn due to the fact that the grammatical structure of L1 increases with AoO (Ruberg, 2013). Dimroth (2007) adds that these categories are exclusively for uncontrolled LA. Besides the factor *age* the input quality and quantity can influence the successful L2A (Ruberg, 2013).

AoO labels the time a child has the first contact to L2 (Birdsong, 2006). In different grammatical features the AoO has a considerable influence on LA (Dimroth, 2007). Singleton and Ryans (2004: 61) claimed in a neuroscientific study that there is evidence for the rule “The younger = the better in the long run”. It should be examined whether the non-grammatical factor affect the data in the study.

3 Method

In this section the data collection procedure and the methodological background of the analysis of the narrative data will be presented.

The data for the present article was collected in the longitudinal study of *BIVEM*. From 2012 to 2015 data of 160 children was gathered. To exclude confounding factors¹ we selected 32 children for the analysis in this article. Every child has told three narratives in three years with a one-year break.

¹ The criteria: homogenous age of children, number of narratives per children, AoO.

Regarding the theory of Ruberg (2013) the data was divided into two AoO groups.

The first narratives were produced by the children at the age of 3;4 to 4;4 years. To guarantee that the children have the grammatical ability to produce narratives, extensive language tests were conducted: productive and receptive lexicon (nouns, verbs), productive and receptive syntax (wh-words, sentence comprehension) and morphology (plural, case and inflexion). To collect the data, one of the two pictorial stimuli of Gagarina et al. (2012) was selected. The pictures of the BB were already shown in the previous section. It is based on the *Cat Story* (Cat) of Hickmann (2002) and was edited by Gagarina et al. (2012) in the context of MAIN regarding the children's background and the comparability of the *story structure* (SS). The second story – the *Baby Goats Story* (BG) – is based on the *Fox Story* (Fox) by Gülzow and Gagarina (2007). BG was modified in the MAIN project. Both stories include three GAOs which are presented on six pictures per story. Every episode contains: ISTe, G, A, O and ISTR.

The assessment of narrative data was conducted through external trained experimenters in preschools in Berlin. Every child was individually tested with one of the two stories of Gagarina et al., (2012). At the beginning of the test the children chose one of three envelopes with a pictorial stimulus in each. To pretend that the experimenter does not know the story. Before the children started telling their narrative they had time to see the whole story. Thereafter, the pictures were folded and the child could see only two of them. The narrative task was initiated by the experimenter's words: 'Erzähle mir die beste Geschichte von den Bildern.'² The narrative continuity should not be interrupted during the testing, but supportive questions and prompts (e.g. 'Anything else?', 'Continue', 'Tell me more', and 'Let's see what else happens in the story.') were allowed. Through these restrictions the comparability of the narratives should be protected. After the data collection, the audio files were transcribed in CHILDES CLAN (MacWhinney, 2000). Further details about the BIVEM study are presented in the Zwischenbericht (2014) of the project.

Based on the transcriptions, the setting and the story structure elements were scored. Example (2a) and (2b) shows two narratives, which mentioned the *time* aspect.

- (2) a) *CHI: Es waren einmal <ein &klei> [//] drei Schafe.
b) *CHI: &em es fängt an einen [: einem] Tag, [...].

Both examples received one point each for setting. The *location* was not named in any of the analyzed narratives. For every element of the episodic structure one

² Engl. translation: 'Tell me the best story you can see on the pictures.'

point was given so that a narrative could maximal have a story structure (SS) score of 17. 2 points for *S*, three times (for each episode) 5 points for each element (Gagarina et al., 2012). Thereafter, the GAO and incomplete episodes (Westby, 2005) were counted and scored. Every GAO received one point, so that one narrative could reach three points for the ES. Additionally, the total number of IST was also counted. Table 2 outlines the components of the SS.

Table 2: Overview of the story structure (SS). Gagarina et al. (2012).

	SS	Points
A1.	S	2
A2.	ISTe 1	1
A3.	G 1	1
A4.	A 1	1
A5.	O 1	1
A6.	ISTr 1	1
A7.	ISTe 2	1
A8.	G 2	1
A9.	A 2	1
A10.	O 2	1
A11.	ISTr 2	1
A12.	ISTe 3	1
A13.	G 3	1
A14.	A 3	1
A15.	O 3	1
A16.	ISTr 3	1
A17.	Total	17

4 Results

In this section the data is analyzed regarding the factors AoO and *time*. Time refers on the age of the children. The different rounds of tests are labelled as post 1, post 2, and post 3 in this article. Initially, the data is examined to give

indication about the influence of AoO. Therefore, the story structure score is statistically analyzed by a Welch t-test regarding the two groups separately for each round of tests. The analyses will be executed with the help of the statistical language R and the R Studio interface. According to Schneider et al. (2006) the story structure score gives the most information about the NAY. Therefore, IST and the ES are not statistically analyzed in this section but still considered in the discussion.

Figure 1 shows the mean score as well as the scores of the individual children regarding the AoO, and the variation of every child in post 1, post 2, and post 3.

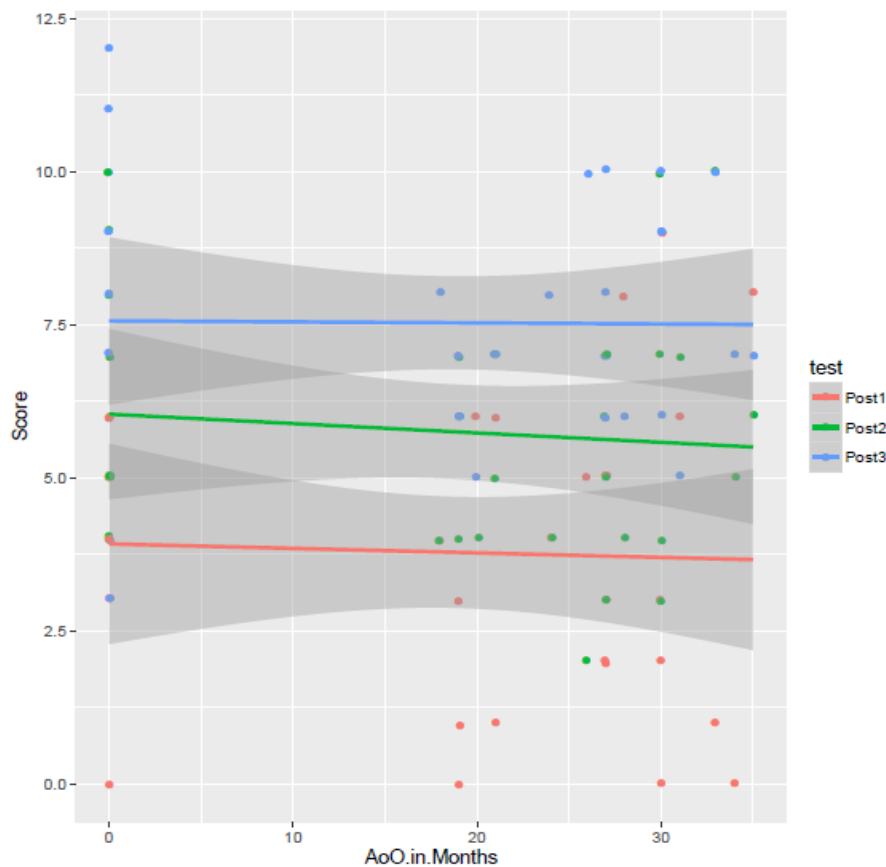


Figure 1: The progress of score in post 1, post 2, and post 3 regarding the factor AoO.

4.1 Post 1

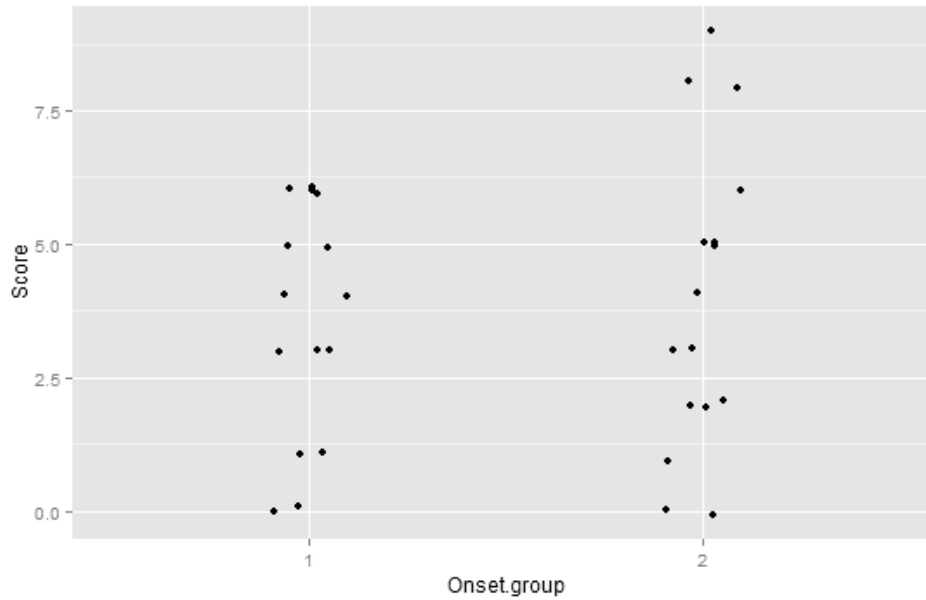


Figure 2: Distribution of the total score in post 1 separated by the groups.

The total score in group 1 differs between 0 and 6 credits, while in group 2 the score is spread between 0 and 9 credits. The mean score in group 1 is 3.63, in group 2 it is 3.94. Hence, the data shows that group 2 achieved a higher score for total and mean. The t- and p-values of this distribution are shown in Table 3:

Table 3: T-test results for the story structure score.

<i>AoO</i> groups	CHI	mean	T-value	df ³	p - value	distribu- tion
Group 1	16	3.63				0 - 6
Group 2	16	3.94				0 - 9
Total	32	3.78	0.352	28.091	0.727	

S was once verbalized in the post 1 context by a child of group 2. In group 1 as well as in group 2, four GAO were told by the children in post 1 testing. Therefore, the mean is 0.25. Further, in group 1 five incomplete episodes contain only AO, four episodes only G and nine incomplete episodes consist of GO or GA. In group 2 five incomplete AO-episodes were verbalized, four episodes contain only G and five episodes consist of GO and GA. In post 1 in

³ Df is the shortcut for degrees of freedom in a statistical analyses.

both groups 18 IST were totally produced, while the distribution between the groups is 10:8. Group 1 verbalized meanly 2.31 IST, and group 2 1.75.

4.2 Post 2

In this test period, the children's ages vary between 4;4 and 5;9 years.

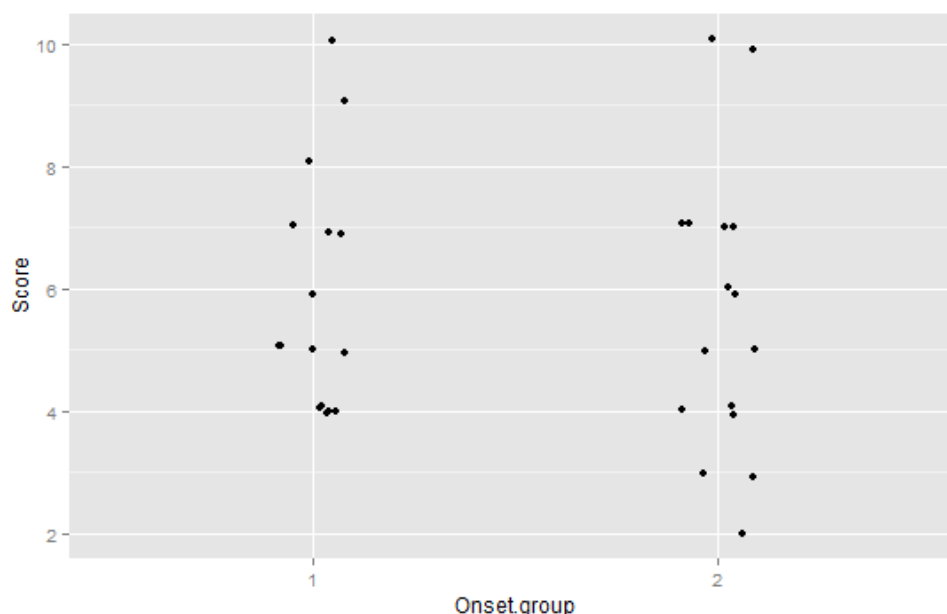


Figure 3: Distribution of score in post 2 separated into the *AoO* groups.

Through the visualization of the total story structure score (Table 4), we can see that the children in group 1 realized from 4 to 10 elements (mean = 5.88), while the distribution in group 2 ranges between 0 and 10 (mean = 5.63).

Table 4: T-test results for the story structure score in post 2.

<i>AoO</i> group	CHI	mean	T-value	df	p-value	distribu- tion
Group 1	16	5.88				4 - 10
Group 2	16	5.63				0 - 10
Total	32	5.75	0.330	28.964	0.744	

Two narratives in post 2 were introduced by S: One in group 1 and the other in group 2. In post 2, eleven *GAO* were produced by the children in group 1, by the children in group 2 there were six *GAOs*. One child in group 2 managed to

produce all *GAO* in his narrative. In group 1 one child verbalized two *GAO*. The mean of *GAO* in group 1 is 0.69, in group 2 it is 0.36. Group 1 produced eleven *AO*-episodes, while group 2 did ten. Eight times *G*-episodes were produced by group 1, group 2 produced four. *GA*- and *GO*-episodes do exist each eight and nine times. In group 1 and group 2, there is a mean occurrence of *IST* of 3.63 and 3.88. The total number vary between zero and eight, and zero and ten, respectively.

4.3 Post 3

In this test period, the children's ages vary between 5;3 and 6;5 years and all children have had constant contact with German for at least three years. In this round of tests the children in group 1 produced narratives with three to twelve story elements. Their distribution is shown in Figure 4. Therefore, the mean value for the narratives is 7.25. In group 2 the average story structure score is 7.81. The number varies between five and ten. The t-value results are presented in Table 5.

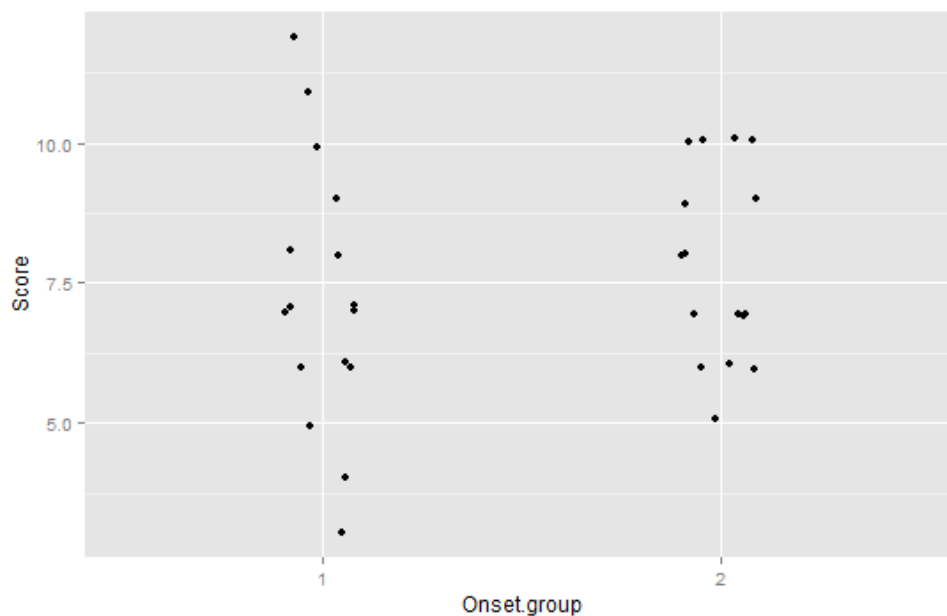


Figure 4: Distribution of score in post 3 separated into the AoO groups.

Table 5: T-test results for the story structure score in post 2.

<i>AoO</i> group	CHI	mean	T-value	df	p-value	distribu- tion
Group 1	16	7.25				3 - 12
Group 2	16	7.81				5 - 10
Total	32	7.53	0.766	26.821	0.4504	

Three narratives were introduced with *S* (2:1). Eight GAO were produced in group 1, while there were eleven in group 2. The mean for group 1 is 0.5 and for group 2 it is 0.875 for GAO in all post 3 narratives. AO-episodes were 19 and 20, respectively, produced in the two groups. In this round of testings there were one G-episode in group 1 and two in group 2. In group 1 GA- and GO-episodes appeared five times and in group 2 two times. ISTs were produced between zero to eight times in the different narratives of group 1, in group 2 there were zero to 10 ISTs. The mean in each group is 4.5 and 4.0.

Statistical analyses were also conducted for all ES and IST values in post 1, post 2, and post 3. However, none of them showed any meaningful result. In all cases the data could have been randomly chosen due to the fact, that the variations between the groups were less than the variations within the group. The statistical IST and ES data is presented in tables in Appendix 2.

In the second hypothesis should be examined if the N_{Ay} equalizes in time. Therefore, the means of every macrostructure component (SS = story structure score; ES = GAO, and IST) at all three rounds of tests are joined in Table 6. Further, the mean differences of all three components were computed and are used as basis of comparison. In the table below, we can see that the difference between group 1 and group 2 in score in the post 1 testing is 0.31. Group 2 reached more points than group 1. Both groups produced meanly the same GAOs in post 1. The IST mean difference in post 1 between the groups is 0.563, while group 1 achieved a higher score. In total, the difference between both groups in post 1 is 0.25. In the post 2 data, the score went higher in both groups, but the difference is: 0.25. In GAO is the difference between the groups 0.31 and in IST it is 0.25. The total difference is 0.31. In the last testing period (post 3), the difference in score between the groups is 0.1, while the difference in GAO is 0.38 and in IST it is 0.5. In total it is 0.03.

Table 6: The mean values and differences of all target variables dependent on the time period. The mathematical operator shows the dominance of group 2.

<i>Macrostructure components</i>					
	<i>AoO</i>	CHI	mean Score	mean <i>GAO</i>	mean <i>IST</i>
Post 1	Group 1	16	3.63	0.25	2.31
	Group 2	16	3.94	0.25	1.75
	Difference		-0.31	0	0.56
	Total Difference			0.25	
Post 2	Group 1	16	5.88	0.69	3.63
	Group 2	16	5.63	0.38	3.88
	Difference		0.25	0.31	-0.25
	Total Difference			0.31	
Post 3	Group 1	16	7.25	0.5	4.5
	Group 2	16	7.35	0.88	4.0
	Difference		-0.1	-0.38	0.5
	Total Difference			-0.03	

Since we use the same groups and considered the same components, we can illustrate our results of progress in Figure 5.

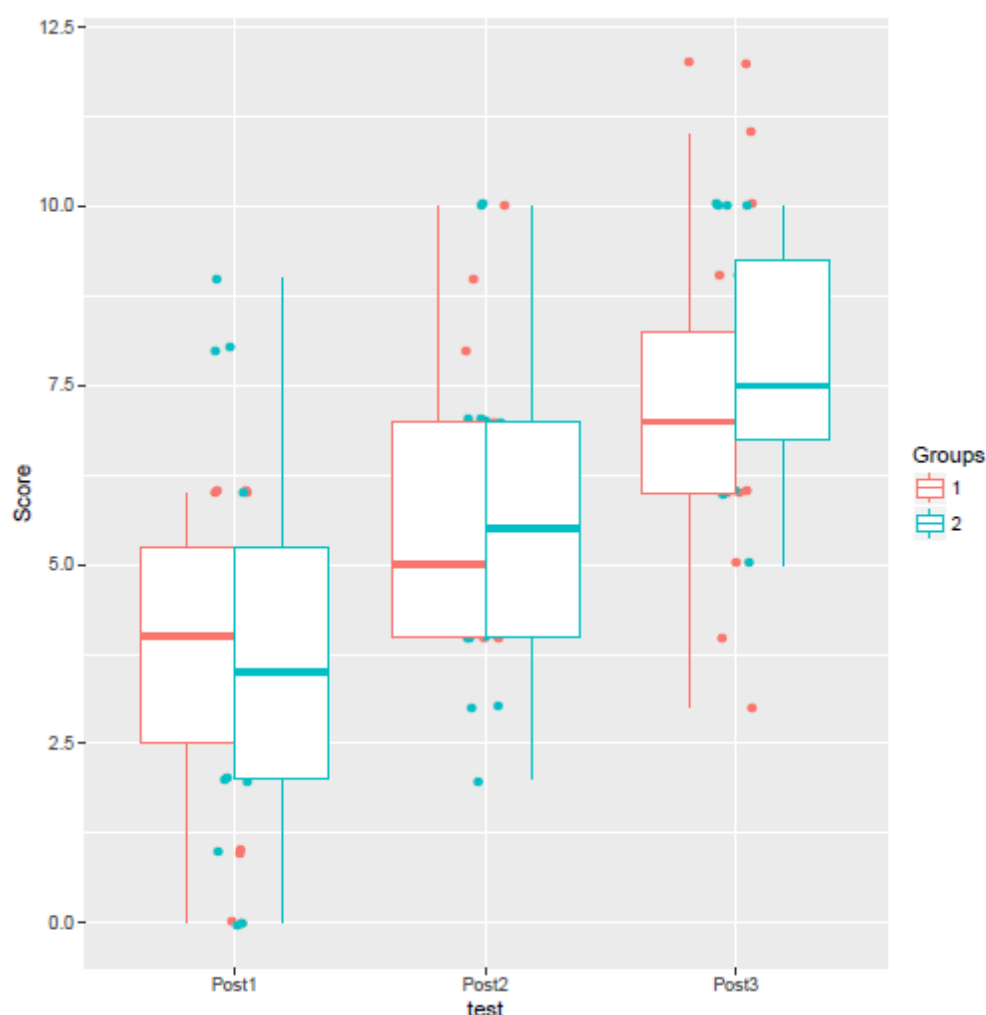


Figure 5: The distribution of macrostructure components in both groups for post 1, post 2, and post 3.

5 Discussion

The goal of this work was to ascertain the influence of the factor AoO on NAYa of bilingual children. First of all, we established two hypotheses to examine the factor AoO. The first hypothesis – children with lower AoO will receive higher scores in the analysis of macrostructure components in their narratives than those with higher AoO – based on the assumption of Singleton and Ryan (2004: 61) “The younger = the better in the long run”. Further, Dimroth (2007) claimed that there is evidence that the factor age influences *L2A*.

Abrahamsson and Hyltenstam (2009) showed in their study on older children with early and late *L2A* that there is no difference between simultaneously and sequentially bilingual children in *LA*. Hence, the second hypothesis was formulated: NAY will be more harmonized between the groups in post 3 than in post 1.

5.1 First Hypothesis

The first hypothesis could not be confirmed by the narrative data. The statistical analysis did not show evidence for the assumption that AoO has a positive effect on NAYa for simultaneously and sequentially bilingual children from 3;4 to 6;5. The difference of the mean value in all macrostructure components changes between post 1, post 2, and post 3. In post 1 and post 3, group 1 achieved a higher mean than group 2, but in the next round of tests it was vice versa. The difference was at no point of time significant. The t-test showed in all cases that the data could also have been randomly chosen. All p-values (except for GAO in post 3) were located between 0.2 and 1. A significant p-value would have been at $p < 0.05$. Further research could examine the NAYa scores for *L1* of the children to give a broader overview of the NAYa of each child. So, possible confounding variables could be also detected.

In this examination the sequentially bilingual children (group 2) reached a higher score in story structure (see Table 3, 4, 5, and 6). In post 1 and post 3 testing group 2 was minimally better. A possible explanation is that the sequential *LA* provided the children the opportunity to acquire the story structure reliably in one language. Hence, the concept could already strengthen before they had contact with the *L2* (Cummins, 1979; 2000). The reliable language knowledge could explain the higher scores in group 2 (Geva & Siegel, 2000). In general, the data confirms earlier research that has shown that NAY is language independent (Berman & Slobin, 1994; Fiestas, 2008; Strömquist & Verhoeven, 2004; Paradis et al., 2010). But to be sure *L1* should be considered in further studies to give a convincing overview of the NAY for each child.

The acquisition of IST is dependent on the language and closely related to the lexicon (Bartsch & Wellmann, 1995). This fact might explain the results for IST in this article. Group 1 performed better in the post 1 and the post 3 testings in IST scores (See Table 6 and Appendix 2). But there are controversial discussions about the lexicon acquisition. For example, Goldberg et al. (2008) examined that *L2*-learners acquire the lexicon faster than *L1*-learners.

Now, further factors which might influence NAY will be considered. According to Chondrogianni and Marinis (2011), the input quality and the input quantity is an important task. Even Oller and Eilers (2002) emphasize that the factor *input* is crucial in *LA* at all and especially in bilingual contexts. Further, parents and teachers can support NAY. For example, they could ask questions about the protagonists and actions in a story during they are reading picture books with the children (McCabe & Peterson, 1991; Peterson & McCabe, 1992; Reese et al., 1993). This task could help children to acquire a sense for the story structure (Uchikoshi, 2005; Verhoeven & Stömquist, 2001). Since the cognitive and linguistic processes in *LA* are dependent on *L1A* (Fiestas, 2008) further

research should accomplish a deeper examination of L1. In this case it is important to examine the story structure qualitatively. Fiestas (2008) and Fiestas and Peña (2004) ascertain in their research that story structure is language dependent. The emotional and cultural backgrounds of a language are transferred by the children into the narratives (Marian & Kaushanskaya, 2004). Due to Marian and Kaushanskaya (2004), the protagonists in Russian narratives of Russian-English bilingual children act more collectivistic than in the English narratives of the same children.

5.2 *Second Hypothesis*

The results in this examination show that the AoO groups differ in post 1 more than in post 3 (see Table 6). However, if each macrostructure component is considered apart, there is no convergence in all of them. Moreover, in the second round of tests, the two groups differ notably greater than in post 1. Therefore, no sequentially convergence is noticeable. In post 1, the total score of group 1 was 0.25 greater than that of group 2. In post 2, the total score increased even more: group 1 had 0.31 in total more than group 2. Only in post 3 there is a convergence between the groups. In that testing, the difference between the groups was 0.03. Group 2 performed better than group 1.

Right now, data supports rather Chondrogianni and Marinis' (2011) assumption that there is a convergence over time than that the two groups will differ constantly. In both of their studies, Chondrogianni and Marinis (2011) did not detect any difference between *LA* of bilingual experimentees (*AoO* 1;0 - 5;0). Further, the factor AoO could only be proven in data of the children's first three years. To be certain about the second hypothesis we need more research on narrative data. For example, a further round of tests could give more information about the long-term progression. On the other hand, Abrahamsson and Hyltenstam (2009) claimed that native-like *LA* could only be reached by children who had acquired the L2 in their first year. Moreover, through the analysis just a tendency can be shown. To present a meaningful statistical analysis, the data should have been normalized and the dependence between the three macrostructure components should be considered.

5.3 *Failure in Narrative Production*

One factor the data was chosen is that there are three narratives from three different rounds of tests (post 1, post 2, post 3) of each child. Some children have only zero or one credit in story structure score. In this section a reflection on how this low score could emerge should be done. There are further reasons that can influence the performance: Chondrogianni and Marinis (2011: 320) call

them external factors. For example, the socioeconomic background of the children can hamper N_{Ay}. Research has found parallels between residential environment, the education and the N_{AyA} (Dickinson & McCabe, 1991; Heath, 1982; Minami & McCabe, 1991). In this case, input quality and quantity could be a crucial factor (Goldberg et al., 2008; Chondrogianni & Marinis, 2011; Klassert & Gagarina, 2010).

Additionally, Uchikoshi (2005) claimed that a low score does not have to show a lack of N_{Ay}. But, there could be other factors which influence the performance, for example the factor *gender*. There are studies that examine narrative data in accordance to gender (Uchikoshi, 2005). There, girls achieved higher scores than boys (Bornstein et al., 1998; Gambell & Hunter, 1999; Karmiloff & Karmiloff-Smith, 2001). Nevertheless, it is important to emphasize that the factor gender is influenced by even more factors, and therefore, the performance is also influenced by them. Moreover, Uchikoshi (2005) highlighted that the factor age is also a crucial influence on the performance. In our data, the age was not balanced regarding the round of tests. In further studies this should be considered.

6 Conclusion

In this paper, the factor AoO on N_{AyA} of bilingual children has been considered. Therefore, 32 narratives of Turkish-German and Russian-German bilingual children of three rounds of tests were quantitatively analyzed and statistically evaluated. Two hypotheses were formulated. The theoretical background of N_{AyA} and LA was given for this article as well as the method of data collection and data analysis. Thereafter, the results were presented and discussed in relation to the hypotheses. Additionally, factors for further examinations, e.g. age, length of acquisition, input quality and quantity as well as the parents' socioeconomics and their education, were introduced. Furthermore, quantitative analyses should usually give generalizable data. In the present article we have seen that the confounding variables and too little data had more influence than the data per se. To decide in individual situations, a qualitative analysis of individual child performance is important (Stölten et al., 2015). Moreover, qualitative analysis could also be used to analyze the data of those children who had a very low score.

In summary, the first hypothesis could not be confirmed through the analysis in this article. There was no evidence that children with lower AoO performed better in the narrative telling task. According to the analysis, the second hypothesis could not either be confirmed. In fact, the two groups differed in post 3 testing less than in post 2 testing, but there was no sequential

convergence. In further analyses, we should consider a bigger sampling as well as a fourth round of tests.

7 References

- Abrahamsson, N., & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2): 249-306.
- Bartsch, K. & Wellmann, H. M. (1995). *Children talk about the mind*. New York: Oxford University Press
- Bericht des Berliner Interdisziplinären Verbunds für Mehrsprachigkeit (BIVEM), Oktober 2011 – Juni 2014. <http://www.zas.gwz-berlin.de/fileadmin/projekte/bivem/bivembericht-kurz.pdf>
- Berman, R. A. & Slobin, D.I. (1994). *Relating Events in Narrative: A Crosslinguistic Developmental Study* (Vol. 1), Lawrence Erlbaum Associates.
- Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language Learning*, 56: 9-49.
- Bornstein, M. H., Haynes, M. O., & Painter, K. M. (1998). Sources of child vocabulary competence: A multivariate model. *Journal of child language*, 25(2): 367-393.
- Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic impairments. *Child Language Teaching and Therapy*, 18: 1-21.
- Chang, C. J. (2004). Telling stories of experiences: Narrative development of young Chinese children. *Applied Psycholinguistics*, 25(1): 83-104.
- Chondrogianni, V. & Marinis, T. (2011). *Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children*. *Linguistic Approaches to Bilingualism*, 1 (3): 318-345.
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of educational research*, 49(2): 222-251.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire* (Vol. 23). Multilingual Matters.
- Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Paul H Brookes Publishing.
- Dimroth, C. (2007). Zweitspracherwerb bei Kindern und Jugendlichen: Gemeinsamkeiten und Unterschiede. In T. Anstatt (Ed.). *Mehrsprachigkeit bei Kindern und Erwachsenen: Erwerb, Formen, Förderung*. Tübingen: Attempto: 115-137.
- Engel, S. (1995). *The Stories Children Tell: Making Sense of the Narratives of Childhood*. W.H. Freeman.
- Fiestas, C. E., & Peña, E. D. (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools*, 35: 155-168.
- Fiestas, C. E. (2008). *The dynamic assessment of narratives: A bilingual study*. ProQuest.

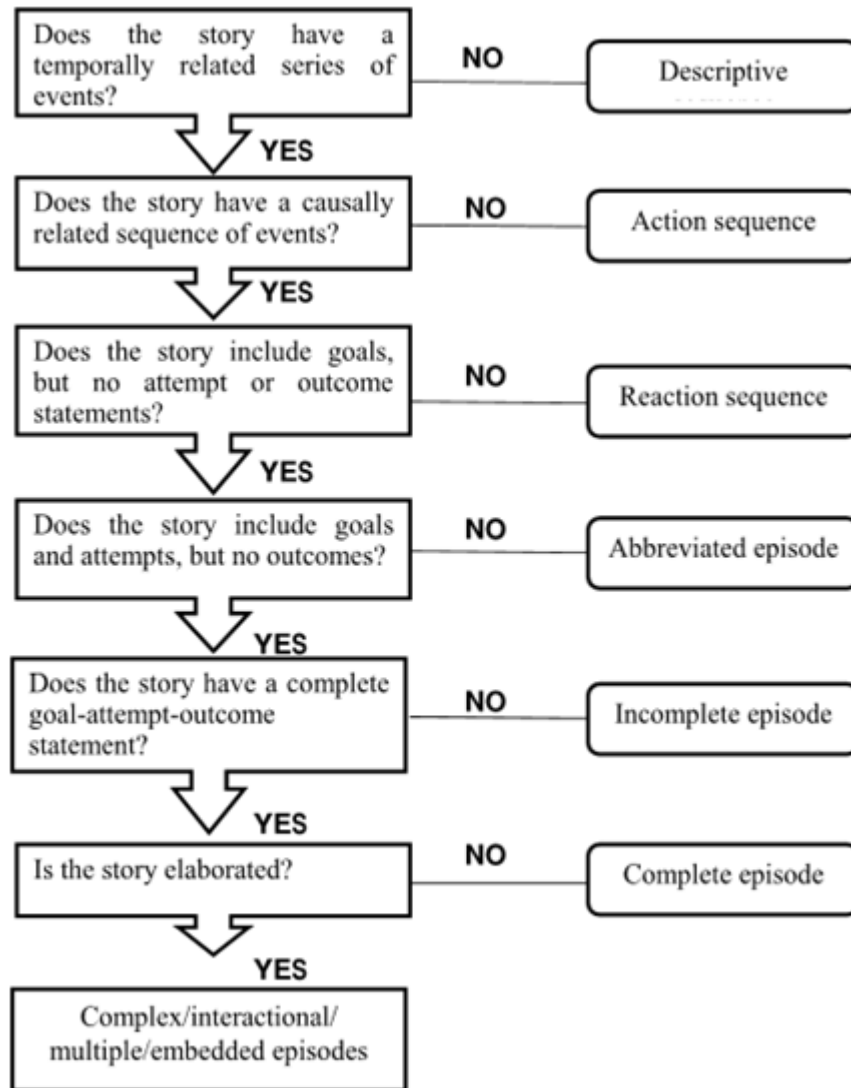
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U. & Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 56.
- Gambell, T. J., & Hunter, D. M. (1999). Rethinking gender differences in literacy. *Canadian Journal of Education/Revue canadienne de l'éducation*: 1-16.
- Geva, E., & Siegel, L. S. (2000). Orthographic and cognitive factors in the concurrent development of basic reading skills in two languages. *Reading and Writing*, 12(1-2):1-30.
- Gillam, R. B., & Johnston, J. R. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech, Language, and Hearing Research*, 35(6): 1303-1315.
- Gillam, R. B. & Pearson, N. A. (2004). *Test of narrative language*. Austin, TX: Pro-Ed.
- Glasgow, C., & Cowley, J. (1994). *The Renfrew Bus Story. North American Edition* (RBS-NA): Centerville School.
- Goldberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority L1 children learning English as a L2. *Applied Psycholinguistics*, 29: 1-25.
- Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood?. *Cognitive Development*, 27(2): 99-111.
- Guelzow, I. & Gagarina, N. (2007). Noun phrases, pronouns and anaphoric reference in young children narratives. *ZAS Papers in Linguistics*, 48: 203-223.
- Heath, S. B. (1982). What no bedtime story means: Narrative skills at home and school. *Language in society*, 11(1), 49-76.
- Hickmann, M. (2002). Children's Discourse: Person, Space and Time across Languages. *Cambridge Studies in Linguistics* (Vol. 98). Cambridge University Press.
- Hughes, D., McGillvray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessments*. Eau Claire, WI: Thinking Publications.
- Karmiloff, K., & Karmiloff-Smith, A. 2001. *Pathways to Language. From Fetus to Adolescent*. Cambridge: Cambridge University Press.
- Klassert, A. & Gagarina, N. (2010). Der Einfluss des elterlichen Inputs auf die Sprachentwicklung bilingualer Kinder: Evidenz aus russischsprachigen Migrantenfamilien in Berlin. *Diskurs Kindheits- und Jugendforschung* 4: 413-425.
- Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular* (Vol. 3). University of Pennsylvania Press.
- Leadholm, B. & Miller, J. F., (1994). *Language sample analysis guide: The Wisconsin guide for identification and description of language impairment in children*. Bulletin 92424.
- Lorusso, M. L., Galli, R., Libera, L. & Gagliardi, C. (2007). Indicators of theory of mind in narrative production: A comparison between individuals with genetic syndromes and typically developing children. *Clinical Linguistics and Phonetics*, 21(1): 37-53.

- McCabe, A. & Peterson, C. (1991). Getting the story: A longitudinal study of parental styles in eliciting oral personal narratives and developing narrative skill. In A. McCabe & C. Peterson (Eds.). *Developing narrative structure*. Hillsdale, NJ: Erlbaum: 217 -253.
- MacWhinney, B. (2000). *The CHILDES project: The database* (Vol. 2). Psychology Press.
- Minami, M., & McCabe, A. (1991). Haiku as a discourse regulation device: A stanza analysis of Japanese children's personal narratives. *Language in Society*, 20(4): 577-599.
- Marian, V., & Kaushanskaya, M. (2004). Self-construal and emotion in bicultural bilinguals. *Journal of Memory and Language*, 51(2): 190-201.
- McNamara, T. F. (1996). *Measuring Second Language Performance*. *Applied Linguistics and Language Study*. London; New York: Longman.
- Miller, J. F. (1991). Quantifying productive language disorders. In J. Miller(Ed.), *Research on child language disorders: A decade of progress*. Austin, TX: Pro-Ed: 211-220.
- Miller, J. F., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. J. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research & Practice*, 21(1): 30-43.
- Oller, D. K., & Pearson, B. Z. (2002). Assessing the Effects of Bilingualism: A Background. In D. K. Oller & R. E. Eilers (Eds.). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters: 3-21.
- Paul, R., & Smith, R. L. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech, Language, and Hearing Research*, 36(3): 592-598.
- Paradis, J., Genesee, F., & Crago, M. (2010). *Dual language development & disorders: A handbook on bilingualism & second language learning* (2nd ed.). Baltimore, MD: Paul H. Brookes.
- Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics*. Springer Science & Business Media.
- Peterson, C., & McCabe, A. (1992). Parental styles of narrative elicitation: Effect on children's narrative structure and content. *First Language*, 12: 299-321.
- Peña, E. D., Gillam, R. B., Malek, M., Ruiz-Felter, R., Resendiz, M., Fiestas, C., & Sabel, T. (2006). Dynamic assessment of school-age children's narrative ability: An experimental investigation of classification accuracy. *Journal of Speech, Language, and Hearing Research*, 49(5): 1037-1057.
- Radford, A. (1995). Phrase structure and functional categories. The handbook of child language. In Fletcher, P. & MacWhinney, B. (Eds.). *The handbook of child language*. Blackwell Oxford: 483-507.
- Reese, E., Haden, C. A., & Fivush, R. (1993). Mother-child conversations about the past: Relationships of style and memory over time. *Cognitive development*, 8(4): 403-430.
- Renfrew, C. E. (1969). *The bus story: A test of continuous speech*. North Place, Old Headington: Oxford.

- RStudio Team (2015). *RStudio: Integrated Development for R*. RStudio, Inc., Boston, MA
URL <http://www.rstudio.com/>.
- Ruberg, T. (2013). Problembereiche im kindlichen Zweitspracherwerb. *Sprache – Stimme – Gehör*, 37: 181-185.
- Rumelhart, D. E. (1975). Notes on a schema for stories. *Representation and understanding: Studies in cognitive science*: 211-236.
- Shapiro, L. R., & Hudson, J. A. (1991). Tell me a make-believe story: Coherence and cohesion in young children's picture-elicited narratives. *Developmental psychology*, 27(6): 960-974.
- Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Retrieved from University of Alberta Faculty of Rehabilitation Medicine.
- Schneider, P., Hayward, D., & Dubé, R. V. (2006). Storytelling from pictures using the Edmonton narrative norms instrument. *Journal of Speech Pathology and Audiology*, 30(4): 224-238.
- Singleton, D. M., & Ryan, L. (2004). *Language acquisition: The age factor* (Vol. 9). Bristol: Multilingual Matters.
- Slobin, D. I. (1985). Crosslinguistic evidence for the language-making capacity. *The crosslinguistic study of language acquisition*, 2: 1157-1256.
- Stein, N. L., & Glenn, C. G. (1977). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New Directions in Discourse Processing*. Norwood, NJ: Ablex: 53-120.
- Strömquist, S., & Verhoeven, L. T. (2004). *Relating events in narrative: Typological and contextual perspectives*. Psychology Press.
- Stölten, K., Abrahamsson, N., & Hyltenstam, K. (2015). Effects of age and speaking rate on voice onset time. *Studies in Second Language Acquisition*, 37(1): 71-100.
- Symons, D. K. (2004). Mental state discourse, theory of mind, and the internalization of self–other understanding. *Developmental Review*, 24(2): 159-188.
- Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- Trabasso, T., & Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes*, 15: 249-275.
- Trabasso, T., & Rodkin, P. C. (1994). Knowledge of goals/plans: A conceptual basis for narrating "Frog where are you?". In R. A. Berman & D. I. Slobin (Eds.). *Relating events in narrative: A cross linguistic developmental study*. Mahwah, NJ: Lawrence Erlbaum Associates: 85-106.
- Uchikoshi, Y. (2005). Narrative development in bilingual kindergarteners: Can Arthur help? *Developmental Psychology*, 41(3): 464-478.
- Verhoeven, L. & Strömquist, S. (Ed.) (2001). *Narrative development in a multilingual context* (Vol. 23). John Benjamins Publishing.

- De Villiers, J. (2007). The interface of language and theory of mind. *Lingua*, 117(11): 1858-1878.
- Westby, C. E. (2005). Assessing and remediating text comprehension problems. In Kamhi, A. G. (Ed.). *Language and reading disabilities*. Boston: Pearson/A & B: 157-232.

Appendix 1



The binary decision tree. Gagarina et al. (2012: 13, based on Westby, 2005).

Appendix 2

Table 7: T-test results of episode structure in post 1 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribu- tion
Group 1	16	0.25				0 - 1
Group 2	16	0.25				0 - 1
Total	32	0.25	0	30	1	

Table 8: T-test results of IST values in post 1 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribu- tion
Group 1	16	2.3125				0 - 8
Group 2	16	1.75				0 - 8
Total	32	2.031	0.6282	29.752	0.5347	

Table 9: T-test results of episode structure values in post 2 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribu- tion
Group 1	16	0.6875				0 - 2
Group 2	16	0.375				0 - 3
Total	32	0.53125	1.2423	28.762	0.2245	

Table 10: T-test results of IST values in post 2 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribution
Group 1	16	3.625				1 - 8
Group 2	16	4.0667				0 – 11
Total	32		-0.4687	24.693	0.6434	

Table 11: T-test results of episode structure values in post 3 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribution
Group 1	16	0.5				0 – 1
Group 2	16	0.875				0 - 2
Total	32	0.6875	-1.8605	29.064	0.07296	

Table 12: T-test results of IST values in post 3 data. R computation.

AoO groups	CHI	mean	T-value	df	P-value	distribution
Group 1	16	4.5				1 - 8
Group 2	16	4				1 - 10
Total	32	4.25	0.60486	29.371	0.5499	