Assessment of Narrative Abilities in Bilingual Children[*]

Natalia Gagarina, Daleen Klop, Sari Kunnari, Koula Tantele, Taina Välimaa, Ingrida Balčiūnienė, Ute Bohnacker, Joel Walters

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U. & Walters, J. 2015. Assessment of narrative abilities in bilingual children. In S. Armon-Lotem, J. de Jong & N. Meir (eds.) *Methods for assessing multilingual children: Disentangling bilingualism from language impairment*. Bristol, UK: Multilingual Matters.

[*]This work was supported in part by Bundesministerium für Bildung und Forschung (BMBF) (Grant Nr. 01UG0711), COST Action IS0804, Israel Science Foundation (Grant No. 770/10) and the Academy of Finland. The authors are grateful to Katrin Reichenbach for assistance in the technical preparation of this chapter.

1 Introduction and overview

The *Multilingual Assessment Instrument for Narratives (MAIN*, Gagarina et al., 2012) was developed to assess narrative production and comprehension abilities of bilingual children from 3 to 9 years (later referred to as *Language Impairment Testing in Multilingual Settings: Multilingual Assessment Instrument for Narratives, LITMUS-MAIN*).

Different types of narratives offer a platform for examining a wide range of linguistic abilities in context (cf. Hughes et al., 1997). These abilities concern story structure, discourse features (e.g. coherence and cohesion), morphosyntax, complex syntax, lexis and uniquely bilingual phenomena such as code-switching and cross-linguistic interference. Children's narratives also provide an index of their cognitive, semantic and social abilities (Liles, 1993). Narrative analysis is considered by researchers and clinicians to be an ecologically valid way to investigate communicative competence (Botting, 2002) and to be less biased against bilingual children than norm-referenced assessment tools (Paradis et al., 2011). Oral narratives provide a rich source of data about a child's language use in a relatively natural context. Finally, narrative analysis allows clinicians to assess multiple linguistic features in context—including macrostructure, e.g. story grammar categories such as goals, attempts and outcomes; microstructure features, e.g. lexical diversity, relational and referential devices—using relatively short language samples (Heilmann et al., 2010a, 2010b).

Narrative skills are important for later success in school, e.g. in literacy and for comprehension of the language of mathematics (Bishop & Edmundson, 1987; Bliss et al., 1998; McCabe, 1996; McCabe & Rollins, 1994; Walach, 2008; Westby, 1991). They form a bridge between oral language and literacy by providing exposure to and experience in using extended, contextualized, cohesive discourse units and abstract texts that children will encounter in written texts (Hadley, 1998; Westby, 2005). According to Kaderavek and Sulzby (2000) and Oakhill and Cain (2007), reading comprehension has its roots in the comprehension of narrative discourse that develops simultaneously with other early language skills prior to formal reading instruction. The ability to tell a story links oral language skills to literacy, since it requires children to plan and produce contextualised and cohesive discourse. Intervention studies have shown that directly teaching narrative skills improves comprehension and production of oral narratives as well as reading comprehension (Hayward & Schneider, 2000; Swanson et al., 2005). Moreover, narrative abilities at the macrostructure level, i.e. composition of cohesive event sequences, reflect capacities which go beyond the specifics of one language. Thus, the assessment of narratives can be seen as especially appropriate for bilingual children: "language tasks that require a cognitive component might also be less biased against dual language children, because the cognitive component could be tapping into language-general capacities" (Paradis et al., 2011:221).

History and overview. Our initial goal was to examine and evaluate different tasks used to elicit narratives and to identify specifically bilingual features in narrative discourse. The review of existing tasks and tests showed that while elicitation procedures and scoring were usually sound, they did not provide sufficiently parallel stimuli for elicitation and assessment of narrative abilities in both of the languages of a bilingual child. Existing pictorial stimuli were not grounded in the present theoretical story grammar framework, nor did they take into account internal state terms. The use of internal state terms provides important information about the narrator's awareness of characters' mental states, motivations, intentions and goals (Nippold et al., 2005). In addition, the protagonists and the lexical items denoting them were often not controlled for frequency of use and perceptual complexity.

This led to the development of new pictorial stimuli, involving more than 200 revisions of pictures and story scripts piloted prior to the design presented here. This design and its accompanying research program comprises the elicitation of narratives in two modes, i) story telling (story generation) and ii) story retelling, as well as assessment of comprehension of macrostructure components and internal state terms. The design allows for another possible mode, storytelling after listening to a model story that is similar in structure yet different in content from the story to be told. The model story is thus a variant of the retelling procedure. In addition, a set of background questions (based on Gagarina et al., 2010) forms part of the assessment procedures to evaluate the quality and quantity of the input to the child in both of her/his languages.

Stimulus pictures and scripts for retelling include a combination of macrostructure, microstructure and internal state features of narrative discourse in order to investigate each child's performance cross-linguistically in a within-subject design. The materials were developed in multiple languages in order to make it possible to screen and identify children at risk for Specific Language Impairment (SLI) in both their languages. The present version of the tool (available online at http://www.zas.gwz-berlin.de/zaspil56.html; Gagarina et al., 2012) has been piloted in 17 languages and 14 language pairs with more than 500 monolingual and bilingual children.

After this general introduction, the remainder of the chapter is organized as follows: Section 2 provides the conceptual framework and a literature overview, emphasizing the distinction between macrostructure and microstructure and providing information about internal state language and bilingual aspects of narrative abilities. Section 3 gives a detailed presentation of the telling and retelling elicitation tasks (3.1), stimulus pictures (3.2), stimulus scripts (3.3), the narrative comprehension task (3.4), and how the instrument is administered (3.5). Section 4 deals with transcription, coding and scoring of the data. Section 5 gives some preliminary findings from the studies which have been conducted thus far, and Section 6 concludes the chapter.

2 Conceptual framework and literature

Methods for collecting and analyzing narrative speech samples are highly varied. There are different sampling procedures (e.g. spontaneous or elicited), different types of narratives (e.g. scripted, personal and fictional stories), and different elicitation methods (e.g. story generation/telling and story retelling). One characteristic feature of narratives is that they contain information at two levels: macrostructure and microstructure. Macrostructure analysis focuses on higher-order hierarchical organization, including episodic structure and story grammar components (Heilmann et al., 2010a) and can be said to be language-independent (Pearson 2002). Microstructure focuses on the linguistic structures used in the construction of coherent discourse, inter alia, number, length and complexity of utterances or communication units, lexical diversity, use of referential noun phrases, connectives, tense, etc. Macrostructure and microstructure abilities represent two distinct but interrelated areas underlying narrative discourse competence (Liles et al., 1995). These abilities are not often examined in a single framework. Rather, most narrative tests focus on language-specific capacities and microstructure features such as vocabulary and/or grammar. *LITMUS-MAIN* examines narrative production of microstructure and macrostructure elements, in combination with the comprehension of macrostructure features. It also includes internal state terms, thereby providing information about skills at the cognitive-linguistic interface.

Various theoretical approaches (e.g. Labov & Waletzky, 1967; Mandler & Johnson, 1977; Stein & Glenn, 1979; Bruner, 1986; Westby, 2005) and elicitation options were examined prior to pilot testing different measures. Narrative assessment calls for a wide-scoped, integrative framework, which includes macrostructure and microstructure as well as production and comprehension. The breadth as well as the particular emphasis on internal states and bilingual features makes *LITMUS-MAIN* unique in scope as well as focus. The components of the instrument can be found in section 4.1 below.

2.1 Macrostructure

The story grammar model (see e.g. Mandler & Johnson, 1977; Stein & Glenn, 1979), according to which all stories have a setting and episode structure claimed to capture a universal organizational pattern for story knowledge (Trabasso & Nickels, 1992), served as our initial theoretical framework. Story grammar research has been conducted with a wide variety of data collection procedures and a wide variety of populations, including bilingual children (Fiestas & Peña, 2004; Gutiérrez-Clellen et al., 2008; Pearson, 2001, 2002; Pearson & de Villiers, 2005; Uccelli & Paéz, 2007) as well as bilingual children with language impairment (e.g. Cleave et al., 2010; Gutiérrez-Clellen et al., 2009; Iluz-Cohen & Walters, 2012; Simon-Cereijido & Gutiérrez-Clellen, 2009).

In the instrument described below, a slight adaptation of the story grammar model is used. The stories begin with a setting statement, which specifies time and place and introduces the protagonist. This component is followed by three episodes. Each episode consists of i) a goal (G) statement for the protagonist, ii) an attempt (A) by the protagonist to reach the goal, iii) an outcome (O) of the attempt in terms of the goal, and iv) internal states which initiate the goal and express reactions. The scripts for the retelling version of each story (see Table 1 in section 3.3) are color-coded to indicate goals, attempts and outcomes as well as internal state terms.

Structural complexity. Analysis of structural complexity provides information about the child's level of narrative development and allows comparison across languages. The approach taken here is grounded in clinical assessment and based on Westby's binary decision tree (Westby, 2005), illustrated in Figure 1, where episodes within the stories are classified into one of three levels of structural complexity: i) sequences, where no goal statement has been generated, ii) incomplete episodes, which include a goal statement, but lack a complete GAO (goal-attempt-outcome) structure due to omission of an attempt or outcome, and iii) complete episodes, which include all three GAO components. The ability to produce well-formed episodes in narratives indicates understanding of narrative schemata, causality, perspective-taking, meta-awareness of the ability to plan, and the need to justify plans and actions (Trabasso & Nickels, 1992; Trabasso & Rodkin, 1994). Additionally, the number of isolated goals (G) is noted in order to provide a more fine-tuned differentiation between the various populations involved.

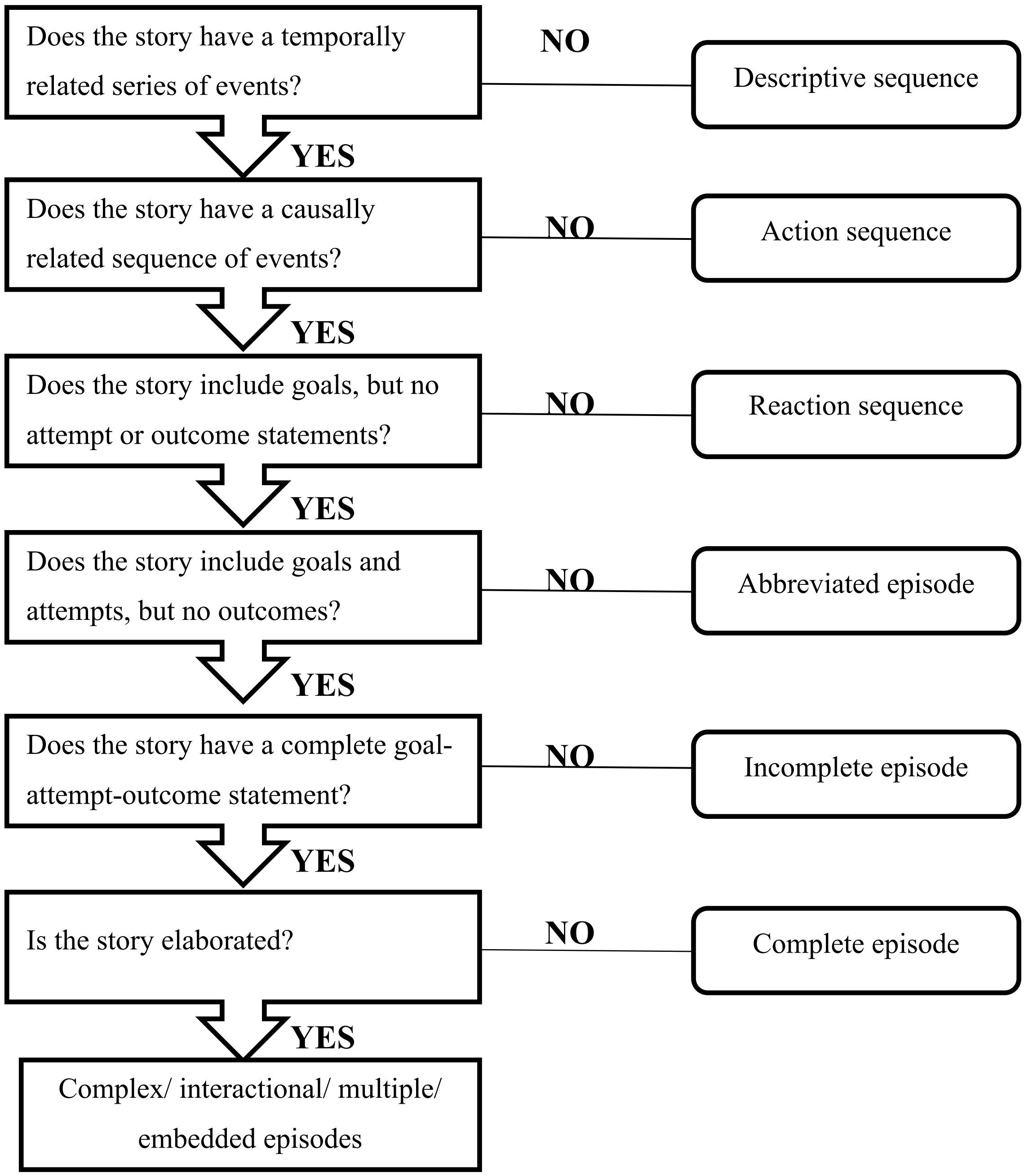


Figure 1. Structural complexity of children's narratives: Decision Tree (adapted from Westby, 2005).

Internal state terms usage. Cohesive and coherent narratives presuppose awareness of others' states of mind on different levels. Story understanding involves interpreting emotions, goals and intentions of protagonists. In addition, story tellers must provide such information for the listener so that s/he can follow and understand a narrative. Mental and internal states attributed to the self and others have been studied in the context of theory of mind (ToM) skills, such as intention-reading, perspective-taking, and repair strategies in instances of communicative breakdown (Berman & Slobin, 1994; Lorusso et al., 2007; Tomasello, 2003).

Analysis of internal state language in children's narratives can provide information about their metalinguistic and metacognitive knowledge as well as ToM abilities. The use of metalinguistic verbs (referring to acts of speaking, such as *shout*, *say*), metacognitive verbs (referring to acts of thinking, such as *think*, *wonder*) and words expressing emotion (e.g. *sad, angry*) can be taken as evidence of awareness of others' states of mind and as indications of cognitive processes required to interpret intentionality and of the ability to make inferences about aspects of stories (Nippold et al., 2005; Westby, 2005). The use of internal state language in narratives is associated with a literate style that forms a crucial aspect of school-based discourse (Curenton & Justice, 2004; Grazzani & Ornaghi, 2012) and the development of complex syntax (Heilmann et al., 2010a; Nippold et al., 2005).

Taxonomies for investigating internal state terms focus primarily on mental state verbs, including motivational verbs (e.g. *want, need*), experiential expressions (e.g. *see, surprised, thirsty*), belief verbs (e.g. *think, know*), verbs of saying and telling (e.g. *say, call, shout*) and emotion terms (e.g. *sad, happy, angry*) (Grazzani & Ornaghi, 2012; Greenhalgh & Strong, 2001). Internal state language is assessed by *LITMUS-MAIN* at the macrostructure level in production as well as in comprehension, and it is interpreted as a marker of children's understanding and awareness of intentionality and goal-directed behavior of protagonists. Since parallel lexical items were selected for all languages, internal state terms should be comparable across languages. All instances of internal state terms (perceptual state terms: *see, hear*; physiological state terms: *thirsty, hungry*; consciousness terms: *alive, awake*; emotion terms: *sad, happy, angry, worried, disappointed*; mental verbs: *want, think, know, forget, decide*; verbs of saying and telling (linguistic verbs): *say, call, shout*; etc.) are calculated in the production section of the scoring form (illustrated in Appendix C) and further analyzed in the comprehension section (also illustrated in Appendix C) in response to inference questions, e.g. *How do you think the cat feels?*

Comprehension. Assessing the comprehension of the main components of story structure is an important complement to story-telling and/or retelling, because some typically developing (TD) bilingual children and bilingual children with SLI might show similar profiles in production, whilst differing in story comprehension. Comprehension is elicited by means of questions which are asked after the production part of the assessment procedure. Ten questions were developed: three of them target the three goals (with the Goal being the main macrostructure component), six target the internal state terms as goals, initiating events, and reactions. The tenth question (*Who does the mother goat like best ... and why?*), as well as some of the earlier questions (e.g. *Why do you think he is feeling …?*) tap Theory of Mind, as they require the child to draw inferences from the pictures, and describe how s/he has come to a certain conclusion.

2.2 Microstructure

Microstructure elements cover a wide range of linguistic aspects including general measures of length and lexis, aspects of morphosyntax and discourse domain, and bilingual phenomena. Microstructure elements are language-specific, and it is inevitable that some will differ across languages. In order to select the most relevant characteristics which might be diagnostic for (bilingual) children with SLI, we reviewed recent studies which made use of narratives elicited with picture stimuli.

It is well known that the narratives of monolingual children with SLI differ from those with typical language development in the area of morphosyntax (e.g. Reilly et al., 2004), verbosity/story length (Strong & Shaver, 1991; Schneider et al., 2006), topic maintenance, event sequencing, informativeness (Lucas, 1980; Roth & Spekman, 1986; Olley, 1989; Biddle et al., 1996), and referencing of events and individuals. Syntactic complexity is another vulnerable area in SLI as suggested by various studies (Kit-Sum To et al., 2010; Liles et al., 1995). Narratives of children with and without SLI have also been found to differ concerning several productivity and complexity measures (Gillam & Johnston, 1992; Newman & McGregor, 2006). These include story length as measured in words, mean length of communication unit (c-unit, where each independent clause with its modifiers forms one c-unit (Loban, 1976)) and number of minimal terminable units (t-units, where each main clause and all the subordinate clauses attached to it form one t-unit (Hunt 1965)), percentage of complex t-units per narrative, frequency of grammatically well-formed t-units, and frequency of subordinate clauses (Gillam & Johnston, 1992).

Studies of microstructure features in the narratives of TD and SLI bilingual children which examine data in both of a child’s developing languages are still relatively scant, as usually only the child’s second language (L2) is studied. There are difficulties in generalizing microstructural markers of SLI across languages. Morpho-syntactic symptoms of SLI are discussed in other chapters of this volume, and while they can be investigated on the basis of narrative data elicited with *LITMUS-MAIN*, they will not be focused on the present chapter, given their inherently language-specific nature. Therefore, a basic set of microstructure features to be analyzed in *LITMUS-MAIN* (including productivity, variety and fluency) has been developed (see Section 4.2) giving the researchers the possibility to include analyses essential for specific research interests.

3 *LITMUS-MAIN* Tasks and Stimulus Materials

3.1 Elicitation Tasks: Telling (story generation) and retelling

Two narrative elicitation tasks, telling and retelling, are used to assess children's narrative abilities along a continuum of complexity. The 'telling' format is presumed to be more difficult, since the child is required to generate his/her own story without the benefit of a prior script. Nevertheless, telling may offer the child more freedom to use his/her imagination and thus may better reflect the child's lexis. Telling formats may therefore provide more information about children's independent narrative formulation abilities than retelling (Schneider et al., 2006). Telling also avoids a potential confound with retelling, as retelling performance is intertwined with story recall, verbal memory and attention (Boudreau, 2007; Dodwell & Bavin, 2008; Graybeal, 1981; Gutiérrez-Clellen, 2002).

Previous studies (Hayward et al, 2007; Liles, 1993; Schneider & Dubé, 2005) have found that children with and without language impairment often provide longer, more detailed, and grammatically more accurate language samples during retelling. This was, however, not confirmed by our preliminary findings, since retelling on MAIN did not result in more detailed or complex narratives than telling. Retelling involves reconstruction and reinterpretation of the story, and is more than just a repetition of the stimulus narrative. Retelling thus provides information about how children modify and assimilate a story's vocabulary and grammatical structures, as well as the content of the story (Gillam & Carlisle, 1997). In addition, retelling offers the researcher control over certain aspects of the narrative, e.g. length, complexity and content, and allows for error analyses and assessment of comprehension by providing a model story (Hadley, 1998; Liles, 1993). However, similar control can also be achieved for telling formats with the help of carefully constructed stimulus materials, e.g. when stories based on parallel picture sequences are elicited and their comprehension is assessed, as the content of the target narrative is then known to the researcher beforehand.

LITMUS-MAIN includes both telling and retelling in two languages, which generates four narratives for each child and allows within-subject, cross-language comparisons. However, several pilot studies (e.g. Klop et al., 2011, 2012) showed carry-over effects from one task to another when children performed both telling and retelling tasks in both languages. In an attempt to minimize carry-over, a third elicitation option, a model story, was introduced. The model story option provides more contextual support than the telling task and a base for asking comprehension questions.

3.2 Development of stimulus pictures

Initially, major narrative assessment instruments which employ picture stimuli were reviewed. These included the *Renfrew Bus Story* (Glasgow & Cowley, 1994), the *Test of Narrative Language* (TNL; Gillam & Pearson, 2004), the *Edmonton Narrative Norms Instrument* (ENNI; Schneider & Hayward 2005), and *Hamburger Verfahren zur Analyse des Sprachstands Fünfjähriger* (Hamburg Procedure for the Analysis of Language Proficiency in Five-year-olds, HAVAS 5; Reich & Roth, 2004) and also the frequently used picture book *Frog where are you?* (Mayer 1969). The pictorial content of these tools may not be appropriate for children from diverse cultural, linguistic and socio-economic backgrounds. Further, none of the aforementioned tools provide enough parallel pictorial stimuli to assess both languages of a bilingual child.

To overcome these shortcomings of existing assessment instruments, four six-picture sequences (see Figures 2 to 5) were developed based on 3 to 9-year-old children's linguistic and cognitive (working memory, attention) abilities. Four separate sequences were needed due to the bilingual and task requirements of the instrument, which implemented a 2x2 factorial design involving Language (L1/L2) and Task (Telling: Baby Birds story & Baby Goats story / Retelling or Model Story: Cat story & Dog story).

The pictorial sequences were designed to portray clearly depicted actions. Working closely with a professional illustrator, each episode was scripted, and careful attention was paid to the protagonists' intentions, emotions and actions, to their relative size, and to the characters' facial expressions. In order to achieve comparability across narratives, we aimed for congruence across the set of scripts as well as between scripts and pictorial content by creating parallel storylines for the different sets of pictures. Details attended to included: (a) protagonists: the number of protagonists, the timing of the introduction of new protagonists, their relative position in the pictures and interaction, their size in relation to other objects, and the angle from which they were looking at the other protagonists; (b) background and foreground: these had to be contrasted and closely connected to and motivated by the actions of the main protagonists, with similar cognitive complexity and visual representation density across pictures and stories; (c) content: comparable onset, development and conclusion of the story-line. In general, the aim was to achieve parallelism of story structure across all four stories.

Content development was based on the components of story grammar, namely initiating events, character's goals, attempts to reach the goal, outcomes of the attempts and reactions following the outcomes. Instructions to the illustrator ensured that these components were explicitly portrayed in the pictures. For example, to portray the goal of the cat in the Baby Birds story, the cat's facial expression, gaze direction and movements towards the birds convey its intention to jump, while the baby birds' facial expression and gaze and the mother's protective stance portray anxiety (see Figure 2). The distance between the characters was designed to imply time for the protagonist's reactions. Special emphasis was put on clearly depicting the emotions of the protagonists in order to justify the use of internal state terms, e.g. the baby birds' beaks are opened and their gaze is directed towards the mother to show that they are hungry.

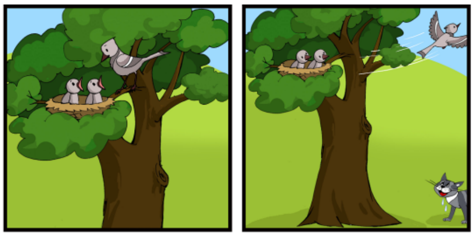
The content of each picture sequence was designed to portray three short episodes. The rationale for portraying three episodes in each picture sequence was to provide more than one opportunity for a child to produce each story structure element targeted for macrostructure analysis. In terms of story grammar (e.g. Stein & Glenn, 1979), this affords the child three opportunities to produce initiating events, goals, attempts, outcomes and reactions.

The four picture sequences – *Baby Birds* and *Baby Goats; Cat* and *Dog* – are all matched for the number of main protagonists and GAO sequences; additionally, each pair of picture sequences (Baby Birds/Baby Goats and Cat/Dog) is parallel in the structure of the plot, internal state terms and general actions performed by the main protagonists (see Figures 2 to 5).

In short, this story design has advantages over longer and more elaborate narrative elicitation methods in that it is carefully structured, allowing identification of the category that has been generated or retold by the child. It also has advantages over shorter narratives where only a single episode is presented.

Cultural and age appropriateness. The findings of pilot studies from 20 countries attest to the cultural robustness of the materials for eliciting veridical narrative data. These studies informed the process of continual revision over a period of two years to arrive at the final stimulus pictures in Figures 2 to 5. As the instrument targets children from 3 to 9 years, protagonists were selected that would be familiar to children from different ages across a wide variety of cultures. Birds, goats, cats and dogs are frequent in children's and child-directed speech and in fairy-tales, and pre-testing showed them to be easily identifiable by a wide range of children.

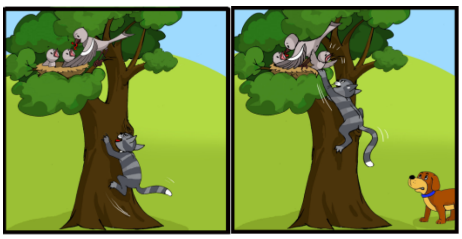
The number of protagonists per picture and the timing/sequence of their appearance in each story were also controlled for. For example, in the first picture in each of the four stories, only the main protagonist is presented (in Baby Birds/Baby Goats the two babies are presented additionally); in all stories the second protagonist appears only in picture 2, where he/she/it is seen only partially in order to convey a process of "entering" or first appearance. Furthermore, plurality was controlled across the set of stories, e.g. in the Baby Birds and Baby Goats stories, there are two baby birds and two baby goats. Finally, background details were kept to a minimum in order not to distract the child from the primary content and structure. (A detailed description of the materials development and the rationale behind can be found in Gagarina et al., 2012: Chapter 2.) In summary, the pictorial content was controlled for macrostructure, characters and their actions and feelings as well as cultural appropriateness and robustness.



Episode 1: Bird Goal – to feed the baby birds

Episode 1: *Bird* Attempt – flies away to get food

Episode 2: Cat Goal – to get the baby bird(s)

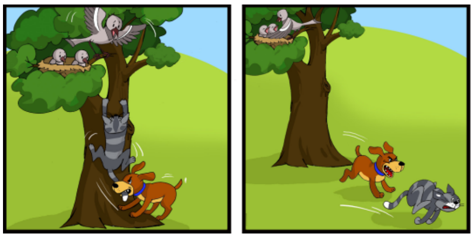


Episode 1: Bird Outcome – returns with food

Episode 2: Cat Attempt – climbs the tree to catch a baby bird

Episode 2: Cat Outcome – catches the baby bird

Episode 3: Dog Goal – to save the baby bird(s)



Episode 3: Dog Attempt – bites, pulls the cat's tail

Episode 3:Dog Outcome – chases the cat away / the cat runs off

Figure 2. Baby Birds stimulus pictures (based on Hickmann, 2003)

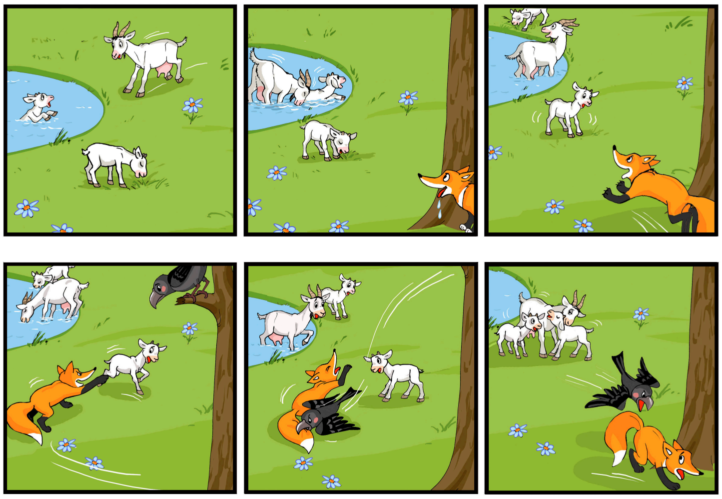


Figure 3. Baby Goats stimulus pictures (based on Gülzow & Gagarina, 2007)

In order to implement the 2x2 (language by task) design, two additional sets of pictures were generated. These also maintain the 3-episode (goal (G)–attempt (A)–outcome (O)) and internal state macrostructure of the Baby Birds and Baby Goats stories, but differ slightly in complexity (see scripts for Cat and Dog story in Table 1 below). Their pictorial stimuli (see Figures 4 and 5) were designed for the retelling task.

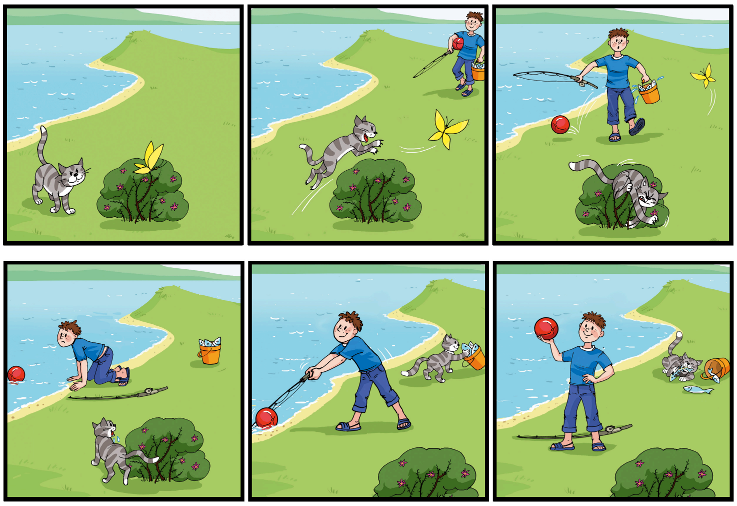


Figure 4. Cat stimulus pictures (story retelling/model story)

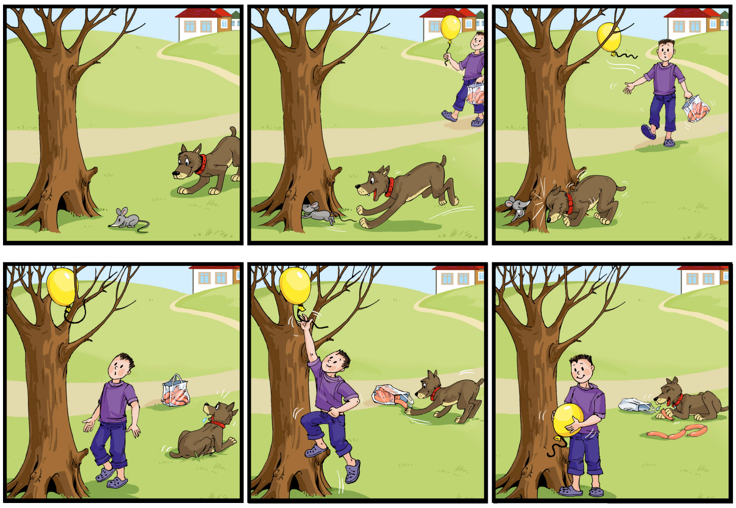


Figure 5. Dog stimulus pictures (story retelling/model story)

3.3 Stimulus scripts for retelling task and comparability across languages

The stimulus scripts were designed to be comparable in terms of both macrostructure and microstructure information. Table 1 displays the English stimulus scripts for the two stories (Cat/Dog) designed for retelling or serving as a model story, color-coded for goals, attempts, outcomes and internal state terms. Cross-linguistic syntactic and lexicalization differences made it impossible to maintain all the microstructure features across languages.

Table 1. English stimulus scripts for the Cat and Dog stories. Color-coded for goal, attempt, outcome, and *internal state terms*.

Cat

Pictures 1/2:

One day there was a *playful* cat who *saw* a yellow butterfly sitting on a bush. He leaped forward because he wanted to catch it. Meanwhile, a *cheerful* boy was coming back from fishing with a bucket and a ball in his hands. He *looked* at the cat chasing the butterfly.

Pictures 3/4:

The butterfly flew away quickly and the cat fell into a bush. He *hurt* himself and was very *angry*. The boy was so *startled* that the ball fell out of his hand. When he *saw* his ball rolling into the water, he *cried*: "Oh no, there goes my ball". He was *sad* and wanted to get his ball back. Meanwhile, the cat *noticed* the boy's bucket and *thought*: "I want to grab a fish."

Pictures 5/6:

At the same time, the boy began pulling his ball out of the water with his fishing rod.

He did not *notice* that the cat had grabbed a fish. In the end, the cat was very *pleased* to eat such a tasty fish and the boy was *happy* to have his ball back.

Pictures 1/2:

One day there was a *playful* dog who *saw* a grey mouse sitting near a tree. He leaped forward because he wanted to catch it. Meanwhile, a *cheerful* boy was coming back from shopping with a bag and a balloon in his hands. He *looked* at the dog chasing the mouse.

Pictures 3/4:

The mouse ran away quickly and the dog bumped into the tree. He *hurt* himself and was very *angry*. The boy was so *startled* that the balloon slipped out of his hand. When he *saw* his balloon flying into the tree, he cried: "Oh no, there goes my balloon". He was *sad* and wanted to get his balloon back. Meanwhile, the dog *noticed* the boy's bag and *thought*: "I want to grab a sausage."

Pictures 5/6:

At the same time, the boy began pulling his balloon out of the tree.

He did not *notice* that the dog had grabbed a sausage. In the end, the dog was very *pleased* to eat such a tasty sausage and the boy was *happy* to have his balloon back.

Macrostructure and internal state terms are identical across languages in *LITMUS-MAIN*, as all languages use the same stimulus pictures and score story structure components in production and comprehension in the same way (see section 4). The number and sequence of Goals, Attempts, Outcomes and internal state terms (as initiating events and as reactions) per protagonist is identical across the parallel versions of the stimulus scripts. This was done for all the languages that *LITMUS-MAIN* has been piloted in, so that bilingual children can be assessed in both of their languages in an identical and reliable fashion.When developing the different language versions, the microstructure in the story scripts was also designed to be as similar as possible across stories. Lists were kept of grammatical/lexical difficulties which occurred during adaptation and of important variations due to language-specific requirements.

The logical sequence of clauses/utterances was kept the same across languages and stories, as were many other linguistic features. For the full set of guidelines for adapting the story scripts to other languages, see Gagarina et al., 2012.

3.4 Narrative comprehension

The comprehension section in *LITMUS-MAIN* provides additional opportunities for the child to demonstrate understanding of macrostructure. Children with limited proficiency in one language may score low on the production measures but, when probed with focused questions, may demonstrate an understanding of the macrostructure elements.

Ten questions are asked for each story. Three questions elicit goal statements, e.g. *Why does the mother bird fly away?* Another six questions elicit internal state terms connected either to the initiating event or to reactions, e.g. *How does the fox feel*? If the child does not provide an explanation or rationale for his/her answer, an additional question is asked, e.g. *Why do you think that the fox is feeling...?* These questions assess inference, i.e. the child's ability to interpret physical and emotional cause-effect relationships and recognize characters' goals, the reasons for these goals and reactions following attempts to reach the goals (Hedberg & Westby, 1993). Finally, the aim of the last question is to see if the child can infer meaning about the story as a whole, e.g. *Who does the mother goat like best, the fox or the bird? Why?*

3.5 Procedures

3.5.1 Presentation modes

The stimulus pictures are presented in fold-out fashion. Fold-out presentation was motivated in part by the results of previous studies on narratives with both typically developing preschoolers and preschoolers with language impairment (Gazella & Stockman, 2003; Liles, 1993). The pictures are dense in pictorial content, and the intent was to facilitate production of the three episodes in each story by presenting the pictures two at a time during both the telling and retelling tasks. Eye movement studies of children listening to stories presented orally also show evidence that children with language impairment attend to less semantically relevant information when all the pictures are presented in one go (Andreu et al., 2011). Thus, it was decided to present the pictures in fold-out fashion, initially showing the whole story with all the six pictures. When the child is ready to tell the story, the pictures are unfolded in pairs of two (pictures 1 and 2 unfolded; pictures 1 to 4 unfolded; pictures 1 to 6 unfolded).

To control for effects of shared knowledge and joint attention, the child chooses a story from one of three envelopes and is instructed not to let the examiner see which story is selected (Serratrice, 2007; Van der Lely, 1996). In addition, during initial viewing and telling/retelling of the story, the pictures are unfolded in such a way that only the child sees them.

An alternative mode involves presentation of the story on a computer screen with audio input via headphones, using the same procedure as described above for the paper version. On the screen, the child first sees three colored envelopes and is asked to choose one. The child 'clicks' on his/her choice and then sees a PowerPoint® presentation, with identical timing for exposure duration and transitions for all languages. The child first sees the entire set of six pictures in the middle of the screen. Then each pair of pictures is presented just like in the paper version, the only difference being that the child pushes a key on the keyboard in order to proceed to the next pair of pictures.

One unresolved difference across research groups is the size of the screen used with the computer version. The 9x9cm size used for each picture in the paper version is difficult to maintain for the computer version due to limits in screen sizes.

3.5.2 Administering the telling and retelling/model story tasks

Instructions for administering the *LITMUS-MAIN* are provided in Appendix A for the Baby Goats story. A detailed manual with guidelines and the full set of instructions is available online at http://www.zas.gwz-berlin.de/zaspil56.html.

Bilingual considerations in the procedures include the following: i) native-speaker examiners for each language (in order to promote a monolingual context (Grosjean, 1998)), ii)

4 to 7 day test intervals between sessions in each language in order to minimize cross-language influence as well as training and carry-over effects, and iii) implementation of the 2x2 design with the Baby Birds/Baby Goats stories for telling and the Cat/Dog stories for retelling/model story to allow counterbalancing of language and task.

4 Transcription, Coding and Scoring

4.1 Transcription and scoring of macrostructure

After a session with a child has ended, his/her narrative(s) are transcribed, and production and comprehension are scored on the scoring sheets. Several transcription systems were initially considered, and CHILDES/CHAT (MacWhinney, 2000) was selected. The full set of macrostructure features assessed is given in Table 2. Coding is carried out manually.

The *LITMUS-MAIN* protocol begins with a brief set of background information concerning the child's birthdate, date of testing, age at the time of testing, gender, the name of the data collector, the language tested, length of exposure to L2, age at the time of enrolment in preschool, and the name of the preschool (see Appendix A). This is followed by the instructions for administering the task. The scoring form is divided into two sections with the following subsections: Section I: Production (A. Story Structure; B. Structural complexity; C. Internal State Terms (ISTs)) and Section II: Comprehension (10 questions)). These are illustrated in Appendices B and C with data from a 5-year and 6-months-old English-Hebrew bilingual child doing the 'telling' task for the Baby Goats picture stimuli in L1 (which was English). The full transcript in CHILDES format is given in Appendix B. The scoring forms for the production and comprehension of Baby Goats are provided in Appendix C. They have been filled in to illustrate how the English-Hebrew child scores on the telling task and the comprehension task.

The protocols and scoring sheets for the four parallel stories of MAIN spell out which kind of prompting to use in order to elicit natural data and avoid echoic narration.

The scoring system for the story structure components (Setting, Goals, Attempts, Outcomes, Internal State Terms) is designed to accommodate different languages and different formulations for the same macrostructure components. The potential range of scores for story structure components is large enough so as to avoid ceiling effects, with a maximum of 17 points for story structure components in production and a maximum of 10 points in comprehension (Gagarina et al., 2012: 109-134). No weighting system is imposed for structural complexity; rather, the scoring includes a measure of how often a child produces partial event sequences (AO, single G), incomplete (GA/GO) and complete episodes (GAO).

Unexpected responses, which appeared in pilot data led to revisions in scoring for story production and comprehension questions. One major goal was to make the scoring system and data evaluation user-friendly. Additional scoring and evaluation guidelines are provided in the manual (http://www.zas.gwz-berlin.de/zaspil56.html; Chapter 3). See Appendices B and C for detailed examples of the coding and scoring of macrostructure categories and internal state terms.

Table 2. Macrostructure features coded and assessed

I. Production

A. Story structure components: Goals, Attempts, Outcomes, goal-related and outcome-related internal state terms (initiating events and reactions)

B. Structural complexity

C. Internal State Terms

II. Comprehension

D. Goals, Internal State Terms, Inferencing, Theory of Mind

4.2 Transcription and coding of microstructure

Based on the literature overview, the following ten features were chosen to serve as the initial basis for the microstructure analysis of narratives in *LITMUS-MAIN* (Table 3).

Table 3. Microstructure features coded and assessed (as an initial basis recommended to be assessed for all languages)

A. Narrative length and lexis

- Total number of word tokens with mazes (TNTm), where only those words related to the pictorial content of the story are included (extraneous material is excluded). Linguistic dysfluency (Sieff & Hooyman, 2006), also called mazes (Loban, 1976) of all four categories – hesitations, fillers, repetitions, and revisions (Fiestas et al., 2005) are considered in this analysis.
- Total number of word tokens without mazes (TNT). It may be informative to compare narrative length once all mazes have been detracted.
- Number of different words (NDW), where words are viewed as equivalent to lemmas or root forms.

- Number of communication units (c-units, also CU). For microstructure of oral language samples, recorded speech must be segmented into one base unit. Options included: utterances (MacWhinney 2000), t-units (= minimally terminable units, Hunt 1965), and c-units (= communication units, CUs, Loban 1976). The c-unit was chosen as the base unit to allow for straightforward comparison of results between research groups.

B. Syntactic complexity and discourse cohesion

- Mean length of c-units (calculated as number of CUs divided by TNT).
- Mean length of the three longest c-units (calculated as 3 longest CUs divided by TNT).
- Proportionof verb-based clauses (calculated as percentage of the total number of verb-based clauses out of CUs).
- Proportion of subordinating constructions (calculated as percentage of subordinate constructions out of CUs).
- Proportion of coordinating constructions, excluding the conjunction *and* (calculated as percentage of coordinating constructions out of CUs), if *and* is used as an (repeated) additive cohesion marker (like in Appendix B, lines 4, 5, 7, etc.: *And then somebody came ... And then he wanted to eat him ... And then the mom was want to drink ...*). This type of use should not be confused with *and* as a coordinating conjunction (e.g. *The mother bird flew away and came back with a worm*).

C. Bilingualism

- Number and percentage of word tokens *not* in the target language of a session (Code switching).

Nothing precludes the possibility for researchers to analyze narratives elicited by *LITMUS-MAIN* from other microstructure perspectives, e.g. lexical richness, literary language style, tense marking, percentage of error-free clauses, percentage of content words vis-à-vis function words, use of different types of noun phrases (i.e. lexical, pronominal, clitic, null) for referent introduction and maintenance, etc.

4.3 Background questions

In addition to the narrative assessment tool, a set of background questions was developed (based on Gagarina et al., 2010) in order to evaluate the acquisition conditions and the quality and quantity of the input to the child in both of her/his languages. The background questions can be used as a questionnaire to be filled in by the parents and/or the preschool/school teacher, on their own or with the help of an experimenter, if desired. They can also serve as a base for a (telephone) interview with parents and/or daycare staff. See Appendix D for the background questions.

5. Preliminary Findings

During the COST Action IS0804, different research groups conducted pilot studies based on the picture sequences and procedures described above, investigating macrostructure in the narratives of monolingual and bilingual children with and without language impairment. This was done for different languages and age groups. Members reported the results of their pilot studies at COST meetings and workshops and in online discussions. This information was used to inform the development and refinement of the *LITMUS-MAIN* materials.

During 2011-2012, narratives were collected from 267 monolingual and 302 bilingual children aged 43—109 months (i.e. 3;6—9;1 years). These include 185 monolingual TD and 82 monolingual SLI children, 285 bilingual TD and 17 bilingual SLI children. 17 languages were represented (Afrikaans, Albanian, Croatian, Cypriot Greek, Dutch, English, Finnish, French, German, Greek, Hebrew, Italian, Lithuanian, Russian, Swedish, Turkish). For the bilingual children, who told the stories in their two languages, 14 different language pairs were represented, as illustrated in Table 4. Slightly over half the children performed both telling and retelling tasks; the rest did telling only, and a few children did the 'model story' and telling.

Table 4A. Monolingual children: Languages, number of children (in order of increasing mean age).

Language	Number of children	Mean age in months
Monolingual TD children		
Afrikaans	28	65
Lithuanian	12	65
Turkish	15	65
Finnish	21	66.8
Russian	15	68
German	10	68.1
Greek	5	73
Albanian	10	78.3
Croatian	20	78.6

Albanian	14	78.8
Cypriot Greek	6	79.8
Swedish	9	80.9
French	12	81.9
French	8	82
*Monolingual SLI children**		
German	18	61.5
Lithuanian	8	65
Russian	9	68
Croatian	20	77.8
Greek	18	100.6
Afrikaans	9	106
Total	Number 267	Mean age 75.5

*Exclusionary criteria: history of hearing, neurological or developmental problems

Table 4B. Bilingual children: Language pairs, number of children (in order of increasing mean age).

Bilingual typically developing children			
L2	L1	Number of children	Mean age in months
German	Turkish	6	43.6
Swedish	Russian	10	64
Finnish	Swedish	10	65
German	Turkish	15	65
Lithuanian	Russian	6	65.3
Italian	English	13	66
German	Russian	10	66.9
English	Hebrew	9	67.3
Finnish	Russian	10	68.2
Hebrew	English	21	70.1
Swedish	English	20	77
Afrikaans	English	10	78
German	Turkish	7	78
Greek	Albanian	6	78
Italian	English	25	78
English	Polish	12	78.5
Afrikaans	English	20	79
Swedish	English	16	81
Swedish	French	21	82
Dutch	Turkish	6	94.3
Afrikaans	English	19	103

German	Russian	22	109
*Bilingual SLI children**			
L2	L1	N	Mean age in months
Swedish	Russian	5	66
Greek	Albanian/German/ Bulgarian/ French/ Romanian	12	111.3
Total		Number 302	Mean age 76.4

*Exclusionary criteria: history of hearing, neurological or developmental problems

Four categories of macrostructure data were assessed: Story structure (including setting, goal, attempt, outcome, internal state terms as initiating events and as reactions), structural complexity, internal state terms, and story comprehension. Only preliminary trends from these pilot studies are reported here.

First, in the narratives of both monolingual TD and bilingual TD children, story structure components in production increase as a function of age, from scores of 4-5 points (out of 17) at 65 months to 9-10 points at 109 months.

Second, for monolinguals with SLI, children from four out of six groups show substantially lower story structure scores than their TD peers. Story complexity (i.e. production of full GAO episodes) proved to be generally low for most children, but still lower for children with SLI as compared to TD children. This measure did in fact differentiate TD and SLI children for some languages (cf., Skerra et al. 2013 for German; Gagarina & Schuktomow, in preparation, for Russian). Internal state terms in the narratives of SLI children were less frequent than for their TD peers.

Third, bilingual children score similarly to monolingual peers, and TD bilingual children score equally high on story structure in both their languages even when parents report one of the child's languages to be weaker than the other. This preliminary finding is particularly important, since it may show that the instrument is less biased against bilingual children than norm-referenced language assessment tools. Similar story structure scores in both languages also provide support for the view that abilities at the macrostructure level, i.e. composition of cohesive event sequences, reflect capacities which go beyond the specifics of one language (Berman & Slobin, 1994; Pearson, 2002; Fiestas & Peña, 2004; see Section 1). The analysis of the narratives' microstructure, which is highly language-specific, will be the next step in the development of the instrument and should be performed initially on a single language level before cross-linguistic comparisons are carried out.

Fourth, the TD bilingual children in these pilot studies score similarly on story comprehension to monolinguals, with a range from 2.5 points for the very young (at 43.6 months) to 8-9 points (out of 10) for older children. TD bilingual children who produce few macrostructure components often still show evidence of good story comprehension.

And finally, bilingual children score very similarly on story comprehension in both their languages, even when one language is reported by parents to be weaker than the other. For 10 out of the 14 bilingual groups in the data, story comprehension scores differed at most 1 point between L1 and L2. If upheld more generally, this would suggest that story comprehension at the macrostructure level is not language-specific, but language universal.

6. Summary and Conclusion

LITMUS-MAIN was developed within COST Action IS0804 as a tool for the evaluation of narrative abilities of bilingual children across languages. The intent was to develop materials for the assessment of narratives in both languages of bilingual children, in order to screen and identify children at risk for Specific Language Impairment (SLI). Two sets of parallel picture sequences and accompanying story scripts that are controlled for macro- and microstructural features were developed, as well as guidelines for implementation, and protocols for administering and scoring. Materials development was informed by the experience gathered while pilot versions of the assessment tool were tested in more than 20 countries, with children speaking 17 different languages and 14 different language pairs. *LITMUS-MAIN* is currently available in 26 languages (<http://www.zas.gwz-berlin.de/zaspil56.html>).

The design of *LITMUS-MAIN* allows for the elicitation of narratives in three modes: i) telling (story generation), ii) retelling and, iii) telling after listening to a model story. A set of comprehension questions which focus on macrostructure components and internal state terms also forms part of the assessment procedure.

The instrument can be used to collect data from bilingual children with and without diagnosed language impairment for a variety of languages and language combinations. This allows for cross-linguistic comparisons and the development of theoretical perspectives. *LITMUS-MAIN* also provides clinicians with a diagnostic tool to guide and inform intervention in children with language impairment.

References

Andreu, L., Sanz-Torrent, M., Guàrdia Olmos, J. & MacWhinney, B. (2011). Narrative comprehension and production in children with SLI: An eye movement study. *Clinical Linguistics and Phonetics, 25*(9), 767-783.

Berman, R. A. and Slobin, D. I. (1994). Narrative structure. In R. A. Berman and D. I. Slobin (Eds.), *Relating events in narrative: A cross-linguistic developmental study* (pp. 39-84). New York/London: Psychology Press.

Biddle, K.A., McCabe, A., and Bliss, L. (1996). Narrative skills following traumatic brain injury in children and adults. *Journal of Communication Disorders, 29,* 447-469.

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

Bliss, L., McCabe, A. and Miranda, E. (1998). Narrative assessment profile: Discourse analysis for school-age children. *Journal of Communication Disorders, 31*, 347-363.

Botting, N. (2002). Narrative as a tool for the assessment of linguistic and pragmatic imairments. *Child Language Teaching and Therapy, 18*, 1-21.

Boudreau, D.M. (2007). Narrative abilities in children with language impairments. In R. Paul (Ed.), *Language disorders from a developmental perspective: Essays in honour of Robin S. Chapman* (pp. 331-356). Mahwah, NJ: Lawrence Erlbaum Associates.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard Univ. Press.

Cleave, P. L., Girolametto, L. E., Chen, X. and Johnson, C. J. (2010). Narrative abilities in monolingual and dual language learning children with specific language impairment. *Journal of Communication Disorders, 43*(6), 511-522.

Curenton, S. M. and Justice, L. M. (2004). *Use of literate language features in low-income preschoolers' narratives*. Paper presented at the Head Start 7th National Research Conference, Washington, DC.

Dodwell, K. and Bavin, E. L. (2008). Children with specific language impairment: An investigation of their narratives and memory. *International Journal of Language and Communication Disorders, 43*(2), 201-218.

Fiestas, C. E. and Peña, E. D. (2004). Narrative discourse in bilingual children:Language and task effects. *Language, Speech, and Hearing Services in Schools, 35*(2), 155-168.

Fiestas, Ch. E., Bedore, L. M., Pena, E. D., and Nagy, V. J. (2005). Use of mazes in the narrative language samples of bilingual and monolingual 4- to 7-year old children. In J. Cohen, K. T. McAlister, K. Rolstad and J. MacSwan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 730-740). Somerville, MA: Cascadilla Press.

Gagarina, N., Klassert, A. and Topaj, N. (2010). Russian language proficiency test for multilingual children. *ZAS Papers in Linguistics, 54*.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U. and Walters, J. (2012). MAIN: Multilingual Assessment Instrument for Narratives. *ZAS Papers in Linguistics, 56.* (Includes manual and MAIN materials for assessment, available online at http://www.zas.gwz-berlin.de/zaspil56.html)

Gagarina, N. and Schuktomow, R. (in prep.). Macrostructure in elicited narratives of preschool monolingual Russian children with and without specific language impairment.

Gazella, J. and Stockman, I. J. (2003). Children's story retelling under different modality and task conditions: Implications for standardizing language sampling procedures. *American Journal of Speech-Language Pathology, 12*(1), 61-72.

Gillam, R. B. and Carlisle, R. (1997). Oral reading and story retelling of students with specific language impairment. *Language, Speech, and Hearing Services in Schools, 28*, 30-42.

Gillam, R. B. and Johnston, J. R. (1992). Spoken and written language relationships in language/learning-impaired and normally achieving school-age children. *Journal of Speech and Hearing Research, 35*(6), 1303-1315.

Gillam, R. B. and Pearson, N. A. (2004). *Test of Narrative Language (TNL)*: PRO-ED.

Glasgow, C. and Cowley, J. (1994). *Renfrew Bus Story - North American Edition (RBS-NA)*: Centerville School.

Grazzani, I. and Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development, 27,* 99-11.

Graybeal, C. M. (1981). Memory for stories in language-impaired children, *Applied Psycholinguistics*, *2*, 269-283.

Greenhalgh, K. S. and Strong, C. J. (2001). Literate language features in spoken narratives of children with typical language and children with language impairments. *Language, Speech, and Hearing Services in Schools, 32*, 114-125.

Grosjean, F. (1998). Transfer and language mode. *Bilingualism: Language and Cognition, 1*(3), 175-176.

Gülzow, I. and Gagarina, N. (2007). Noun phrases, pronouns and anaphoric reference in young children narratives. *ZAS Papers in Linguistics, 48*, 203-223.

Gutiérrez-Clellen, V. F., 2002. Narratives in two languages: Assessing performance of bilingual children. *Linguistics and Education, 13(2*), 175-197.

Gutiérrez-Clellen, V., Simon-Cereijido, G. and Wagner, C. (2008). Bilingual children with language impairment:A comparison with monolinguals and second language learners. *Applied Psycholinguistics, 29*, 3-20.

Gutiérrez-Clellen, V. F., Simon-Cereijido, G. and Erickson Leone, A. (2009). Code-switching in bilingual children with specific language impairment. *International Journal of Bilingualism, 13*(1), 91-109.

Hadley, P. A. (1998). Language sampling protocols for eliciting text-level discourse. *Language, Speech, and Hearing Services in Schools, 29*, 132-147.

Hayward, D. and Schneider, P. (2000). Effectiveness of teaching story grammar knowledge to preschool children with language impairment. An exploratory study. *Child Language Teaching and Therapy, 16*, 255-284.

Hayward, D. V., Gillam, R. B. and Lien, P. (2007). Retelling a script-based story: Do children with and without language impairments focus on script and story elements? *American Journal of Speech-Language Pathology, 16*, 235-245.

Hedberg, N. L. and Westby, C. E. (1993). *Analyzing storytelling skills. Theory to practice*. Tucson, AZ: Communication Skill Builders.

Heilmann, J., Miller, J. F. and Nockerts, A. (2010a). Sensitivity of narrative organization measures using narrative retells produced by young school-age children. *Language Testing, 27*, 603-626.

Heilmann, J., Miller, J. F., Nockerts, A. and Dunaway, C. (2010b). Properties of the narrative scoring scheme using narrative retells in young school-age children. *American Journal of Speech-Language Pathology, 19*, 154-166.

Hickmann, M. (2003). *Children's discourse: person, space and time across languages*. Cambridge: Cambridge University. Press.

Hughes, D., McGillvray, L. and Schmidek, M. (1997). *Guide to narrative language: Procedures for assessments*. Eau Claire, WI: Thinking Publications.

Hunt, K. (1965). *Grammatical structures written at three grade levels*. NCTE Research report No. 3. Champaign, IL: National Council of Teachers of English.

Iluz-Cohen, P. and Walters, J. (2012). Telling stories in two languages: Narratives of bilingual preschool children with typical and impaired language. *Bilingualism: Language and Cognition, 15*(1), 58-74.

Kaderavek, J. N. and Sulzby, E. (2000). Narrative production by children with and without specific language impairment: Oral narratives and emergent readings. *Journal of Speech, Language, and Hearing Research, 43*, 34-49.
Kit-Sum To, C., Stokes, S. F., Cheung, H. and T'sou, B. (2010). Narrative Assessment for Cantonese-Speaking Children. *Journal of Speech, Language and Hearing Research, 53(3)*, 648-669.
Klop, D., Visser, M. and Oosthuizen, H. (2011). Narratives in bilingual South African children. *Paper presented at 4th Meeting of COST Action IS0804*. Eskişehir, Turkey.
Klop, D., Visser, M. and Oosthuizen, H. (2012). Narrative profiles of 20 bilingual normally developing 6-7 year old South African children. *Paper presented at 6th Meeting of COST Action IS0804.* Berlin, Germany.
Labov, W. and Waletzky, J. (1967). Narrative analysis. In J. Helm (Ed.), *Essays on the verbal and visual arts* (pp. 12-44). Seattle: University of Washington Press.
Liles, B. Z. (1993). Narrative discourse in children with language disoders and children with normal language:A critical review of the literature. *Journal of Speech and Hearing Research, 36*, 868-882.
Liles, B. Z., Duffy, R. J., Merritt, D. D. and Purcell, S. L. (1995). Measurement of narrative discourse ability in children with language disorders. *Journal of Speech and Hearing Research, 38*, 415-425.
Loban, W. (1976). *Language development: Kindergarten through grade twelve.* Urbana, IL: National Council of Teachers of English.
Lorusso, M. L., Galli, R., Libera, L., Gagliardi, C., Borgatti, R. and Hollebrandse, B. (2007). Indicators of theory of mind in narrative production:acomparison between individuals with genetic syndromes and typically developing children. *Clinical Linguistics and Phonetics, 21*(1), 37-53.
Lucas, E.V. (1980). *Semantic and pragmatic language disorders.* Rockville, MD: Aspen.
MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
Mandler, G. and Johnson, N. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9*, 111-151.
Mayer, M. (1969). *Frog, Where Are You?* New York: Dial Books.
McCabe, A. (1996). Evaluating narrative discourse skills. In K. Cole, P. Dale and D. Thal (Eds.), *Assessment of communication and language* (pp. 121-142). Baltimore, MD: Paul H. Brookes.
McCabe, A. and Rollins, P. R. (1994). Assessment of preschool narrative skills. *American Journal of Speech-Language Pathology, 3*(1), 45-56.
Newman, R. M. and McGregor, K. K. (2006). Teachers and laypersons discern quality differences between narratives produced by children with and without SLI. *Journal of Speech Language and Hearing Research, 49*(5), 1022-1036.
Nippold, M. A., Ward-Lonergan, J. M. and Fanning, J. L. (2005). Persuasive writing in children, adolescents, and adults: A study of syntactic, semantic, and pragmatic development. *Language, Speech, and Hearing Services in Schools, 36*, 125-138.
Oakhill, J. and Cain, K. (2007). Issues of causality in children’s reading comprehension. In D. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 47-72). New York: Erlbaum.
Olley, L. (1989). Oral narrative performance of normal and language impaired school aged children. *Australian Journal of Human Communication Disorders, 17*, 43-65.
Paradis, J., Genesee, F. and Crago, M. (2011). *Dual language development and disorders: A handbook on bilingualism and second language learning* (2nd ed.). Baltimore, MD: Paul H. Brookes.

Pearson, B. Z. (2001). Language and mind in the stories of bilingual children. In L. Verhoeven and S. Strömqvist (Eds.), *Narrative development in multilingual contexts* (pp. 373-398). Amsterdam, NL: John Benjamins.

Pearson, B. Z. (2002). Bilingual Infants:What we know,what we need to know. In M. Suarez-Orozco and M. Paéz (Eds.), *Latinos: Remarking America* (pp. 306-320). Berkeley, CA: University of California Press.

Pearson, B. Z. and de Villiers, P. A. (Eds.). (2005) *Encyclopedia of language and linguistics* (2nd ed.). Oxford, UK: Elsevier.

Reich, H. H. and Roth, H.-J. (2004). *Hamburger Verfahren zur Analyse des Sprachstands Fünfjähriger - HAVAS 5*.

Reilly, J., Losh, M., Bellugi, U. and Wulfeck, B. (2004). "Frog, where are you?" Narratives in children with specific language impairment, early focal brain injury, and Williams syndrome. *Brain and Language, 88*, 229-247.

Roth, F. P. and Spekman, N. J. (1986). Narrative discourse: Spontaneously generated stories of learning-disabled and normally achieving students. *Journal of Speech and Hearing Disorders, 51,* 8-23.

Schneider, P. and 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. and Hayward, D. (2005). The Edmonton Narrative Norms Instrument (Publication: http://www.rehabresearch.ualberta.ca/enni).

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

Serratrice, L. (2007). Referential cohesion in the narratives of bilingual English-Italian children and monolingual peers. *Journal of Pragmatics, 39*, 1058-1087.

Sieff, S. and Hooyman, B. (2006). Language-based disfluency: A child case study. *Hearing Association Annual Convention*. Miami, FL.

Simon-Cereijido, G. and Gutiérrez-Clellen, V. (2009). A cross-linguistic and bilingual evaluation of the interdependence between lexical and grammatical domains. *Applied Psycholinguistics, 30*(02), 315-337.

Skerra, A., Adani, F. and Gagarina, N. (2013). Diskurskohäsive Mittel in Erzählungen als diagnostischer Marker für Sprachentwicklungsstörungen. In T. Fritzsche, C. B. Meyer, A. Adelt and J. Roß (Eds.), *Spektrum Patholinguistik, 6*, 123–154. Potsdam: Universitätsverlag Potsdam.

Stein, N. L. and Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Discourse processing: Multidisciplinary perspectives* (pp. 53-120). Norwood, NJ: Ablex.

Strong, C. J. and Shaver, J. P. (1991). Stability of cohesion in the spoken narratives of language-impaired and normally developing school-aged children. *Journal of Speech and Hearing Research, 34*(1), 95-111.

Swanson, L., Fey, M., Mills, C. and Hood, L. (2005). Use of narrative-based language intervention with children who have specific language impairment. *American Journal of Speech-Language Pathology, 14*, 131-143.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Trabasso, T. and Nickels, M. (1992). The development of goal plans of action in the narration of a picture story. *Discourse Processes, 15*, 249-275.

Trabasso, T. and Rodkin, P. C. (1994). Knowledge of goals/plans:A conceptual basis for narrating "Frog where are you?". In R. A. Berman and D. I. Slobin (Eds.), *Relating* *events in narrative: A cross linguistic developmental study* (pp. 85-106). Mahwah, NJ: Lawrence Erlbaum Associates.

Uccelli, P. and Paéz, M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associationsamong English and Spanish skills. *Language, Speech, and Hearing Services in Schools, 38*, 225-236.

Van der Lely, H. K. J. (1996). Grammatical specific language impaired children: Evidence for Modularity. In C. Koster and F. Wijnen (Eds.), *The Groningen Assembly on Language Acquisition* (pp. 273-282). Groningen, Netherlands: University of Groningen Press.

Walach, G. P. (2008). *Language intervention for school-age students: Setting goals for academic success*. St. Louis, MO: Mosby.

Westby, C. E. (1991). Assessing and remediating text comprehension problems. In A. Kamhi and H. Catts (Eds.), *Reading disabilities: A developmental language perspective*. Boston, MA: Allyn & Bacon.

Westby, C. E. (2005). Assessing and facilitating text comprehension problems. In H. Catts and A. Kamhi (Eds.), *Language and reading disabilities* (pp. 157-232). Boston, MA: Allyn & Bacon.

Appendix A. Instructions for administering MAIN, Baby Goats story (Telling)
(For detailed guidelines and instructions, see Gagarina et al., 2012)

Protocol for Baby Goats (Telling)

Name of child: ____________________
Date of birth: ____________________
Date of testing: ____________________
Age of testing (in months): ____________________
Gender: ____________________
Name of examiner: ____________________
Exposure to L2 (in months): ____________________
Kindergarten entry date: ____________________
Name of kindergarten: ____________________

Be sure that all the envelopes are on the table before testing begins.
Prepare the audio recorder in order to record the session. Begin recording before warming up.

Warming-up
Ask for example: Who is your best friend? What do you like to watch on TV? Do you like telling stories? Do you like listening to stories?

Instructions:
Sit opposite the child. Say to the child: Look, here are three envelopes. There is a different story in each envelope. Choose one and then you can tell me a story. *Unfold the pictures so that the whole sequence is visible to the child only.* First look at the whole story. Are you ready?

Unfold the first two pictures. *Say to the child*: Now I want <u>you</u> to tell the story. Look at the pictures and try to tell the <u>best</u> story you can. *Allowable prompt if the child is reluctant to begin*: “Tell me a story about this picture” (*point to picture)*. *When the child has finished telling the first two pictures, unfold the next (so that all pictures from 1 to 4 are visible). Repeat the process until the end of the story. Allowable prompts if the child is silent in the middle of the story:* “Anything else?”, “Continue”, “Tell me more”, “Let’s see what else is in the story”. *If the child stops talking without indicating that he/she has finished, say:* “Tell me when you are finished”.
When the child has finished, praise the child and then ask the comprehension questions.

Appendix B. An example of a transcript of the Baby Goats story told by an English-Hebrew bilingual child age 5 years 6 months

1. *CHI: The mom wants to get the baby.
2. *CHI: And (be)cause he fell inside the water.
3. *CHI: He wanted to drink.
4. *CHI: And then somebody came.
5. *CHI: And then he wanted to eat him.
6. *CHI: Then he got his foot.
7. *CHI: And then the mom was want to drink.
8. *CHI: And then saw the baby was drinking.
9. *CHI: And the bird he &he said that.
10. *CHI: The bird bited his tail.
11. *CHI: And then and they saw it.
12. *CHI: And then they saw the bird and biting his tail.
13. *CHI: And then saw the mom.
14. *CHI: And then saw the baby.

Appendix C. Scoring protocol for Baby Goats story
The narrative provided in Appendix B is scored here to provide an example of how the scoring is done.

Section I: Production

A. Story Structure; B. Structural complexity; C. Internal State Terms (IST)

A. Story Structure

		Examples of correct responses[1]	Score	Comments[2]
A1.	Setting	Time and/or place reference, e.g. once upon a time/one day/long ago... In a forest/in a meadow/at the lake/at the pond…	0 1 2[3]	
Episode 1: Mother/Goat (episode characters: baby goat and mother/goat)				
A2.	IST as initiating event	**Baby goat** was scared/in danger/drowning/needed help/cried/called the mother **<Mother/Goat etc.>** saw that baby goat was scared/in danger/drowning/couldn't swim/was worried about the baby goat in the water	0 1	
A3.	Goal	**Mother goat** wanted to help the baby/to save/rescue the baby/to push the baby out of the water	0 1	1 point "wants to get the baby"
A4.	Attempt	**Mother goat** ran/went into the water/is pushing	0 1	
A5.	Outcome	**Mother goat** pushed the baby out of the water/saved/rescued the baby **Baby goat** was saved/out of the water	0 1	
A6.	IST as reaction	**Mother goat** was happy/relieved **Baby goat** was relieved/satisfied/happy/glad/not scared any more	0 1	
Episode 2: Fox (episode characters: fox and baby goat)				
A7.	IST as initiating event	**Fox** saw mother looking away/saw that the baby was alone/saw that there was food/fox was hungry	0 1	
A8.	Goal	**Fox** wanted to eat/catch/kill the baby goat	0 1	1 point "wanted to eat him"
A9.	Attempt	**Fox** jumped towards/jumped up/jumped out/tried to reach/grab/catch the baby goat	0 1	
A10.	Outcome	**Fox** got/grabbed/caught the baby goat	0 1	1 point "got his foot"
A11.	IST as reaction	**Fox** was happy **Baby** goat was scared	0 1	

1 If in doubt or the response of the child is not on this scoring sheet consult the manual.
2 Write down responses here or indicate No response.
3 Zero points for wrong or no response, 1 point for one correct response, 2 points for reference to both *time* and *place*.

Episode 3: Bird (episode characters: bird, fox and baby goat)				
A12.	IST as initiating event	**Bird** saw that the goat was in danger **Baby goat** was in danger	0 1	
A13.	Goal	**Bird** decided/wanted to stop the fox / help/protect/save the baby goat	0 1	
A14.	Attempt	**Bird** bit/dragged the fox's tail/ attacked/chased the fox	0 1	1 point "bited his tail"
A15.	Outcome	**Bird** chased the fox away **Fox** let go of the baby goat/ran away **Baby goat** was saved / rescued	0 1	
A16.	IST as reaction	**Bird** was relieved/happy/proud to have saved / rescued the baby goat/ **Fox** was angry/disappointed **Baby goat/goats** was/were relieved/happy/safe	0 1	
A17.			**Total score out of 17:**	4 out of 17

B. Structural complexity

(Note that this results from subsection A.)

Number of AO Sequences	Number of single G (without A or O)	Number of GA / GO sequences	Number of GAO sequences
B1.	B2.	B3.	B4.
(0)	(1)	(1)	(0)

C. Internal State Terms (IST)

C1.	Total number of IST in **tokens**. IST include: Perceptual state terms, e.g. *see, hear, feel, smell;* Physiological state terms, e.g. *thirsty, hungry, tired, sore*; Consciousness terms, e.g. *alive, awake, asleep*; Emotion terms, e.g. *sad, happy, angry, worried, disappointed*; Mental verbs, e.g. *want, think, know, forget, decide, believe, wonder, have/make a plan*; Linguistic verbs/verbs of saying/telling, e.g. *say, call, shout, warn, ask*	10 (want, wanted, wanted, want, saw, said, saw, saw, saw, saw)

Section II: Comprehension

The right-hand column shows how the child from Appendices B and C responded to the comprehension questions and how these were scored. (Question D10 was not asked.)

		Examples of correct responses	Examples of wrong responses	Score	Comments
0	Did you like the story?	Warm-up question, not scored			
D1.	Why was the mother goat in the water? (point to pictures 1-2) (Episode 1: Goal/IST as initiating event)	Wants to save/to help/rescue/ worried about the baby/the baby goat is in danger/ drowning/scared/ the baby was crying	Is swimming/ playing/wants to take a bath/to wash herself/to wash the baby goat	0 1	To get the baby (1 point)

		for help			
D2.	How does the baby goat feel? (point to baby goat in the water, picture 1) (IST as initiating event)	Bad/scared/in danger/horrified	Good/fine/ happy/ playing/freezing/ cold/hungry/thirsty/ dirty/clean/stupid/refreshed	0 1	Sad (1 point)
D3.	*(only ask D3 if the child gives a correct response without explanation/rationale in D2. If a correct explanation is provided in D2, then give a point in D3 and proceed to D4)* Why do you think that the baby goat is feeling bad/scared/in danger etc.?[4]	Because he has fallen into the water/is not able to get out of the water/is drowning/ cannot swim	Because he is hungry/swimming/ playing in the water/wasn't allowed to stand there	0 1	(Be)cause he's stuck inside the water (1 point)
D4.	Why does the fox leap forward? (point to picture 3) (Episode 2: Goal)	Wants/to get/to kill/to eat the baby goat/couldn't resist to eat the baby goat/ takes the opportunity when mother is not looking/is far away	To play with the baby goat	0 1	(Be)cause he wants to eat him (1 point)
D5.	How does the fox feel? (point to picture 5-6) (IST as reaction)	Bad/sad/angry/ mad/scared/still hungry/hurt/stupid/disappointed	Good/fine/happy/ playful	0 1	Hungry (1 point)
D6.	*(only ask D6 if the child gives a correct response without explanation/ rationale in D5. If a correct explanation is provided in D5, then give a point in D6 and proceed to D7)* Why do you think that the fox is feeling bad/scared/hungry/ disappointed etc.?[5]	Because he did not get the baby goat/he was still hungry/ afraid/scared of the bird/the bird was biting/ chasing him	Because the bird saw that the goat was in danger/the fox is running away/I don't know	0 1	
D7.	Why does the bird bite the fox's tail? (point to picture 5) (Episode 3: Goal)	Wants/decided to save/rescue the baby goat/wants to stop the fox/to make the fox let the goat go/saw that the goat was in danger	Wants to eat the fox/eat the goat/ play with the fox	0 1	(Be)cause he didn't want the fox to eat the baby (1 point)
D8.	Imagine that the bird sees the goats. How does the bird feel?	Good/fine/happy/ relieved/satisfied/ proud/like a hero	Bad/sad/angry/mad /sorry/stupid/"I have to get the fox"	0 1	Hm, sad (0 points)

4 Use the same IST provided by the child in response to D2.
5 Use the same IST provided by the child in response to D5.

	(point to picture 6)				
D9.	*(only ask D9 if the child gives a correct response without explanation/rationale in D8. If a correct explanation is provided in D8, then give a point in D9 and proceed to D10)* Why do you think that the bird is feeling good/fine/happy etc.?[6]	Because he stopped the fox/got the fox out of there/ saved/rescued the goat/sees that the goats are happy/unharmed/now the fox won't come back	Because he is smiling/angry at the fox/wants to eat the baby goat himself	0 1	(Be)cause he wanted to; (be)cause he took his foot (0 points)
D10.	Who does the mother goat like best, the fox or the bird? Why?	The bird – give at least one reason (he saved/helped the baby goat/ chased the fox away)	The fox/I don't know / other irrelevant answer	0 1	This question was not asked
D11.				**Total score out of 10:**	6 out of 10

6 Use the same IST provided by the child in response to D8.

Appendix D: Background questions

1. Child's name (forename, surname) ____

2. Date of birth ____

3. Does your child currently go to a kindergarten/ day care/ school?

○ Yes, kindergarten from _____	○ Yes, school from _________
○ No	○ No
If yes, what kind of kindergarten?	If yes, what kind of school?
○ Bilingual	○ Bilingual
○ Monolingual L1 = child's native language	○ Monolingual L1 = child's native language
○ Monolingual L2 = child's second language	○ Monolingual L2 = child's second language
○ Other. What kind of other? ________ ___________	○ Other. What kind of other? _________________________________

4. In what country was your child born?

○ In country of L1, which is ____

○ In country of L2, which is ____

○ In other country, which is ____

5. Since when has your child lived in the country of L2? ________________(Year, Month)

6. Birth order

○ 1

○ 2

○ 3

○ Put the number

7. How old was your child when he/she spoke the first words? ___________ (year(s), months)

8. Have you ever been concerned about your child's language?

○ No

○ Yes. Specify why? ____

9. Has anyone in your family had any speech or language difficulties?

○ No

○ Yes. Specify who, e.g., mother, father, sibling(s)? ______________________

10. Has your child ever had hearing problems?
Hearing impairment?

○ No

○ Yes

Frequent ear inflections?

○ No

○ Yes. How many? ____

○ Grommets (ear tubes)

11. In your opinion, does your child hear normally?

○ No

○ Yes

12. Information about the parents

	Specify your native language (L1)	Specify your second language (L2)	Specify other languages you speak	How long have you been living in XX country?	Your education	Your occupation
Mother						
Father						

13. What language do you speak with your child?

Mother

○ My native language (L1)

○ My second language (L2)

○ Both native and second language

○ Other language(s). Specify which? ______________________________

Father

○ My native language (L1)

○ My second language (L2)

○ Both native and second language

○ Other language(s). Specify which? ______________________________

14. What languages does your child speak now?

○ Child's L1, which is ______________________________

○ Child's L2, which is ______________________________

○ Other languages, which are ______________________________

15. What languages is your child exposed to?

○ Child's L1, which is ______________________________

○ Child's L2, which is ______________________________

○ Other languages, which are ______________________________

16. At what age did your child's exposure for L2 begin?

○ From birth

○ Before age 1

○ Before age 3

○ Before age 5

○ From age ______________________________

17. Is your child exposed to L2 in

○ Kindergarten or school

○ With friends

○ With siblings/ parents/ other relatives

○ TV/ computer/ books

○ Other ______________________________

18. Estimate, in terms of percentages, how often your child is exposed to different languages per day (in all daily activities combined)?

His/ her native language (L1)	His/ her second language (L2)	Other languages
○ 25%	○ 25%	○ 25%
○ 50%	○ 50%	○ 50%
○ 75%	○ 75%	○ 75%
○ 100%	○ 100%	○ 100%

19. Please estimate your child's language skills by ticking the appropriate box.

	Very well	Quite well	Quite badly	Very badly
How well does your child <u>understand</u> his/ her native language (L1)				
How well does your child <u>understand</u> his/ her second language (L2)				
How well does your child <u>speak</u> his/ her native language (L1)				
How well does your child <u>speak</u> his/ her second language (L2)				

20. In your opinion, which language does your child speak best?

○ His/ her L1

○ His/ her L2

○ Other language, which is ______________________________

21. In your opinion, does your child like/ prefer any of the languages more than others?

○ No

○ Yes, which is ______________________________

22. Please indicate the frequency of the following activities carried out with your child during the last month

	His/ her native language (L1)				His/ her second language (L2)			
	Never	Twice a month	Once or twice a week	Almost every day	Never	Twice a month	Once or twice a week	Almost every day
Telling stories								
Reading books								
Listening to songs or singing								
Watching TV/ DVD/ computer games								