Structural Development of Narratives in Arabic: Task Complexity, Age, and Cultural Factors

Article *in* Language, Speech, and Hearing Services in Schools · February 2020

DOI: 10.1044/2019_LSHSS-19-00044

3 authors, including:

Abdessatar Mahfoudhi
Australian College of Kuwait
71 PUBLICATIONS **372** CITATIONS

LSHSS

Research Article

Structural Development of Narratives in Arabic: Task Complexity, Age, and Cultural Factors

Fauzia Abdalla,[a] Abdessattar Mahfoudhi,[b,c] and Shouq Alhudhainah[d]

Purpose: This study examines the effect of age and task complexity on the macrostructure of story production in preschool- and school-age Kuwaiti Arabic–speaking children. It also compares the children's production of core and complementary macrostructure story elements.
Method: A descriptive, cross-sectional research design was used to explore the participants' narrative skills. A total of 122 monolingual speakers of Kuwaiti Arabic (97 children and 25 adults) participated in this study. The children aged 4;0 to 7;11 (years;months) were randomly recruited from public schools across Kuwait. There were 24 four-year-olds (Kindergarten 1), 23 five-year-olds (Kindergarten 2), 23 six-year-olds (Grade 1), and 27 seven-year-olds (Grade 2). A group of adults was also included to establish a benchmark. Storytelling was elicited from all the participants using two sets of sequential pictures from the Edmonton Narrative Norms Instrument: a one-episode story and a more complex three-episode story (Schneider et al., 2005). Across-group comparisons were conducted to explore the effect of age, story complexity, and type of macrostructure elements on story production.
Results: The findings revealed a progression by age in the development of story macrostructure, but there was no effect of task complexity. Within all age groups, the core macrostructure components were mastered before the complementary elements.
Conclusions: The results of this study confirmed that cross-linguistic narrative measures could be used in contexts that are culturally and linguistically different with minor adaptations. The piloting of two picture-based stories showed that the shorter one-episode version may be sufficient to evaluate the language development of this age group.

Narratives are factual or fictional recounts of events presented in chronological order (McCabe et al., 2008). Oral narratives are emerging as a naturalistic method for investigating children's linguistic and cognitive development due to the numerous advantages they offer. Story narration exposes children to a distinct form of language that is holistic, rich, complex, and omnipresent in everyday social, educational, and recreational contexts (Bloch, 1999). Specific organizational principles govern comprehension and production of oral narratives, and their development can be observed, analyzed, and described (Schneider, 1996).

Narrative analysis can be carried out at the microstructure or macrostructure level. The microstructure of the narrative discourse refers to the linguistic complexity of the word and sentence levels, including lexical diversity, morphological structure, and sentence complexity (e.g., Justice et al., 2006). In contrast, macrostructure is a schematic organization that is independent of the specific content of the story (Hughes et al., 1997). The macrostructure of children's fictional stories is the focus of the current study.

The story grammar model is a popular macrostructural analysis that includes the structural elements of a story, such as the setting, characters, and plot (N. Stein & Glenn, 1979). Essentially, story grammar units form brackets of information provided in chronological order within one or more episodes. These units make up the story's central constituent and are generally considered essential for standard storytelling. A complete episode minimally consists of a foundational unit involving "the initiating event" (what motivates the character to act), "the attempt" (actions the

[a]Department of Communication Disorders Sciences, Kuwait University, Kuwait City
[b]English Language Department, Australian College of Kuwait, Kuwait City
[c]Research and Test Development Unit, Center for Child Evaluation and Teaching, Kuwait City, Kuwait
[d]Speech and Hearing Unit, Adan Hospital, Ministry of Health, Hadiya, Kuwait
Correspondence to Fauzia Abdalla: fauziamam@yahoo.com

Editor-in-Chief: Holly L. Storkel
Editor: Kerry Danahy Ebert
Received February 18, 2019
Revision received May 16, 2019
Accepted November 1, 2019
https://doi.org/10.1044/2019_LSHSS-19-00044

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

character takes in response to the initiating event), and "the outcome" (consequence of the action; Schneider & Dubé, 2005). Typically, these core structural units are more often included in a story than elements considered complementary, such as "the internal responses" (the character's thoughts and feelings about the initiating event), "the plans" (the character's ideas for actions), and "the reactions" (inner thoughts or feelings about the action's outcome; Schneider & Dubé, 2005). Core units are so named also because they tend to develop earlier than the complementary ones (e.g., Berman & Slobin, 1994; Castilla-Earls et al., 2015; Trabasso et al., 1992).

Acquisition and knowledge of story schema appear to be age related, with stories of older children resembling adult competence (Gagné & Crago, 2010; Hayward et al., 2009; Schneider et al., 2006). Basic narrative structure emerges in the preschool years (5–6 years; Hughes et al., 1997). As children mature, their stories become more complex in terms of episodes and inclusion of core and complementary story grammar components, including the late-developing internal state and emotional reactions (e.g., Berman & Slobin, 1994; Castilla-Earls et al., 2015). These late-developing story grammar elements tend to differentiate between children with developmental language disorder (DLD) and their peers with typical language development (e.g., Fichman et al., 2017; Kupersmitt & Armon-Lotem, 2019).

Whether at the microlevel of morphosyntactic complexity or at the macrolevel of story grammar, narratives have been found to have clinical and educational utility (Guo & Schneider, 2016). A growing body of research has demonstrated that narratives can identify or predict language and literacy delays as well as serve as a teaching or treatment tool. For example, children with DLD have less developed story grammars in oral story production than their typically developing peers (Fichman et al., 2017; Merritt & Liles, 1989; Chan, 2003, for Cantonese; Soodla & Kikas, 2010, for Estonian). In a longitudinal study of the language development of children with DLD, Bishop and Edmundson (1987) found that the ability to retell a picture story (the bus picture story by Renfrew, cited in Bishop & Edmundson, 1987) was the best predictor of language improvement. In a longitudinal study that measured children's performance at the age of 2 years and later at the age of 4 years, Paul and Smith (1993) also found that the macrostructure of oral narrative skills, using the information score of the bus story test, predicted language delays. Similar results were found in relation to narrative skills predicting literacy and academic development (Griffin et al., 2004; Pinto et al., 2016). It is, therefore, crucial to develop or adapt tools to assess children's narratives, especially in contexts where language measures and standardized assessment procedures are lacking, as is the case of Kuwaiti Arabic.

Another objective of this study is to explore the effect of internal task complexity on the production of narrative macrostructure (one-episode vs. three-episode story), a question that has not been investigated so far in the first language literature. The available research evidence is limited to the effect of across-task variation. Previous studies found discrepancies in the complexity of child narratives depending on the type of task, but the results remain heterogeneous. For instance, there is evidence that story retelling leads to the generation of longer stories with more story grammar components (e.g., Merritt & Liles, 1989). There is also evidence that personal narratives are easier to perform than fictional narratives, especially among children from lower socioeconomic classes (Shiro, 2003). However, McConnell (2011), who examined children's narratives using two slightly different tasks of story retelling, once with oral retelling only and once with the support of pictures, found that narratives produced using the picture-supported technique yielded more complete stories (story grammar elements) that were lexically more complex. Retelling was also found to be easier than telling in macrostructure development, although the difference was not statistically significant (Maviş et al., 2016). Lofranco et al. (2006) examined the effect of using a model versus no model in wordless picture storytelling tasks in English-speaking children who were also exposed to Tagalog. They found a significant effect of story macrostructure complexity in favor of the storytelling with a model (Lofranco et al., 2006).

Another aspect that the current study explores is whether there are differences in the production of certain story elements, such as internal plan, in a different language and culture. Although story grammar has been studied mainly in a Western English–speaking context, there are several studies that examined narratives from a cross-linguistic perspective (e.g., Berman & Slobin, 1994; Maviş et al., 2016). While some studies of narratives that compared macrostructure across cultures reported similar cross-cultural and cross-linguistic performance, others found differences. Story grammars of 7-year-old Bhutanese children were comparable to children from rural Pennsylvania (C. L. Stein, 2004). On the other hand, using a wordless picture book narrative, Pankratz et al. (2007) found that minority children (African American and Hispanic children grouped together) performed significantly below their European children counterparts on macrostructure elements (but not on sentence length). Similar results were reported in the work of van Kleeck et al. (2011), who also examined story retelling performance in African American and European American children.

This Study

Data from Arabic are essential because of the structural variation of the language and, more importantly, because of the rather unique cultural, Arabic educational context. For instance, Arabic-speaking children in Kuwait, where the study was conducted, do not seem to receive adequate exposure to narratives because of diglossia (Ferguson, 1959), as well as cultural and educational practices. Diglossia refers to the co-existence of two varieties of a language that are functionally in complementary distribution—a high variety that is primarily a written language that is learned at school and is used in most formal written and spoken communication and a low variety that is mainly a spoken vernacular that is the real native language of any Arabic language

speaker. Despite the fact that they are genetically related, the two varieties are lexically, morphologically, and syntactically different. The high variety, Standard Arabic, has preserved most of the morphosyntactic and lexical features of its historical predecessor, Classical Arabic, because it has been the carrier of religious, literary, and cultural heritage. The less prestigious variety, spoken Arabic, has developed naturally from the same ancestor language into numerous local varieties. In their early years at home and in nursery and kindergarten, children are sometimes exposed to narratives in their native language (the spoken variety) before they are initiated to narratives in Standard Arabic later during their formal training.

The previous few studies in Arabic, both those that focused only on narrative microstructure (e.g., Kelly, 2010; Saidi, 2014) and others that examined both the microstructure and macrostructure of narratives (Leikin et al., 2014; Ravid et al., 2014), were done in other varieties of Arabic that are structurally and sociolinguistically different from Kuwaiti Arabic.

The main objective of this study is to develop a narrative instrument that could be validated later, on a larger sample of children with typical language development and participants with DLD, to assess language acquisition in the Kuwaiti context where assessment tools are lacking. Tests currently implemented are either in other varieties of Arabic or in the standard variety, which is not appropriate for preschool children. Measures or procedures used in other Arabic dialects are not appropriate because of linguistic and cultural differences. Although Arabic dialects share some lexical and grammatical features, some are mutually unintelligible. It is, therefore, necessary to develop assessment tools that are in the local variety and are validated on the local population.

Since this is the first study of narratives in Kuwaiti Arabic, data from a group of adults were also gathered to establish a benchmark for child narrative development and cross-linguistic and cross-cultural comparison with the more prevalent Western/English norms reported in the literature (e.g., Schneider et al., 2006). In addition to studying the story grammar development in the narratives of Kuwaiti children in the early school years, this study investigates the effect of task complexity on story grammar to inform assessment and teaching practice using a measure that was developed in a Western context.

The following research questions guided the current study.

1. Is there a difference in story grammar performance between the different age groups?
2. Does story task complexity (simple vs. complex story) affect story grammar production?
3. Are there differences in the core versus complementary story grammar categories in the participants' story formulation?
4. Are there certain aspects of story grammar that pose a challenge for the participants?

Method

Participants

The participants for this study were 97 typically developing, monolingual, Kuwaiti Arabic–speaking children aged 4;0 to 7;11 (years;months) and 25 Kuwaiti adults (N = 122). The adults who served as controls were between 19 and 60 years old. All the children attended Arabic preschool and elementary public school in four grade levels: Kindergarten 1, Kindergarten 2, Grade 1, and Grade 2. The public schools represented all six governorates of Kuwait: Kuwait City, Hawalli, Farwaniya, Mubarak Al-Kabeer, Ahmadi, and Jahra.

For the analyses, the children were subdivided into four groups according to age (4-, 5-, 6-, and 7-year-olds). These age groups also generally corresponded to the school grades that the children attended, with the 4-year-olds in Kindergarten 1, 5-year-olds in Kindergarten 2, 6-year-olds in Grade 1, and 7-year-olds in Grade 2. For the children, there were slightly more boys (n = 51) than girls (n = 46), whereas the opposite was true for the adult group with eight men and 17 women. Table 1 presents the age groups of the participants and the number of males and females in each group. A chi-square test for independence indicated no significant association between gender and age group, $\chi^2(4, n = 122) = 3.76$, $p = .44$, phi = .18.

Permission to conduct the testing was obtained from the Ministry of Education, Kuwait. Once a list of all the schools that could be accessed was obtained from the Ministry of Education, 32 schools (eight for each of the four grade levels) from various districts of Kuwait were randomly selected, taking into account that Kuwaiti public schools are segregated. Therefore, separate schools for girls and boys for each grade level were included. The principals of the schools facilitated testing after clearance from the children's parents. The teachers gave the children the choice to participate, and only those who agreed were tested. Approximately 7% of the children either opted out of participating or parental consent was not secured.

Admission criteria for governmental mainstream schools in Kuwait include cognitive abilities that are within normal limits. The children are required to satisfy the

Table 1. Descriptive data on the participants.

Age group	Mean age in years (*SD*)	Gender (*n*)		Total (*N*)
		Male	Female	
4-year-olds	4.53 (0.29)	13	11	24
5-year-olds	5.52 (0.32)	13	10	23
6-year-olds	6.56 (0.32)	11	12	23
7-year-olds	7.44 (0.26)	14	13	27
Adults	40.56 (11.7)	8	17	25
N		59	63	122

Note. Standard deviations (*SD*) are presented within parentheses, next to the means; n = sample size for each age group.

academic achievement standards for their grades. According to the social services records in the school, the participants met the current study's inclusionary criteria of being Kuwaiti citizens with no history of cognitive delay, socioemotional problems, DLD, and vision or hearing impairments. Four children were excluded from the study because they did not complete the storytelling task.

Story Elicitation Stimuli

The current study used the Edmonton Narrative Norms Instrument (ENNI), an assessment tool designed to gather story narration from children aged 4–9 years (for more details, see Schneider et al., 2005, 2006). Reportedly, ENNI demonstrated construct and concurrent validity (Schneider et al., 2006). Authorization to adopt the instrument for the present investigation was obtained from the developers, Professor Phyllis Schneider and Dr. Denyse Hayward. As noted by the authors, the ENNI stories use picture scenes to measure the children's narrative abilities. These fictional story illustrations correspond to N. Stein and Glenn's (1979) story grammar categories. The ENNI procedures also included a reliable scoring scheme for measuring the children's formulation of story macrostructure (for more details, see Schneider et al., 2005).

The ENNI comprises two sets of stories (A and B), each with three levels of complexity in terms of number of episodes. Two narrative tasks of varying complexity from Set A were selected from the ENNI protocol. The stories were reportedly illustrated by a cartoonist, each on a separate page. The first story selected for the current study was A1, "The Ball," a one-episode simple story with two characters, giraffe (male) and elephant (female). The plot of the story was that the pair was playing with a ball beside a swimming pool when their ball fell into the water, causing the elephant to be sad. The giraffe then decided to jump into the water to fetch the ball and give it back to the elephant. The second story A3, "The Airplane," was more complex, consisting of three episodes and four characters. Giraffe (the male character) was playing with a toy airplane near a swimming pool when the elephant (the second character, a young female) snatched the plane. The plane fell into the water, and the third character, a male adult elephant lifeguard, tried to retrieve it but was unsuccessful. A fourth character, a female adult elephant with a net, came to their rescue and ended up recovering the plane and handing it back to giraffe. A1 and A3 stories were chosen for this study as they have already been used with Canadian children in English (Schneider et al., 2006) and French (Gagné & Crago, 2010), thus enabling cross-cultural comparisons.

Procedure

This study adopted the stories and the elicitation procedures described in the works of Schneider et al. (2005, 2006). The protocol included administration of a training story followed by A1 and A3 stories. The third author and three trained examiners tested each child individually in a quiet room at their respective preschool or elementary school. A five-picture training story (also from the ENNI website) was presented to acquaint the child with the storytelling procedure. Following this step, the examiner showed the child a series of pictures depicted in a booklet, first for A1 story and then for A3 story. The examiner followed the ENNI instructions as presented in the work of Schneider et al. (2005) and told the participant in Kuwaiti Arabic, "First I'll show you all the pictures. Then we'll go back to the beginning of the story. And then I want you to look at the pictures and tell me the story that you see in the pictures. I won't be able to see the pictures, so you need to tell me the story really well so I can understand it."

Transcription, Scoring, and Reliability

Each participant's story production was audio-recorded using an Olympus portable digital voice recorder. A native Arabic-speaking research assistant transcribed the participants' narrative formulations verbatim in Arabic. A second assistant, who was blind to the purpose of the study, independently transcribed 25 recordings (approximately 20% of the total; five transcripts for each age group). Word-by-word agreement was 97.8%.

The story transcripts were then scored for the inclusion of target story grammar categories using the ENNI scoring criteria with slight modifications (ENNI scoring scheme can be found in the work of Schneider et al., 2005). In the ENNI procedures, a score of 0 represents a missing story grammar category, a score of 1 represents the inclusion of a complementary unit, and a score of 2 represents a core element. In the current study, to facilitate comparisons across story types, a binary scheme of 0 (absent category) and 1 (unit present) was adopted without according extra weight for the core story grammar elements. As described for French in the study of Gagné and Crago (2010), there are potential contradictions between the gender of the noun (giraffe) in Arabic and the actual gender of this character as depicted in the ENNI story pictures. In Arabic, the noun "giraffe" (zarafah) ends with the feminine marker "at"; however, the character "giraffe" in the ENNI pictures is a male. Additional allowances were made for the participants whereby gender contradiction was ignored and a score of 1 was accorded whether giraffe was treated as a girl or a boy.

The A1 story ("The Ball") consisted of a maximum total raw score of 10 points. A sample story of a preschool participant can be found in the Appendix. The sample provides a transcription of the child's A1 story (English translation under each utterance) and scoring for the story using the adapted ENNI procedure. The A3 story ("The Airplane") had a maximum total score of 28 points. The participant's total score was converted into a percentage to obtain an overall measure of story grammar elements that were included in story formulation. Twenty-five transcripts were randomly selected and independently scored by a second research assistant to calculate interrater reliability. Kappa measure of agreement was used to check the interrater reliability for story grammar categories and was

found to be $k = .89$, with a significance of $p < .001$. According to Peat (2001), a value of above .8 represents very good agreement.

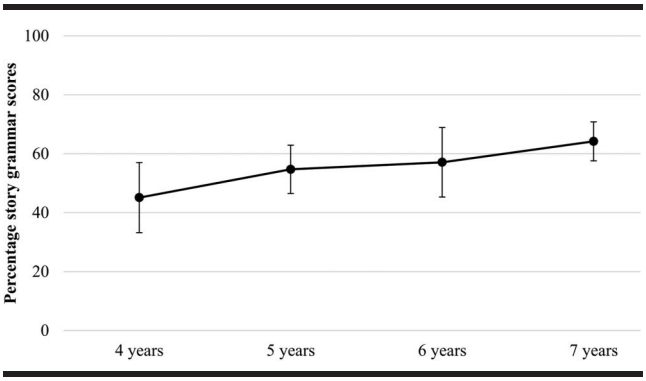
Results

The story macrostructure categories that the participants generated, for two stories varying in length and complexity, were analyzed for possible age- and complexity-related effects. As noted in the Method section, the children were subdivided into four age groups: Group 1: 4-year-olds, Group 2: 5-year-olds, Group 3: 6-year-olds, and Group 4: 7-year-olds. For the ensuing analyses, the independent variable was age with four levels (4-, 5-, 6-, and 7-year-olds). The dependent variable was the mean percentage scores for the two stories (i.e., amount of story grammar elements included in the children's stories). The descriptive scores of the adult participants who served as a benchmark are also reported.

Development of Story Macrostructure

A developmental increase in the quantity of story grammar information that the children would include in their stories was predicted based on the story grammar model and findings from similar studies (Gagné & Crago, 2010; Schneider et al., 2006). A trend analysis was performed to explore whether the mean scores increase in increments or a varying pattern is present across the ordered set of groups. Figure 1 presents the mean performance (total for both stories) of the 4-year-olds (45.1%), 5-year-olds (54.7%), 6-year-olds (57.1%), and 7-year-olds (64.2%). There was a significant linear trend, $F(1, 93) = 46.8$, $p < .001$, indicating that, as the children's ages increased, the story grammar scores increased proportionately. Quadratic trend was not significant, $F(1, 93) = 0.39$, $p = .54$, which suggests that the increase was even across the age groups. Comparatively, the scores of the adults were higher than those of the children with a mean of 79.7% ($SD = 6.9$) and a range of 71%–100%.

Figure 1. Mean percentage scores of story grammar of the children across age levels. Bars represent standard deviations.



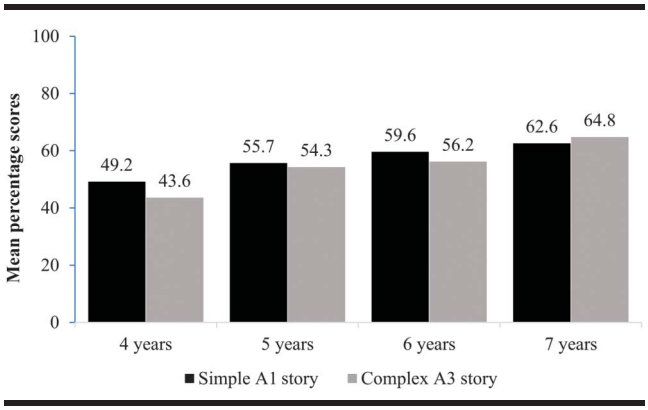
Story Task Complexity

The two stories presented to the participants varied in story content information, length and gender, and number of characters. As mentioned earlier, A1 was a single-episode story with two characters, while A3 had three episodes and four characters. A two-way, between-groups analysis of variance was conducted to explore the impact of age and story type (two levels: simple vs. complex) on the mean story grammar scores. Figure 2 presents mean comparative scores for the four age groups. There was a statistically significant main effect for age, $F(3, 186) = 14.9$, $p < .001$ (partial $\eta^2 = .19$). Post hoc comparisons using the Tukey's honestly significant difference test revealed that the youngest group (4-year-olds) was significantly different from the 5-year-olds ($p < .01$) and the 6- and 7-year-olds ($p < .001$). In turn, the 5-year-olds scored significantly lower than the 7-year-olds ($p < .007$) but not the 6-year-olds ($p = .72$). The two oldest groups (6- and 7-year-olds) did not differ significantly ($p = .13$). The main effect of story type, $F(3, 186) = 1.1$, $p = .29$, did not reach statistical significance. The interaction effect between age and story type was also not statistically significant, $F(3, 186) = 0.79$, $p = .49$. A descriptive analysis showed that the adult controls scored higher than the children for both stories, A1 ($M = 83.6$, $SD = 9.1$) and A3 ($M = 78.3$, $SD = 7.6$). Moreover, the adults performed better on the simple A1 story than the more complex A3 story, $t(24) = 2.93$, $p < .007$.

Core Versus Complementary Story Grammar

For this analysis, the elements of the story macrostructure were broken down into core versus complementary components. As mentioned above, the core elements included the "initiating event" (what motivated the main character to an action), "attempt" (the action that the main character took in response to the initiating event), and "outcome" (the consequence of the action). On the other hand, the complementary aspects of the story comprised the rest of the macrostructures, such as the "setting," the "characters,"

Figure 2. Mean percentage story grammar scores for A1 versus A3 stories.



and the "reactions" (the characters' feelings or thoughts about the outcome of their actions). Subcategories within story grammar were explored to determine whether there was a variation in the participants' core versus complementary story grammar scores.

As evident from Figure 3, all four groups of children (4-, 5-, 6-, and 7-year-olds) were more likely to include the core units of the story grammar (initiating event, attempt, and outcome) when telling a story (A1 and A3) than units that are considered complementary. A mixed analysis of variance with repeated measures on element type (core vs. complementary) was conducted to compare the children's performance across age groups. There was a significant main effect for element type (Wilks' lambda = .19), $F(1, 93) = 385.4$, $p < .001$, partial $\eta^2 = .81$. These findings suggest that the mean percentages of the core elements included in the story were significantly different than those of the complementary units. The analyses also revealed a statistically significant main effect of age group, $F(3, 93) = 17.7$, $p < .001$, partial $\eta^2 = .36$. This indicates that the children's performance on the core versus complementary elements differed significantly. However, no significant interaction effect was found, $F(3, 93) = 1.99$, $p = .12$. A comparable pattern was found for the adults whereby their core percentage score ($M = 99.7$, $SD = 1.7$) was significantly better than their complementary score ($M = 70.5$, $SD = 10.3$), $t(24) = 13.74$, $p < .001$.

Story Grammar Subunits

As displayed in Figure 3, the performance of the children, as well as the adults, was lower (range: 37.7%–70.5%) for the complementary story content than for the core elements (range: 65.3%–99.7%). Table 2 presents further details on the subtypes of the core and complementary units that the participants included during their story formulation. A further examination of the specific complementary items presented in Table 2 revealed that all participants

Figure 3. Mean percentage story grammar scores for core versus complementary elements. Bars represent standard deviations.

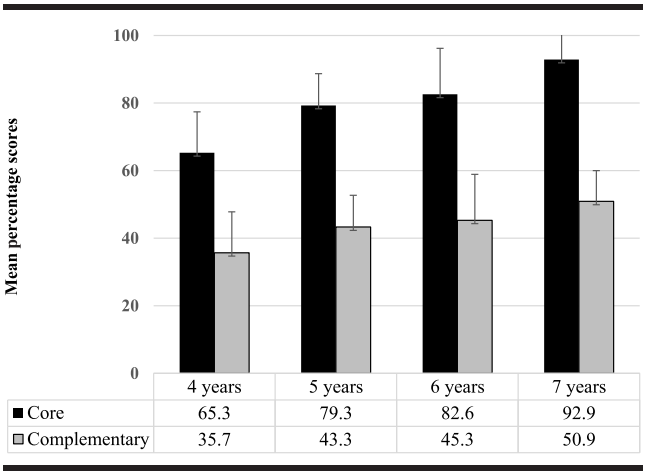


Table 2. Mean percentage of each story subunit included in the participants' narratives across age groups.

Story subunit	4 years	5 years	6 years	7 years	Adults
Character	72.9	84.1	81.9	93.2	100
Outcome	*68.8*	*81.5*	*80.4*	*88.9*	*100*
Attempt	*66.7*	*76.1*	*85.9*	*96.3*	*100*
Internal event	*60.4*	*80.4*	*81.5*	*93.5*	*99*
Setting	64.6	60.9	78.3	83.3	100
Reactions	25.4	36.1	37.4	42.2	63.6
Internal response	20.8	22.8	17.4	26.9	52
Internal plan	6.3	10.9	21.7	17.6	47

Note. Core elements are in italics.

were more likely to exclude certain elements like internal plan, internal response, and reactions of the characters than the setting and characters. For example, the mean for inclusion of internal plan was only between 6.3% and 21.7% for the children. Even the adults who served as a benchmark scored lowest on these three complementary elements: internal plan, internal response, and reactions (range: 47%–63.6%). However, this contrasts with the children's relatively higher counts and the adults' near-perfect scores for all the core elements and two complementary units, namely, "character" and "setting."

A further analysis was conducted to investigate certain story grammar categories that were less frequently included by the participants. Table 3 illustrates the story grammar elements that were included by 20% or less of the participants. For example, 4.2% of the 4-year-old children included the internal plan of the third episode in the A3 story. The data are in line with the aforementioned results, whereby the complementary units pertaining to the internal plan, the internal response, and the reactions of the characters seemed to pose a challenge for the children. First, many of the children appeared to consistently leave out the "internal plan" when telling the events of the first episode of the A3 story. Second, several groups of children, including the adults, displayed a tendency to omit the "internal response" when telling the multi-episode story (A3). Finally, the "reactions" of the elephant characters (Characters 2, 3, and 4) were more often ignored than those of the giraffe (Character 1).

Discussion

The current study examined the development of story macrostructure in Arabic-speaking children (4;0–7;11) with a focus on factors related to age and task complexity. These factors are also examined from a cross-linguistic, cross-cultural perspective to inform theory and guide practice.

Acquisition of Macrostructure Production

The results of the trend analysis revealed that the quantity of story grammar elements included in the children's stories increased with age. In general, the children's

Table 3. Percentage of participants using story grammar units.

Unit	Story and Episode (E)	4 years	5 years	6 years	7 years	Adult
Internal plan	A1	8.3	13	13	11.1	
	A3 E1	0	0	8.7	7.4	
	A3 E2	12.5	17.4			
	A3 E3	4.2	13		14.8	
Internal response	A3 E1	16.7	13			
	A3 E2	16.7	13	13		
	A3 E3	16.7		8.7	11.1	16
Reaction	A1 Character 1				14.8	
	A1 Character 2	8.3				
	A3 E1 Character 2	16.7		4.3		
	A3 E2 Character 2	8.3	8.7	8.7	14.8	
	A3 E2 Character 3	4.2	0	8.7		12
	A3 E3 Character 2	8.3		17.4		
	A3 E3 Character 4	4.2	4.3	8.7	3.7	20

Note. Only elements with scores of 20% or less were included.

story grammar scores were lower than those of the adult controls. The age progression in story grammar scores that emerged in the current study corroborates findings of other studies that have used a similar methodology, namely, storytelling in Canadian English–speaking children (Schneider et al., 2006) and Canadian French–speaking children (Gagné & Crago, 2010). Table 4 juxtaposes the mean story grammar raw scores of the participants from the current study with those from the ENNI reported in Schneider et al. (2006). A perusal of the data in Table 4 shows that the overall performances of the two cultural groups are rather comparable across ages.

The results are also in agreement with previous studies in Arabic (Ravid et al., 2014) and other languages. This suggests that the story grammar model could potentially be used to measure structural development of story formulation in Arabic. For example, they can be used as a criterion-referenced measure, along with other tools, to assess language development in Arabic-speaking children. The similarity of the scores of this sample to the original Canadian samples suggests that the measure could be adapted to other languages and cultures.

Story Grammar and Task Complexity

The results did not yield a task complexity effect on the generation of story macrostructure. Essentially, the children's mean story grammar scores for the simple A1 story were not statistically different from the multi-episode A3 story. Conversely, the adults were significantly better in telling the simple story (A1) than the more complex story (A3). This suggests that it was easier for the adults to handle the simpler story. The fact that the children's performance was similar across tasks is probably due to the fact that they are still developing their story grammar.

As can be seen in Figure 2, the performance of the different age groups was similar across the two stories, which suggests that both versions are comparable. Tentatively, either of the stories can serve as a potential assessment tool, hence eliminating the need to administer both versions of the narrative measures (A1 and A3). However, data from children with DLD would help establish whether administering only the simple A1 story would, indeed, suffice, if the test is to be later developed to be used for diagnostic purposes.

Table 4. Mean raw Edmonton Narrative Norms Instrument (ENNI) story grammar scores cross-linguistically.

Variable	4 years	5 years	6 years	7 years	Adults
A1 English (out of 12)	6.78 (2.53)	7.82 (2.62)	9.5 (1.18)	10.2 (0.76)	
A1 Arabic (out of 13)	7.04 (2.29)	7.96 (1.87)	8.7 (1.69)	9.15 (1.43)	11.36 (0.91)
A3 English (out of 37)	17.68 (6.11)	21.74 (4.21)	24.1 (3.16)	25.52 (2.76)	
A3 Arabic (out of 37)	18.12 (5.68)	22.35 (3.66)	22.91 (5.09)	26.4 (2.39)	30.88 (2.13)

Note. Comparison is between typically developing Canadian English–speaking children (Schneider et al., 2006) and the Kuwaiti Arabic–speaking children from the current study. Standard deviation is in parentheses. Both sets of data are based on the ENNI scoring scheme (i.e., 2 for core, 1 for complementary, and 0 when a unit is absent).

The final question explored distinct aspects of story grammar, such as differences in core versus complementary categories. There was a significant within-group difference in favor of core elements. The children and the adults were more likely to include the core units of the story grammar (initiating event, attempt, and outcome) when formulating a story than units that are considered complementary. This may be related to the fact that the core units of story grammar typically develop earlier than the complementary elements (e.g., Castilla-Earls et al., 2015; Hughes et al., 1997; Schneider et al., 2006; N. Stein & Glenn, 1979). As evident from Tables 2 and 3, the complementary components, especially the "internal plan," "internal response," and "reactions" of the characters, were clearly less frequent even in the story formulations of the adult control group. This corroborates previous research findings in other languages (e.g., Maviş et al., 2016; see review in Soodla & Kikas, 2010).

The exclusion of these elements is not due to any simplification on the parts of the adults, because during data gathering, they told the stories to adults and not to children. A possible reason for this is that feelings and plans are less imageable and salient in a picture story and were less likely to be used by participants. This explanation is less relevant for the adults, whose performance on the mentioned complementary elements was very low. This could be a cultural difference that is specific to (Kuwaiti) Arabic speakers who tend to focus on actions and overlook the internal states of the characters. Previous research shows that Eastern cultures (e.g., Chinese) focus more on behavior rather than thoughts and emotions in storytelling (e.g., Doan & Wang, 2010). This suggests that these elements of macrostructure could be excluded from the scoring criteria in future adaptations of these stories to Arabic.

A variation within these complementary units across story types surfaced as can be discerned in Table 3. For example, the reactions of the elephant characters (Characters 2, 3, and 4) were less likely to be included than that of the giraffe (Character 1). This may be related to the story design whereby three elephant characters differing in age and gender were used, but only one salient giraffe character. This observation cannot be ascertained as the normative ENNI data from English (Schneider et al., 2006) did not include further analyses of these types of complementary units.

In summary, these results confirm that the core story grammar units are better developmental indicators than the complementary elements for this age group. Comparing the adults' story macrostructure components to those of the children is rather revealing. Without examining an adult group, we would erroneously assume that the children were lacking in their development, given that certain aspects of the story grammar appeared to be missing in their narrations. The story grammar model is based on Western European tradition, and cultural variations need to be taken into account when examining other languages. This is especially true in the case of the low performance in some of the complementary elements.

Nonetheless, the results of the comparison between the single-episode story and the three-episode story should be interpreted with caution as the administration of the two tasks was not counterbalanced. Future studies on the effect of complexity should ensure counterbalancing in task administration to control for order and practice effects. Furthermore, future research that includes a sample of children with DLD could shed more light on some of the results and tentative conclusions of this study. Further exploration of story grammar development in Arabic using a larger sample may be warranted, given the large within-subject variations in some of the participants' scores.

Although the current study showed that the wordless picture story task yielded outcomes that are comparable to the results found in English and French in Canada, it would be useful to examine other elicitation techniques (e.g., Gagarina et al., 2012) that could facilitate the development of assessment tools that are sensitive to age and cultural variables. In the context of Arabic, it is important to use tasks that are not dependent on literacy (books) because in early years and particularly in less advantaged families, access to literacy, including the use of books, is limited. Because of the diglossic situation in Arabic, it is also crucial to study the storytelling practices at home, in nursery, and in school environments, in addition to the systematic use of dialect and the standard variety to facilitate more effective language assessment measures.

In conclusion, the results of this investigation revealed a progression by age in the development of story macrostructure in Kuwaiti Arabic–speaking children. On the other hand, no effect of story complexity was found, which indicates that the children were probably still in the process of developing their story grammar. Across all age groups, the core macrostructure elements were mastered before the complementary units. A particularly interesting finding that seems to be specific to (Kuwaiti) Arabic, compared to the results found in English (Schneider et al., 2006) and French (Gagné & Crago, 2010), is that the complementary units of "internal plan," "internal response," and "reactions" do not seem to be essential elements in the stories of the adults. However, the other two complementary elements of "character" and "setting" were included in all the adults' stories and, to a large extent, in the children's narratives. This suggests that the notions of core and elementary should be revised.

In general, the results of this study confirmed that cross-linguistic measures could be used in contexts that are culturally and linguistically different with some adaptations. The findings hold crucial clinical implications such as the need to revise the scoring system and analysis of story grammar subunits when adapting standardized measures like the ENNI into Arabic. Finally, data from children with DLD can confirm this study's tentative conclusion that the less complex, one-episode version alone may potentially be sufficient to assess the language development of this age group.

Acknowledgments

We extend a special gratitude to the children and adults who participated in this study and to the Ministry of Education

and the public school teachers for facilitating the project. We gratefully acknowledge the assistance of Reem Almutairi, Jenan Hassan, and Hanan Aljeeran for their efforts in recruiting, screening, and testing the participants. We thank Saleh and Abyan Abshir and Sarah and Fatima Alquraini for their assistance in transcribing and coding the data. An earlier version of the paper was presented as a poster at the 30th World Congress of the IALP held in Dublin, Ireland, in 2016.

References

Berman, R. A. & Slobin, D. I. (Eds.). (1994). *Relating events in narrative: A cross-linguistic developmental study*. Hillsdale, NJ: Erlbaum.

Bishop, D. V. M., & Edmundson, A. (1987). Language impaired 4-year-olds: Distinguishing transient from persistent impairment. *Journal of Speech and Hearing Disorders, 52*(2), 156–173.

Bloch, C. (1999). Literacy in the early years: Teaching and learning in multilingual early childhood classrooms. *International Journal of Early Years Education, 7*(1), 39–59.

Castilla-Earls, A. P., Petersen, D., Spencer, T., & Hammer, K. (2015). Narratives and story grammar development in preschool Spanish-speaking children. *Early Education and Development, 26,* 1166–1186.

Chan, L.-L. S. (2003). *Production of story grammar in Cantonese-speaking children with and without specific language impairment* (Unpublished BSc thesis). The University of Hong Kong, Pokfulam, Hong Kong.

Doan, S. N., & Wang, Q. (2010). Maternal discussions of mental states and behaviors: Relations to emotion situation knowledge in European American and immigrant Chinese children. *Child Development, 81*(5), 1490–1503.

Ferguson, C. A. (1959). Diglossia. *Word, 15*(2), 325–340.

Fichman, S., Altman, C., Voloskovich, A., Armon-Lotem, S., & Walters, J. (2017). Story grammar elements and causal relations in the narratives of Russian–Hebrew bilingual children with SLI and typical language development. *Journal of Communication Disorders, 69,* 72–93.

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,* 1–140. http://publikationen.ub.uni-frankfurt.de/frontdoor/index/index/docId/34782

Gagné, A., & Crago, M. (2010). The use of the ENNI to assess story grammar competency of school-aged French speaking children with and without specific language impairment. *Canadian Journal of Speech-Language Pathology and Audiology, 34*(4), 231–245.

Griffin, T. M., Hemphill, L., Camp, L., & Wolf, D. P. (2004). Oral discourse in the preschool years and later literacy skills. *First Language, 24*(2), 123–147.

Guo, L. Y., & Schneider, P. (2016). Differentiating school-aged children with and without language impairment using tense and grammaticality measures from a narrative task. *Journal of Speech, Language, and Hearing Research, 59*(2), 317–329.

Hayward, D., Schneider, P., & Gillam, R. (2009). Age and task-related effects on young children's understanding of a complex picture story. *Alberta Journal of Educational Research, 55*(1), 54–72.

Hughes, D., McGillivray, L., & Schmidek, M. (1997). *Guide to narrative language: Procedures for assessment*. Eau Claire, WI: Thinking Publications.

Justice, L. M., Bowles, R. P., Kaderavek, J. N., Ukrainetz, T. A., Eisenberg, S. L., & Gillam, R. B. (2006). The index of narrative microstructure: A clinical tool for analyzing school-age children's narrative performances. *American Journal of Speech-Language Pathology, 15*(2), 177–191.

Kelly, K. L. (2010). Interpreting the "Sounds of Silence" in dual language preschools in Qatar: Teacher's use of story retelling and language sampling. *SIG 14 Perspectives on Communication Disorders and Sciences in Culturally and Linguistically Diverse (CLD) Populations, 17*(1), 11–24.

Kupersmitt, J. R., & Armon-Lotem, S. (2019). The linguistic expression of causal relations in picture-based narratives: A comparative study of bilingual and monolingual children with TLD and DLD. *First Language, 39*(3), 319–343.

Leikin, M., Ibrahim, R., & Eghbaria, H. (2014). The influence of diglossia in Arabic on narrative ability: Evidence from linguistic and narrative structure of discourse among pre-school children. *Reading and Writing, 27*(4), 733–747.

Lofranco, L. A. L., Peña, E. D., & Bedore, L. M. (2006). English language narratives of Filipino children. *Language, Speech, and Hearing Services in Schools, 37*(1), 28–38.

Maviş, İ., Tunçer, M., & Gargarina, N. (2016). Macrostructure components in narrations of Turkish–German bilingual children. *Applied Psycholinguistics, 37*(1), 69–89.

McCabe, A., Bliss, L., Barra, G., & Bennett, M. (2008). Comparison of personal versus fictional narratives of children with language impairment. *American Journal of Speech-Language Pathology, 17*(2), 194–206.

McConnell, G. E. (2011). *Story presentation effects on the narratives of preschool children from low and middle socioeconomic homes* (Unpublished doctoral dissertation). The University of Kansas, Lawrence, KS, United States.

Merritt, D. D., & Liles, B. Z. (1989). Narrative analysis: Clinical applications of story generation and story retelling. *Journal of Speech and Hearing Disorders, 54*(3), 438–447.

Pankratz, M. E., Plante, E., & Insalaco, D. M. (2007). The diagnostic and predictive validity of *The Renfrew Bus Story*. *Language, Speech, and Hearing Services in Schools, 38*(4), 390–399.

Paul, R., & Smith, R. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech and Hearing Research, 36*(3), 592–598.

Peat, J. (2001). *Health science research: A handbook for quantitative methods*. Sydney, Australia: Allen & Unwin.

Pinto, G., Tarchi, C., & Bigozzi, L. (2016). Development in narrative competences from oral to written stories in five- to seven-year-old children. *Early Childhood Research Quarterly, 36,* 1–10.

Ravid, D., Naoum, D., & Nasser, N. (2014). Narrative development in Palestinian Arabic. In E. Saiegh-Haddad & M. Joshi (Eds.), *Handbook of Arabic literacy: Insights and perspectives*. Dordrecht, the Netherlands: Springer.

Saidi, D. (2014). *Développement de la compétence narrative en arabe tunisien: rapport entre formes linguistiques et fonctions discursives* [Development of narrative competence in Tunisian Arabic: Relationship between linguistic forms and discourse functions] (Unpublished PhD dissertation). Université Lumière Lyon 2, Lyon, France.

Schneider, P. (1996). Effects of pictures versus orally presented stories on story retellings by children with language impairment. *American Journal of Speech-Language Pathology, 5*(1), 86–96.

Schneider, P., & Dubé, R. V. (2005). Story presentation effects on children's retell content. *American Journal of Speech-Language Pathology, 14*(1), 52–60.

Schneider, P., Dubé, R. V., & Hayward, D. (2005). *The Edmonton Narrative Norms Instrument*. Retrieved from University of Alberta Faculty of Rehabilitation Medicine website: http://www.rehabmed.ualberta.ca/spa/enni

Schneider, P., Hayward, D., & Dubé, R. V. (2006). Storytelling from pictures using the Edmonton Narrative Norms Instrument. *Journal of Speech-Language Pathology and Audiology, 30*(4), 224–238.

Shiro, M. (2003). Genre and evaluation in narrative development. *Journal of Child Language, 30*(1), 165–195.

Soodla, P., & Kikas, E. (2010). Macrostructure in the narratives of Estonian children with typical development and language impairment. *Journal of Speech, Language, and Hearing Research, 53*(5), 1321–1333.

Stein, C. L. (2004). Analysis of narratives of Bhutanese and rural American 7-year old children: Issues of story grammar and culture. *Narrative Inquiry, 14*(2), 369–394.

Stein, N., & Glenn, C. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *New directions in discourse processing* (Vol. 4, pp. 53–120). Norwood, NJ: Ablex.

Trabasso, T., Stein, N. L., Rodkin, P. C., Munger, M. P., & Baughn, C. R. (1992). Knowledge of goals and plans in the online narration of events. *Cognitive Development, 7*(2), 133–170.

van Kleeck, A., Lange, A., & Schwartz, A. L. (2011). The effects of race and maternal education level on children's retells of the Renfrew Bus Story–North American edition. *Journal of Speech, Language, and Hearing Research, 54*(6), 1546–1561.

Appendix

Sample Story Transcription and Scoring From a Kuwaiti Arabic–Speaking Child

Child's Name: Mansoor **Age:** 4;4 (years;months) **Gender:** Male **Grade:** Kindergarten 1

Once upon a time, long time ago.	كان يا مكان في قديم الزمان	C
There was a horse (means giraffe), he dropped an apple (means the ball) in the swimming pool.	كان حصان طيح التفاحه في الماي السباحه	C
Then the elephant got upset with him.	بعدين الفيل هاوشته	C
Then he stepped into the water/swimming pool.	بعدين دخل داخل ماي السباحه	C
Then (xxx) he gave it to the elephant.	بعدين (xxx) عطاها الفيل	C
Then she said, "thank you!"	بعدين قالت شكرا	C
That's it!	و بس	C

Adapted Story Grammar Scoring Sheet[a]
Based on the Edmonton Narrative Norms Instrument (ENNI)

A1 Story (ball goes in the water)

Child's Code: 120M

Child's Name: Mansoor **Age:** 4;4 (years;months) **Gender:** Male

SG unit	Acceptable	Score	
Character 1 (C1)	Giraffe/male or female/boy or girl (any type of animal such as horse) (Not acceptable: pronoun).	0	(1)
Character 2 (C2)	Elephant/male or female/boy or girl (or any type of animal such as cow) [not pronoun]	0	(1)
Setting	Swimming pool	0	(1)
Initiating event	Had a ball/playing with ball/want to play ball. Acceptable if ball is substituted with apple or some object. Ball goes in water/pool/sand/mud/ocean. Ball is in water. They see a ball.	0	(1)
Internal Response	One/both want to get ball. Elephant says, e.g., "look what happened" "what am I going to do?" Elephant upset/sad.	0	(1)
Internal Plan	Giraffe decides to/thinks he will get the ball.	(0)	1
Attempt	Giraffe jumps in pool/swims toward ball/tries to get ball. [not: giraffe swimming (without goal); giraffe falls in water]	0	(1)
Outcome	Giraffe gets ball/gives ball to elephant. [not: elephant gives ball to giraffe, unless it is noted as unexpected, e.g., 'but instead, Elephant gets it and gives it to him']	0	(1)
Reaction C1	Giraffe is happy/proud/smiles. Giraffe says "You're welcome". Giraffe's teeth are chattering/giraffe is cold/wet.	(0)	1
Reaction C2	Elephant is happy/is grateful/says thank you. Elephant hugs the ball [not: holds/has the ball].	0	(1)
Reaction of both or Unknown	"They" are happy [code only as replacement for reaction of Character 1 or 2; there should not be more than two reactions total].	(0)	1
	Total Raw Score	**8**	

[a]In the original ENNI, the story grammar (SG) components pertaining to internal event, attempt, and outcome were each 2 points and not 1 point as in the current study. Gender confusion, for example, labeling the "giraffe" character as a female was accepted for linguistic reasons given that the noun "giraffe" (zarafa) in Arabic ends with a feminine marker.