

# Local reference data for the Multilingual Assessment Instrument for Narratives in Pristina, Kosovo

**Charlotte Hilker**

University of Iowa, USA

**Ardita Uka-Rexha**

Instituti Kosovar për Logopedi, the Republic of Kosovo

**Marigona Kozmaqi**

Instituti Kosovar për Logopedi, the Republic of Kosovo

**Ardea Çakmaku**

Instituti Kosovar për Logopedi, the Republic of Kosovo

**Albina Krasniqi**

Instituti Kosovar për Logopedi, the Republic of Kosovo

**Philip Combiths**

University of Iowa, USA

---

Currently, there are no standardized or norm-referenced language assessment tools appropriately contextualized for use by speech-language therapists in Kosovo. In this community-based study in collaboration with the non-governmental organization Instituti Kosovar për Logopedi, local reference data for the Albanian adaptation of the Multilingual Assessment Instrument for Narratives (MAIN) were collected in Pristina, Kosovo. Participants were 58 school-aged children (ages 5;7–10;11), who completed tellings and retellings of two stories (Cat and Baby Goats) from the MAIN in Albanian. Information about the children's demographic and language history and environment was collected using parent-report questionnaires. Reference data for MAIN scores are provided with summary statistics and confidence interval estimates by age group (5–6, 7, 8, 9, and 10 years old). The results indicate that scores from the story structure and

comprehension sections of the retelling task may differentiate MAIN performance across age groups most effectively in this population.

---

## 1 Introduction and background

Narrative language skills, including the ability to comprehend and convey stories, are a major component of and can provide insights into the overall language development of school-aged children. Speech-language therapists use reference data on narrative language abilities as a tool in assessing receptive and expressive language, diagnosing language disorders, and identifying treatment needs for students. In Kosovo, a country in the Balkan region of Europe, there are not yet any normative data on school-age language development, including narrative language skills, to support evidence-based speech-language therapy. Additionally, the linguistic environment in Kosovo is unique (as described in section 1.3, below) and reference data from neighboring countries may not be generalizable to Kosovan populations. This study, conducted in partnership between the non-governmental organization *Instituti Kosovar për Logopedi* and the University of Iowa, seeks to support speech-language therapists in Pristina, Kosovo by collecting and analyzing narrative samples with the Albanian version of the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al., 2019) from 58 school-aged children. In the remainder of Section 1, we review the existing literature on school-age language in Kosovo and provide historical context for language use in Kosovo. We then describe the planning and methodology of this community-based participatory research (i.e., Israel et al., 1998) to establish reference data collected for speech-language therapists in the context of Pristina, Kosovo (Section 2). Next, we describe the reference data (Section 3) and discuss how these data could be utilized in clinical practice through application to a case study (Section 4). Finally, we discuss the limitations of this study and future directions (Section 4).

### 1.1 Existing data on school-age language in Kosovo

Currently, only one study is known to describe the prevalence of speech and language disorders in Kosovo. Nešić, Minić, and Jakšić (2011) distributed surveys to 36 teachers in the northern regions of Kosovo (Mitrovica, Zvečan, Leposaviq, and Zubin Potok) and estimated the prevalence of speech and language disorders in 10- and 11-year-old children at approximately 7.64%, based on a written questionnaire distributed to 36 teachers of 3<sup>rd</sup> and 4<sup>th</sup> grade classrooms. This study collected data from teachers in a region of Kosovo with a high population of ethnic Serbs; therefore, the linguistic makeup of the sample group would likely not generalize to the ethnic Albanian majority (i.e., the Gheg Albanian-speaking areas of Kosovo, such as Pristina). More data is needed on Gheg-Albanian speaking school age children in Kosovo, as existing data is insufficient to serve the needs of speech-language therapists. The current study will begin to address this gap by providing initial reference data collected in the capital city, Pristina.

## **1.2 Reference data for MAIN in other populations**

Several studies have utilized the MAIN to assess language abilities in school-aged children, providing reference data for distinct linguistic communities. Most studies have focused on ages 4-7, fewer having focused on ages 8 and older, and age effects have been found in monolinguals and bilinguals (see the review by Lindgren et al., 2023). For example, Lindgren (2019) investigated age effects in Swedish monolingual children and found that though there was a large increase in performance between 4 years 4 months and 5 years 10 months, performance seemed to plateau around 6 years old. Bohnacker (2016) found age effects in production and comprehension between 5- and 6-to-7-year-olds in Swedish-English bilingual children. Not all studies have found age effects for comprehension (Roch & Hržica, 2020). Studies documenting MAIN performance in Kosovo do not currently exist.

## **1.3 Language use in Kosovo**

Diagnosis and treatment of language disorders is most effective when the cultural and linguistic context of the child and their environment is taken into consideration (Hyter & Salas-Provence, 2021). In Kosovo, the cultural and linguistic environment is unique, yet there are currently no diagnostic tools developed with data collected in Kosovo available to assess school-age language in Kosovan children.

Kosovo's unique linguistic landscape is a result of past language policy, a history of oppression, and a necessarily multilingual environment. The most widely spoken language in Kosovo is Albanian, specifically the Gheg dialect of Albanian. Gheg Albanian is mutually intelligible with Tosk Albanian, the other major dialect of Albanian spoken in Kosovo. Despite many similarities, there are substantial linguistic distinctions between the Gheg and Tosk varieties of Albanian. For example, a Tosk-Albanian speaker may say *Unë dua më shkoj* 'I want to go', whereas a Gheg-Albanian speaker would say *Unë du më shku* 'I want to go (Mëniku & Campos, 2011). Kosovo was an autonomous state of Yugoslavia when steps were taken to standardize Albanian. Tosk Albanian, the main dialect spoken in Albania, was established as the standard in Kosovo (alongside Serbo-Croatian) two years after a 1972 congress on standard Albanian that was vetted and approved by both the governments of Yugoslavia and Albania. This decision was motivated by the geopolitics of Albania and Yugoslavia in the 1960s and 1970s. Currently, there exists diglossia among Gheg speakers such that Gheg is often spoken at home and in everyday life, but Tosk is taught in schools, printed in books and newspapers, and used in more formal settings (Kamusella, 2016).

Kosovo's unique linguistic environment is also heavily impacted by the conflicts of the past 50 years. The year 1989 marked the end of Kosovo's autonomy in Yugoslavia and led to the suppression of Albanian language and culture in Kosovo. Schools were segregated and students in Kosovo were forced into underfunded schools with very few resources or underground homeschools, taught by their community members. Books printed in the Albanian language were also destroyed by the Serbian military. Printed material was thus available only in Serbian, if available at all (Shahini, 2016). Additionally, many girls did not attend school at this time due to concerns for their safety. Consequently, as of 2001, 40% of the women between 16 and 35 years of age were illiterate, and nearly 80% of the surveyed women did not complete

the compulsory 8 years of Basic Education (Haneman, 2006). Literacy and language use among parents can affect the language development of their children; thus, the linguistic environment in Kosovo during and after the late 20<sup>th</sup> century is likely to still impact language use today (Regis College, 2023).

The linguistic and cultural suppression of Albanian in Kosovo has influenced the languages spoken in Kosovo and resulted in a dynamic multilingual environment with differing expectations on each generation. Prior to the oppression that began in 1989, Kosovan Albanian students learned Serbian in schools (Quell, 2017). However, with the political tension between Serbs and Albanians in present-day Kosovo, English has been increasingly used as a lingua franca between the two groups. English also serves as a lingua franca between Kosovo and the international community as it has globalized in the 21st century (Lohaj, 2018). Today, it is common for children in Kosovo to speak both Gheg and Tosk Albanian, in addition to English. It is common for their parents to speak Serbian in addition to these languages. Additionally, many Kosovans also speak Turkish and other European languages, contributing to a unique and diverse multilingual environment.

Although Kosovo formally declared independence from Serbia in 2008, not all countries currently recognize Kosovo's sovereignty, including Serbia, Russia, China, and five European Union (EU) member states (UK Parliament House of Commons Library, 2024). Kosovo is currently recognized as sovereign and independent by 90 countries, including 22 EU member states, Japan, Canada, and the United States (US Department of State, n.d.).

#### **1.4 Community collaborator**

The present study is a community-based participatory research project (following Israel et al., 1998) that was conceptualized, planned, and conducted based on the immediate needs identified by members of the community this research seeks to support. The community collaborator for this project is *Instituti Kosovar për Logopedi* (IKL), a non-governmental organization offering free speech therapy services in three public schools and an SOS Children's Village school, which primarily serves orphaned or abandoned children. IKL provides direct one-on-one services for children who have speech or language disorders primarily through student volunteers. Speech therapy in Kosovo is not widely accessible to children through public services, so IKL fills a necessary gap in care with their free services. Speech-language therapy is also a relatively new field in Kosovo, and IKL began its advocacy in 2016 and its free speech-language services in the schools started in 2022.

Speech-language therapists at IKL and the first author developed a reciprocal and collaborative relationship during the year in which this study was designed and conducted. The team at IKL identified the initial resources needed to strengthen the services provided by IKL. In subsequent discussions with the first author, a standardized language assessment tool with reference data collected locally in Pristina, Kosovo was identified as the priority to support the language-related needs of school-age children in their community. The team at IKL were integral to study design, participant recruitment, assessment administration, and reliable assessment scoring. Their collective knowledge of the community was crucial for the completion of this project in an effective and culturally responsive manner.

### **1.5 The present study**

The unique linguistic environment of Kosovo is such that generalization of language development norms from nearby communities is not appropriate. Therefore, this project aims to collect local reference data to better understand and differentiate the language abilities of school-age children in Pristina, Kosovo. Given the localized, community-based nature of this study, the population served by IKL is sampled directly: Gheg Albanian-speaking children aged 5 to 10 years old attending urban-area schools in Pristina, Kosovo. The data collected consist of answers to parental questionnaires and expressive language samples in the form of narratives that were elicited using the Albanian adaptation of MAIN (Gagarina et al., 2019), as it is a dynamic and practical standardized tool for language assessment in this population. Using this narrative assessment, we seek to provide insight into the language development of school-age children in the urban capital city of Kosovo with direct application to language assessment and screening for developmental language disorder in speakers of Albanian.

## **2 Methods**

This study was conducted in accordance with the ethical standards of the Helsinki declaration and its later amendments (World Medical Association, 1964). The research protocol was reviewed and approved by the Institutional Review Board (IRB) of the University of Iowa (IRB Approval Number: 202311500). Written informed consent from parents and verbal assent of children were obtained for all participants included in the study.

### **2.1 Participants**

Participants were between age 5;7 and 10;11 ( $N=58$ , mean age=8.4). Packets containing an informed consent form, a Demographic Questionnaire, and a Home Language Questionnaire were distributed to the families of every child between the ages of 5 and 11 years old enrolled at the three sites. Sixty-two packets were returned. Of these 62, three participants did not assent to participating in the study; therefore, 59 subjects participated. Only one student in the oldest age group (11 years old) and two students in the youngest age group (5 years old) completed the MAIN. Initially, these participants with outlying ages were combined with the nearest age group (i.e., 5–6 years old and 10–11 years old). However, the 11-year-old participant had significantly different scoring from the 10-year-old group on the *Cat* story task (independent samples *t*-test,  $p < .01$ ). Therefore, the 10- and 11-year-old data were not combined, and the singular 11-year-old data were excluded from this study. The 5- and 6-year-old data, however, were not significantly different on any of the MAIN sub scores ( $p > .05$ ). Therefore, data from the two 5-year-old participants were included in a combined 5–6-year-old group. In total, one participant was excluded, and 58 participants were included in the study. Age ranges, age averages, and gender distributions groups are illustrated in Table 1.

**Table 1.** Participant breakdown by age and gender (n=58).

Age group	N (total)	n (male)	n (female)	Mean age (years)	Age range (years)
5–6	11	4	7	6.45	5.58–6.92
7	13	6	7	7.49	7.00–7.92
8	9	3	6	8.56	8.00–8.92
9	14	7	7	9.37	9.08–9.83
10	11	3	8	10.53	10.00–10.92

About 93% of parents reported their child’s ethnicity to be Albanian; one child (i.e., 2%) was reported to be Albanian, Italian, and German, and 5% of parents did not respond to this item on the questionnaire. It should be noted that the ethnicity in Kosovo is estimated to be about 92.9% Albanian, 1.6% Bosniak, 1.5% Serb, 1.1% Turk, 0.9% Ashkali, 0.7% Egyptian, 0.6% Gorani, 0.5% Romani, and 0.2% other/unspecified (2011 estimate; CIA, n.d.). Thus, the current sample may be more representative of the population IKL directly serves (i.e., the Albanian population of Pristina, Kosovo), rather than the entire country. For instance, areas with a higher population of Kosovan Serbs would not be represented by these data, as they are more likely to use Serbian as a primary language at home and at school.

Before completing the MAIN, students completed a short task to estimate their primary spoken variety of Albanian. In the tasks, students were asked to name pictures of items that had distinct names in Gheg and Tosk Albanian (i.e., *green, phone, roof, jacket*). In Gheg and Tosk respectively, these words are “e gjelbert”/“e jeshile”, “cellular”/“mobil”, “çati”/“kulem” and “jakne”/“jaket”. All students used Gheg Albanian to name each item and participants were thus considered to be speakers of Gheg Albanian.

**2.2    *Participants’ language and demographics***

A *demographic questionnaire*, written in Albanian, was sent home to parents in a packet with the informed consent. The questionnaire asked parents to report their age, their education level, and how long they had lived in Kosovo. The results of this questionnaire offered some information on the historical, cultural, and linguistic context of Kosovo that may have impacted our participants and their families.

The *Home Language Questionnaire* (Combiths, 2023) was adapted to Albanian for this study. The questionnaire asked parents to report the languages spoken by their child, their child’s proficiency in these languages, and what languages their child was exposed to at home and at school for each year of their lives. This questionnaire also captured information on the child’s language environment by asking the languages spoken by other people in the household and how often each person spends time with the child. Some limited developmental information was gathered about the children, such as when they spoke their first word, when they were first exposed to English, and an opportunity to express concerns about their child’s language. These questions are displayed in Table 2.

**Table 2.** Questionnaire items related to speech-language development.

Item
How old was your child when they first started speaking words?
Do you have any concerns about your child's hearing?
How does your child express their needs?
Does your child talk like other kids in your community or in your family that are the same?
Do you have any concerns about your child's speech or language?
Do you have any concerns about your child's health or development?

**2.3 Albanian adaptation of the Multilingual Assessment Instrument for Narratives**

MAIN (Gagarina, et al., 2019) was developed by the Narrative and Discourse working group within COST (European Cooperation in Science and Technology) Action IS0804 as part of the set of assessment tools Language Impairment in a Multilingual Society: Linguistic Patterns and the Road to Assessment. It is part of a test battery known as the Language Impairment Testing in Multilingual Settings (LITMUS), also developed within COST Action IS0804. First published in 2012 and revised in 2019, MAIN assesses the child’s ability to understand, tell, and/or retell a story using a sequence of six pictures for each of the four stories (*Cat, Dog, Baby Birds, Baby Goats*), which were designed to be age-appropriate and culturally relevant across different languages and cultures.

MAIN has been adapted into Tosk Albanian, the dialect primarily spoken in Albania, by Enkeleida Kapia based on the revised version in English. This is the version used in this study. Prior to beginning testing or recruitment, a meeting was held with the speech-language therapists at IKL to discuss adaptation of the MAIN to Gheg Albanian from the Tosk Albanian version. The consensus was that the MAIN did not need to be changed from its Tosk Albanian form to be administered in Kosovo. As described above, Gheg Albanian and Tosk Albanian are mutually intelligible. Tosk Albanian is also taught and used in schools and widely used in written contexts, such as textbooks and newspapers. Thus, Tosk Albanian is understood by Gheg-Albanian-speaking children attending school in Kosovo. Second, the scoring manual does not restrict the language variety children use in their responses, and thus children could provide responses in either Tosk or Gheg Albanian without restriction. During testing, children were encouraged to use whichever language variety was most comfortable. If children were hesitant to respond, examiners were instructed to prompt with: “Tell me just like you would talk at home”.

**2.4 Testing procedures**

Each child completed the *Baby Goats* story as a retelling task followed by the *Cat* story as a telling task from the Albanian MAIN. Two examiners (university students studying speech-language therapy employed by IKL) administered the tasks to all children in a quiet room at their school. The administration procedures were followed as specified in the MAIN manual (Gagarina et al., 2019), with any differences specified below. The child sat opposite the examiner. Examiners began by building rapport with each child (asking about the child’s favorite TV show). The picture stimuli were cut out as a sequence of story cards, presented as

three identical stacks of picture story cards. This is a slight deviation from MAIN manual procedures which suggest using three identical picture strips in envelopes. The child was asked to pick any one of the piles; however, all three piles contained six identical picture sequences of the same target *Baby Goats* story. The child thus believed they were making a meaningful choice between three possible stories but in fact would always begin with the *Baby Goats* story. This method minimized the effect of shared knowledge during the narrative task, encouraging the child to elaborate on the details of the story for a potentially unfamiliar listener (Gagarina et al., 2019). The cards were presented as the picture strips would have been presented per the manual, revealing part of the story at a time and not allowing the examiner to see the pictures. For the retelling task, the child was told the story by the examiner. For the telling task, the same methods were repeated using the *Cat* story cards, except that the child did not hear the examiner tell the story. After the child told or retold the story, comprehension questions as provided in the MAIN manual were asked of the child. The *Cat* telling task was administered immediately following the *Baby Goats* retelling task. All language samples were audio recorded using Tascam DR-07X portable audio recorders at a sampling rate of 44.1 kHz and saved in uncompressed WAV format for later review.

## 2.5 MAIN scoring

The MAIN assessments were scored according to the MAIN protocol. Examiners accepted lexical choices and syntactic structures in either Gheg Albanian or Tosk Albanian as correct responses. Inter-rater reliability of MAIN scoring was assessed for 9 out of 58 assessments (16% of the total sample) by having a second examiner, who did not have access to the original examiner's scores, independently rescore them. Agreement between the two examiners was 96%.

The MAIN produces four different scores for each child: a story structure score, a structural (or episodic) complexity score, an internal state terms score, and a comprehension score (Bohnacker & Gagarina, 2019). The first three scores are based on the child's narrative retelling or telling, and the comprehension score is determined from the child's ability to answer the ten comprehension questions about the story after the narrative portion is completed.

The MAIN stories presented to the participants consisted of three episodes, each with an initiating event, goal, attempt, outcome, and reaction. The *story structure score* was derived as a combined score of these sections, in which the child received one point for each component (initiating event, goal, attempt, outcome, reaction) they provided for each episode as well as up to two points for specifying setting (one point for time, one point for place). The maximum score for story structure was 17.

The *structural complexity scores* were determined by the story sequences the child produced (i.e., an attempt-outcome sequence, a goal-attempt sequence, a goal-outcome sequence, or a goal-attempt-outcome sequence). The structural complexity score was the total count of all attempt-outcome, goal-attempt, goal-outcome, and goal-attempt-outcome sequences. Thus, the more story sequences the child included, the higher their resulting structural complexity score. For example, a child who produced a goal-attempt-outcome, an attempt-outcome, or a goal-attempt sequence would receive 1 point for that episode. There were

three possible episodes in the story, so the maximum structural complexity score for a story was 3.

The *internal state terms score* was derived as a count of the number of words (i.e., tokens) the child used in their narrative to describe the internal states of the characters (e.g., *see, feel, thirsty, hungry, asleep, happy, sad, want, think, say, ask*). A child could produce any number of internal state tokens; therefore, there was no maximum *internal state terms score*.

The *comprehension* section began with a warm-up question (“Did you like the story?”) that was not scored, followed by ten “why” and “how” questions about the story targeting goals and internal states (e.g., “Why does the cat grab the fish?”). The child received one point for each correct response with a maximum possible *comprehension score* of 10.

## 2.6 Analysis procedure

Data analyses were conducted using R, with all results grouped by age, scoring section sections (story structure, structural complexity, internal state terms, comprehension), and task (*retelling, telling*). Participants were grouped by year of age (e.g., the 7-year-old group included ages 7 years, 0 months through 7 years, 11 months). Exceptions to this were the youngest ages (5 and 6 years old), which were grouped together due to fewer participants in this age range (i.e., only 2 participants were 5 years old). The participant age groups were thus 5–6, 7, 8, 9, and 10 years old.

Descriptive statistics (i.e., mean, standard deviations, minimum score, maximum score, and quartiles) were generated using the *dplyr* package (Wickham et al., 2023). Confidence intervals for the means and standard deviations were estimated with bootstrapping using the *boot* package (Canty & Ripley, 2024). This non-parametric approach resamples the data with replacement to simulate the distribution of each descriptive statistic (Hinkley & Davison, 1997). For each measure, 5,000 samples were generated ( $R = 1000$ ). The BCa (bias-corrected and accelerated) estimation method was used. This method adjusts for bias and skewness in the bootstrap distribution for more robust interval estimates with data that may not meet the assumptions of normality. To determine the effect of age group on MAIN scores in this sample, we conducted a one-way analysis of variance (ANOVA) for each of the four scores (story structure, structural complexity, internal state terms, comprehension) within each task (*retelling, telling*) using the *car* package (Fox & Weisberg, 2019).

## 3 Results

### 3.1 Demographic questionnaire

Fifty out of 58 parents completed the Demographic Questionnaire. 80% (40/50) reported that they have lived their entire lives in Kosovo, whereas 12 % (6/50) reported that they have not lived in Kosovo their entire lives, and 8% (4/50) did not respond to this item. Parents each reported their highest level of education with reports ranging from middle school completion to graduate degrees. About one quarter of parents completed middle school as their highest educational attainment. About half of parents who responded to this item completed a bachelor's degree or higher. The education levels of parents are shown in Table 3.

**Table 3.** Highest level education of parents reported in Demographic Questionnaire (n=50).

Parental Education Level	n	%
Middle school (up to year 8)	12	24
High school (up to year 12)	12	24
College (bachelor's degree)	12	24
Graduate degree (master's degree or above)	14	27

### 3.2 Home language questionnaire

Fifty-four out of 58 parents completed at least part of the Home Language Questionnaire. All 54 parents (100%) indicated Albanian as their child's first language. Fifty-one parents reported a proficiency level for their child's first language. Forty-nine parents reported Albanian as the child's first language with "Very good" proficiency. Two parents reported Albanian as the child's first language with "Good" proficiency, and both of these children were 5 years old. Three parents did not answer this question. All parents who filled out the questionnaire except one (98%) reported English as their child's second language.<sup>1</sup> The average age of initial exposure to English was 3 years, 2 months (range: 0-6 years; SD: 1.38).

Ten children were reported to also speak a third or fourth language. Third and fourth languages were reported as follows: five children spoke German, three children spoke Turkish, one child spoke Bulgarian, one child spoke Russian, one child spoke Bosnian, and one child spoke Arabic.

No parent reported concerns about their child's hearing. Parents reported their child's first words as early as 6 months old and as late as 3 years old ( $M=1.37$ ,  $SD=0.57$ ). In response to the question "Does your child talk like other kids in your community or in your family that are the same age?", 3 parents indicated "No" and 51 indicated "Yes".

Six parents reported developmental or language concerns on the questionnaire, as listed below:

- *Mendoj që nganjëherë nuk mund të shprehet shumë mirë në gjuhën shqipë, por mendon në gjuhë angleze pastaj e përkthem atë në trurim e saj.* 'I think that sometimes she can't express herself very well in Albanian, but she thinks in English and then translates it in her brain.'
- *Probleme me dialektin e Kosoves.* 'Problems with Kosovan dialect.' (from a family recently relocated)
- *Mos shqiptimi i mirë i fjalëve.* 'Not pronouncing words well.'
- *Sjellja ndaj tij me presion të lart resepektivisht ton të lart ja humb -- aftësin për përgjigje adekuate.* 'Behaving towards him with high pressure or high tone, respectively, loses the ability for an adequate response.'
- *Është natyre më e ndishme, dhe ndonjëherë i duhet pak kohë që të shprehet me lirshe me përsona të panjohur për të.* 'He is more sensitive in nature, and sometimes takes a while to express himself freely with unfamiliar people.'
- *Unë shqetson pesha e ulët.* 'I'm worried about the low weight.'

<sup>1</sup> The English proficiency was reported as "Poor" (n=1), "Fair" (n=8), "Good" (n=20), "Very Good" (n=20), or no response (n=5).

3.3 MAIN results

All 58 children completed both MAIN narrative tasks. Narratives were elicited using the *Baby Goats* story as a retelling task and *Cat* story as a telling task. Children from all age groups were able to complete both tasks.

Descriptive statistics, with estimated confidence intervals where applicable, are provided by age group, narrative task (retelling, telling), and MAIN scoring section (story structure, structural complexity, internal state terms, and comprehension) in Table 4 for the retelling task and Table 5 for the telling task.

Table 4. Results for the narrative retelling task *Baby Goats* story.

Age	Score	M	M 95% CI	SD	SD 95% CI	Min	Q1	Med	Q3	Max
5–6 (n=11)	SS	5.60	3.60 – 7.70	3.50	2.53 – 5.25	0	3.25	6	7.75	12
	SC	1.10	0.70 – 1.30	0.57	0.42 – 0.92	0	1	1	1	2
	IST	1.40	0.50 – 2.1	1.35	1.07 – 1.58	0	0	1.5	2.75	3
	C	8.18	6.91 – 9.00	1.72	1.25 – 2.29	5	7	9	9.5	10
7 (n=13)	SS	6.54	5.46 – 7.46	1.90	1.35 – 2.77	3	6	7	7	10
	SC	1.15	–	0.38	–	1	1	1	1	2
	IST	1.92	1.08 – 2.46	1.32	0.99 – 1.69	0	1	2	3	4
	C	8.92	8.08 – 9.23	1.04	0.49 – 1.79	6	9	9	9	10
8 (n=9)	SS	8.00	6.11 – 9.56	2.78	1.81 – 3.74	4	7	8	10	12
	SC	1.22	–	0.44	–	1	1	1	1	2
	IST	2.22	1.22 - 2.78	1.20	0.71 – 1.79	0	2	2	3	4
	C	9.11	–	1.05	–	8	8	10	10	10
9 (n=14)	SS	8.46	7.23 – 9.38	1.98	1.55 – 2.56	5	7	9	10	11
	SC	1.23	–	0.44	–	1	1	1	1	2
	IST	2.92	2.23 – 3.54	1.32	0.99 – 1.73	1	2	3	4	5
	C	9.93	–	0.27	–	9	10	10	10	10
10 (n=11)	SS	9.11	7.89 – 10.89	2.26	1.27 – 3.50	7	7	9	10	14
	SC	1.56	–	0.53	–	1	1	2	2	2
	IST	3.00	1.35 – 4.44	2.50	1.87 – 3.05	0	1	3	6	6
	C	9.20	7.70 – 9.80	1.48	0.32 – 2.00	6	9.25	10	10	10

Note. *M* = mean, *CI* = confidence interval, *Q* = quartile, *Min* = minimum, *Med* = median, *Max* = maximum, *SD* = standard deviation, *SS* = Story Structure, *SC* = Structural Complexity, *IST* = Internal State Terms, *C* = Story Comprehension. Confidence intervals for the means and standard deviations were estimated with bootstrapping using the R *boot* package (Canty & Ripley, 2024). This non-parametric approach resamples the data with replacement to simulate the distribution of each descriptive statistic (Hinkley & Davison, 1997). For each measure, 5,000 samples were generated (*R* = 1000). The BCa (bias-corrected and accelerated) estimation method was used.

**Table 5.** Results for the telling task, the *Cat* story.

Age	Score	M	M 95% CI	SD	SD 95% CI	Min	Q1	Med	Q3	Max
5–6 (n=11)	SS	5.36	4.45 – 6.27	1.69	1.21 – 2.23	3	4	5	6.5	8
	SC	0.89	–	0.33	–	0	1	1	1	1
	IST	2.00	1.18 – 2.55	1.18	0.69 – 1.68	0	2	2	2.5	4
	C	8.09	6.73 – 8.91	1.92	1.27 – 2.82	4	7	9	9.5	10
7 (n=13)	SS	6.31	5.38 – 6.92	1.44	0.95 – 2.19	3	6	6	7	8
	SC	0.85	–	0.38	–	0	1	1	1	1
	IST	1.38	0.62 – 2.08	1.39	1.24 – 1.56	0	0	2	3	3
	C	8.85	7.62 – 9.46	1.63	1.04 – 2.36	5	8	10	10	10
8 (n=9)	SS	6.22	5.33 – 7.11	1.39	0.78 – 2.19	4	6	6	7	9
	SC	1.11	–	0.33	–	1	1	1	1	2
	IST	1.78	0.78 – 2.33	1.20	0.73 – 1.51	0	1	2	3	3
	C	9.00	7.56 – 9.56	1.32	0.53 – 2.03	6	9	9	10	10
9 (n=14)	SS	8.23	7.23 – 9.15	1.92	1.39 – 2.77	5	7	8	10	12
	SC	1.08	0.69 – 1.31	0.64	0.41 – 0.86	0	1	1	1	2
	IST	2.92	1.75 – 3.75	1.83	1.27 – 2.50	0	2.5	3	4	6
	C	9.57	8.65 – 9.86	0.85	0.36 – 1.44	7	9.25	10	10	10
10 (n=11)	SS	8.67	6.33 – 11.00	3.77	2.35 – 6.15	2	8	8	10	16
	SC	1.11	–	0.33	–	1	1	1	1	2
	IST	1.33	0.33 – 3.50	1.97	0.52 – 2.74	0	0	0.5	1.75	5
	C	9.10	6.30 – 9.80	2.18	0.32 – 3.54	3	9.25	10	10	10

Note. *M* = mean, *CI* = confidence interval, *Q* = quartile, *Min* = minimum, *Med* = median, *Max* = maximum, *SD* = standard deviation, *SS* = Story Structure, *SC* = Structural Complexity, *IST* = Internal State Terms, *C* = Story Comprehension. Confidence intervals for the means and standard deviations were estimated with bootstrapping using the R *boot* package (Canty & Ripley, 2024). This non-parametric approach resamples the data with replacement to simulate the distribution of each descriptive statistic (Hinkley & Davison, 1997). For each measure, 5,000 samples were generated (*R* = 1000). The BCa (bias-corrected and accelerated) estimation method was used.

To determine if scores differed significantly between the age groups, we conducted a one-way analysis of variance (ANOVA) for each score of the two different tasks grouped by age (5–6, 7, 8, 9, and 10 years old). The results of these analyses are provided in Table 6. The story structure scores in both tasks (*Baby Goat* story retelling and *Cat* story telling) were, overall, significantly different between age groups. The structural complexity and internal state terms scores were not, overall, significantly different across age groups in either of the tasks. The comprehension scores only showed a significant effect of age in the retelling task.

**Table 6.** ANOVA results for each MAIN score by age group age (5–6, 7, 8, 9, and 10 years old).

Task	Score	<i>F</i>	<i>p</i>
Retelling ( <i>Baby Goats</i> )	SS	3.46	.01*
	SC	1.37	.26
	IST	1.99	.11
	C	3.50	.01*
Telling ( <i>Cat</i> )	SS	4.71	<.01*
	SC	0.93	.46
	IST	1.97	.12
	C	1.34	.27

Note. \* indicates statistically significant differences between age groups (*p* <.05). *SS* = Story Structure, *SC* = Structural Complexity, *IST* = Internal State Terms, *C* = Story Comprehension.

## 4 Discussion and conclusion

The primary aim of this project was to provide the speech-language therapists at IKL with reference data for a language assessment tool to monitor language abilities over time and identify students who may need more support with language. Given the preliminary nature and community-based scope of this study, we recommend that these data be interpreted in conjunction with the range of potential values indicated by their confidence intervals, as illustrated in the case study provided below in Section 4.2. That is, a child's raw test score could be interpreted as representing a possible range where their performance lies above or below the mean, and these data should be considered along with available non-standardized or criterion-referenced assessment techniques, including parent and teacher report (e.g., Alberta Language and Development Questionnaire; Paradis et al., 2010), language-general processing measures, such as nonword repetition tasks (Chiat, 2015; Ortiz, 2021), and language sample analysis (Ramos et al., 2022) to provide converging evidence of language disorder (Castilla-Earls et al., 2020). Analyses of variance between age groups (5–6, 7, 8, 9, and 10 years old) revealed overall significant differences in story structure scores across the age groups (5–6, 7, 8, 9, and 10 years old) for both the *Baby Goats* (story retell) task and the *Cat* (story telling) task. This suggests that the story structure scores in both story tasks may provide greater ability to differentiate performance across age groups in this sample when compared to the other three scores (structural complexity, internal state terms and story comprehension). Additionally, there was also a significant effect of age group on the comprehension scores from the *Baby Goats* retelling task. Taken together, these results indicate that the story structure scores from either story task are the most reliable indicators of different developmental levels of narrative ability in this sample, followed by the comprehension score from the *Baby Goats* story retelling task.

These results suggest that the Albanian adaptation of the MAIN assessed skills that improve as language and academic skills develop and may be useful to monitor progress over time in this population of children in Pristina, Kosovo. Additionally, the greater range of the story structure scores, and the stability of means and standard deviations across age groups that increased with age highlights this score as a diagnostically useful measure of language ability that is likely to capture developmental changes over time.

### 4.1 Comparison to previous research

This study contributes to a growing body of work that has adapted and localized the MAIN for different languages and populations (for an overview, see Lindgren et al., 2023). As previously mentioned, age effects have been found in performance on comprehension and production between ages 4;4 and 5;10 (Lindgren, 2019). In our study, significant age effects were found in comprehension performance for the retelling task, but not for the telling task. Significant age effects were also found in production for both telling and retelling (Story Structure score). Our findings align with previous studies in demonstrating age-related differences in narrative performance, though we did not directly compare specific age groups as in some earlier studies. Notably, we found a significant effect of age group on comprehension in the *Baby Goats* condition, but not in the *Cat* condition. This pattern is consistent with prior research indicating that contrasts like *Baby Goats*/*Baby Birds* tend to yield lower comprehension scores compared

to Cat/Dog, suggesting that the observed age effects may be driven in part by the stimulus material used (Lindgren, 2018).

4.2 Case Study: A practical interpretation of results

To illustrate how the data from the present study could be used in a clinical context, we will analyze one participant, which we refer to as Child A. In the Home Language Questionnaire (see Section 4.2), Child A’s parents reported concerns about her language, writing, “I think that sometimes she can’t express herself very well in Albanian, but she thinks in English and then translates it in her brain.”

From the report given by the parents of this child, it is not clear whether the parents were referring to typical characteristics of multilingual language development, such as code-switching (Miccio et al., 2009) or concerns that may be indicative of language delay or impairment (Paradis et al., 2010). It is possible that this report may not indicate difficulty with expressive language. For discussion, we review the story structure and comprehension scores of Child A in relation to means and standard deviations for their respective age group, as shown in Table 4 and Table 5.

Child A was 9 years, 2 months old. Table 7 provides her story structure and story comprehension scores adjacent to the corresponding statistics for scores in her age group. For story structure, Child A’s narrative retelling score was 6 and her telling score was 7 (out of a maximum of 17). Given the 95% confidence interval around the mean and standard deviation for the 9-year-old age group, Child A’s story structure score in retelling would likely fall between 0.88 and 1.14 standard deviations below the mean for her age group. Her story structure score in telling would likely fall somewhere between 0.15 and 0.93 standard deviations below the mean for her age group. For comprehension, Child A scored 10 points (the maximum score) on both the narrative retelling and telling tasks.

**Table 7.** Story Structure (SS) and Story Comprehension (C) scores for Child A with group score statistics for her age group (Age 9).

Measure	Task	Score	<i>M</i>	<i>M</i> 95% CI
SS	Telling	7	8.46	7.23 – 9.38
	Retelling	6	8.23	7.23 – 9.15
C	Telling	10	9.57	8.65 – 9.86
	Retelling	10	9.93	-

Both of Child A’s story structure scores were below the mean for the 9-year-old age group. Using the most conservative estimate, her scores would be 0.88 and 0.15 standard deviations below the mean on retelling and telling, respectively. Using  $\geq 1$  standard deviation below the mean as a typical cutoff, this would be considered on the lower end of the normal range but would not alone indicate the need for clinical intervention from a speech-language pathologist. At the opposite end of this possible range, her scores would be 1.14 and 0.93 standard deviations below the mean on retelling and telling, respectively. Depending on the paradigm applied, these could fall at or near a clinically significant level of low performance. In addition, Child A demonstrated strong receptive language abilities with maximum scores on both comprehension

tasks. Given the range of possible interpretations, follow-up with this child, including additional assessment measures to acquire a complete diagnostic view of expressive language abilities, would be warranted.

In sum, these reference data for the Albanian MAIN provide context for expected narrative language abilities in school-aged children living in Pristina, Kosovo. These data are most appropriately interpreted within this local context. Based on the current data, story structure scores may be the most informative scores for differentiating performance between age groups in this population. If individual performance is to be interpreted in the context of age group score statistics, the entire range of the confidence interval for a given reference value should be considered. Clinicians are encouraged to weigh clinical judgement, parent report, teacher report, and other available measures, along with these reference data, when evaluating language abilities.

### **4.3 Conclusion**

More research must be conducted to understand school-age language development in Kosovo. The body of research on this topic is severely limited, which negatively affects students with language disorders who may be difficult to identify and consequently unlikely to receive support. Furthermore, a gold standard tool or set of tools for identification of developmental language disorder among children in Kosovo is not yet available, and only limited parental report data was collected in this study. Consequently, we highlight that the reference data presented here for the Albanian adaptation of the MAIN should be used for screening and/or in combination with additional language measures. Note that comparison to any group reference values in these data should consider the reported confidence interval for that value.

In this study, we collected reference data for the Albanian adaptation of the MAIN to aid in diagnostic decision-making and progress monitoring for speech-language therapists in Pristina, the capital city of Kosovo. Presently, no other comparable language assessment tool for school-age children in Kosovo exists. Across tasks, the story structure score was found to be the most reliable for differentiating narrative language performance across age groups (5–6 years, 7 years, 8 years, 9 years, 10 years). Comprehension scores on the *Goat* retell task also increased significantly with age. Together, these indicate a degree of construct validity and potential for these scores to be used to monitor progress over time and differentiate levels of language ability in this population. Completion of this community-based collaboration between the University of Iowa and *Instituti Kosovar për Logopedi* (IKL) and IKL's expanding provision of speech and language services in Pristina since 2020 highlight the recent growth of speech-language therapy in Kosovo and the urgent need for accessible and evidence-based speech and language assessment and intervention resources for the diverse communities within this country.

### **Acknowledgements**

The authors wish to acknowledge the contributions of the participating schools that provided space for this study and assisted in recruitment of participants. The authors also express their gratitude to the parents and children who participated in the study. Additionally, the authors

thank Saloni Upadhyay and Adrian Bradley for their support. Finally, the authors express a very special thank you to Dr. Shanley Allen of University of Kaiserslautern-Landau and Dr. Enkeleida Kapia of Ludwig Maximilian University of Munich, without whom this project would not have been possible.

## References

- Bohnacker, U. (2016). Tell me a story in English or Swedish: Narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics*, 37(1), 19–48. <https://doi.org/10.1017/S0142716415000405>
- Bohnacker, U., & Gagarina, N. (2019). Background on MAIN – Revised, how to use it and adapt it to other languages. *ZAS Papers in Linguistics*, 63, 4–12. <https://doi.org/10.21248/zaspil.63.2019.516>
- Canty, A., & Ripley, B. D. (2024). *boot: Bootstrap R (s-plus) functions* [Manual].
- Castilla-Earls, A., Bedore, L., Rojas, R., Fabiano-Smith, L., Pruitt-Lord, S., Restrepo, M. A., & Peña, E. (2020). Beyond scores: Using converging evidence to determine speech and language services eligibility for dual language learners. *American Journal of Speech-Language Pathology*, 29(3), 1116–1132. [https://doi.org/10.1044/2020\\_AJSLP-19-00179](https://doi.org/10.1044/2020_AJSLP-19-00179)
- Chiat, S. (2015). Non-word repetition. In S. Armon-Lotem, J. de Jong, & N. Meir (Eds.), *Methods for Assessing Multilingual Children: Disentangling Bilingualism from Language Impairment* (pp. 125–150). Multilingual Matters.
- CIA, Central Intelligence Agency. (n.d.). *Kosovo*. Central Intelligence Agency. Retrieved January 10, 2024 from <https://www.cia.gov/the-world-factbook/countries/kosovo/>
- Combiths, P. (2023). Home Language Questionnaire. Clinical Linguistics and Disparities Lab, University of Iowa. <https://cld.lab.uiowa.edu/resources/resources-slp/home-language-questionnaire>
- Fox, J., & Weisberg, S. (2019). *An R Companion to Applied Regression* (3<sup>rd</sup> ed.). Sage.
- Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U. & Walters, J. (2019). MAIN: Multilingual Assessment Instrument for Narratives – Revised. *ZAS Papers in Linguistics*, 63, 1–36. <https://doi.org/10.21248/zaspil.63.2019.516>
- Hanemann, U. (2006). Literacy in conflict situations. *UNESCO Education for All Global Monitoring Report*.
- Hinkley, D. V., & Davison, A. C. (1997). *Bootstrap methods and their applications*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511802843>
- Hržica, G., & Kuvač Kraljević, J. (2020). The Croatian adaptation of the Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 64, 37–44. <https://doi.org/10.21248/zaspil.64.2020.555>
- Hyter, Y. D., & Salas-Provance, M. B. (2021). *Culturally Responsive Practices in Speech, Language, and Hearing Sciences, Second Edition*. Plural Publishing.
- Israel, B. A., Schulz, A. J., Parker, E. A., & Becker, A. B. (1998). Review of community-based research: Assessing partnership approaches to improve public health. *Annual Review of Public Health*, 19(1), 173–202. <https://doi.org/10.1146/annurev.publhealth.19.1.173>
- Kamusella, T. (2016). The idea of a Kosovan language in Yugoslavia's language politics. *International Journal of the Sociology of Language*, 242, 217–237. <https://doi.org/10.1515/ijsl-2016-0040>
- Lindgren, J. (2018). *Developing narrative competence: Swedish, Swedish-German and Swedish-Turkish children aged 4-6*. Acta Universitatis Upsaliensis.

- Lindgren, J. (2019). Comprehension and production of narrative macrostructure in Swedish: A longitudinal study from age 4 to 7. *First Language*, 39(4), 412–432. <https://doi.org/10.1177/0142723719844089>
- Lindgren, J., Tselekidou, F., & Gagarina, N. (2023). Acquisition of narrative macrostructure: A comprehensive overview of results from the Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics*, 65, 111–132. <https://doi.org/10.21248/zaspil.65.2023.623>
- Lohaj, F. (2018). English language as a communication bridge in post war Kosova. *Universal Journal of Educational Research*, 6(8), 1675–1683. <https://doi.org/10.13189/UJER.2018.060807>
- Miccio, A. W., Hammer, C. S., & Rodríguez, B. (2009). Code-switching and language disorders in bilingual children. In A. J. Toribio & B. E. Bullock (Eds.), *The Cambridge Handbook of Linguistic Code-switching* (pp. 241–252). Cambridge University Press.
- Mëniku, L., & Campos, H. (2011). *Discovering Albanian I Textbook*. The University of Wisconsin Press.
- Nešić Blagoje V., Minić Jelena Lj., Jakšić Aleksandra S. (2011). Prevalence of speech and language disorders in children in northern Kosovo and Metohija. *Zbornik Radova Filozofskog Fakulteta u Prištini*, 41, 587–601.
- Ortiz, J. A. (2021). Using nonword repetition to identify language impairment in bilingual children: A meta-analysis of diagnostic accuracy. *American Journal of Speech-Language Pathology*, 30(5), 2275–2295. [https://doi.org/10.1044/2021\\_AJSLP-20-00237](https://doi.org/10.1044/2021_AJSLP-20-00237)
- Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: Using parent report on first language development. *Journal of Communication Disorders*, 43(6), 474–497. <https://doi.org/10.1016/j.jcomdis.2010.01.002>
- Quell, C. (2017). *Official Languages in Kosovo: A Canadian contribution to the implementation of language rights in Kosovo*. University of Ottawa.
- Ramos, M. N., Collins, P., & Peña, E. D. (2022). Sharpening our tools: A systematic review to identify diagnostically accurate language sample measures. *Journal of Speech, Language, and Hearing Research*, 65(10), 3890–3907. [https://doi.org/10.1044/2022\\_JSLHR-22-00121](https://doi.org/10.1044/2022_JSLHR-22-00121)
- Roch, M., & Hržica, G. (2020). Narrative comprehension by Croatian-Italian bilingual children 5–7 years old: The role of receptive vocabulary and sentence comprehension. In U. Bohnacker & N. Gagarina (Eds.), *Developing Narrative Comprehension: Multilingual Assessment Instrument for Narratives* (pp. 172–196). John Benjamins.
- Regis College. (2023, June 29). *Child Illiteracy in America: Statistics, Facts, and Resources*. Regis College. <https://online.regiscollege.edu/blog/child-illiteracy/#:~:text=Parents%20and%20caretakers%20with%20a,according%20to%20Comic%20Relief%20US>
- Shahini, B. (2016, October 14). Lessons in resistance: Kosovo's parallel education system in the 1990s. *Prishtina Insight*. <https://prishtinainsight.com/lessons-resistance-kosovos-parallel-education-system-1990s-mag/>
- Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A Grammar of Data Manipulation*. <https://dplyr.tidyverse.org>
- World Medical Association. (1964). Declaration of Helsinki. Retrieved from <https://www.wma.net/what-we-do/medical-ethics/declaration-of-helsinki/>
- UK Parliament House of Commons Library. (2024). *Kosovo: development since 1999 and relations with Serbia*. House of Commons Library. Retrieved December 26, 2024, from <https://commonslibrary.parliament.uk/research-briefings/cbp-10012/>

U.S. Department of State. (n.d.). *A Guide to the United States' History of Recognition, Diplomatic, and Consular Relations, by Country, since 1776: Kosovo*. U.S. Department of State. Retrieved December 26, 2024, from <https://history.state.gov/countries/kosovo>