Why Roma Children Need Language Assessments in Romani

Hristo Kyuchukov  
*University of Silesia*

Jill de Villiers  
*Smith College*, jdevilli@smith.edu

Andrea Takahesu Tabori  
*Smith College*

Follow this and additional works at: [https://scholarworks.smith.edu/psy_facpubs](https://scholarworks.smith.edu/psy_facpubs)

Part of the [Philosophy of Language Commons](https://scholarworks.smith.edu/psy_facpubs), and the [Psychiatry and Psychology Commons](https://scholarworks.smith.edu/psy_facpubs)

**Recommended Citation**

Kyuchukov, Hristo; de Villiers, Jill; and Tabori, Andrea Takahesu, "Why Roma Children Need Language Assessments in Romani" (2017). Psychology: Faculty Publications, Smith College, Northampton, MA.
[https://scholarworks.smith.edu/psy_facpubs/24](https://scholarworks.smith.edu/psy_facpubs/24)

This Article has been accepted for inclusion in Psychology: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu
In this paper we make one major point: that Roma children in Europe need to be tested in their mother tongue before school placement. Roma children are in a particularly perilous position with respect to their education. We describe the problematic linguistic situation of Roma children, who are bilingual and often bidialectal, but are frequently evaluated in the language of the state for educational placement, a process that has been shown to significantly compromise their chance of success. We then review the considerable empirical evidence that bilingual children must be evaluated in both languages to give a fair assessment of their knowledge and skills. Furthermore, strength in the mother tongue has demonstrable transfer to skills in the second language. We provide a brief summary of a new assessment for Romani that has been used successfully to evaluate children aged 3 to 6 years, and present the results of a new study using it in Slovakia on 29 children aged 3 to 6 years.

**Key words**: Roma children, assessment, Romani, bilinguals

---

**What is Romani?**

Romani belongs to the new-Indian languages formed outside of India during the last 1000 years. At least 1000 years ago, Muslim incursions into India made a group of Indian people leave the territory to escape the war. The refugees belonged to different tribes, practiced different professions, and spoke different languages and dialects. On route from India to the European continent, the different tribes and groups intermixed and developed a kind of “Lingua-Franca” using...
different languages and dialects, and at the same time adapted different words, phrases and grammatical categories from several European languages - mainly from Armenian, Greek, Romanian and Slavic languages. The history of the language is well documented in several scholarly works (Bakker & Kyuchukov, 2000; Matras, 2002). Romani is now an established language in its own right, used for communication by the Roma people in different countries throughout Europe. In the last 3 or 4 decades there were attempts to codify the Romani from different countries that used different alphabets, in order to document the Romani grammar. In a recent publication, Kyuchukov (2016) analyses those attempts and presents the Romani alphabet most used for written communication.

There has been a long history of prejudice against not only the Roma people, but also the language they speak. As early as 1685, Portuguese Roma were forbidden to speak Romani (Fraser, 1992). Today, the status of Romani as an independent language is not disputed by modern linguists. In some European countries, courses in Romani are taught at universities such as Bucharest University, Charles University in Prague, Graz University, Manchester University, and Begonia University. Many international linguistic conferences have focused on the Romani language (Kyuchukov, 2009). However, like all natural languages, it contains regional dialects and variations. The varieties of Romani have not been codified for writing (Hancock, 1995). Many documents of international institutions such as the Council of Europe, European Commission, ERRC, OSCE, and Amnesty International are translated into Romani, though it was not recognized as an official language until 2000. Despite this recognition, there is still considerable prejudice against the language among ordinary people, educators, psychologists and even linguists in Europe. A number of publications show that different forms of linguistic discrimination ("linguicism") exist towards Romani in Bulgaria (Kyuchukov, 2013; Kyuchukov, 2015), in Slovakia (Kyuchukov & Balvin, 2013), and in Poland (Gawlicz, Rudnicki, & Starnawski, 2015; Grzymała-Moszczyńska, Barzykowski, Dzida, Grzymała-Moszczyńska, & Kosno, 2011).

How Roma Children Learn Romani as a Mother Tongue

Roma children typically grow up exposed to at least two languages, one Romani dialect and the official language of their country of residence. It is also common for Roma children to acquire additional languages spoken by other minority groups. While visiting a small town in the Northeast of Bulgaria, the first author met a 3-4 year old girl who was using three languages interchangeably: Romani which she learned from home, Bulgarian because she was attending a kindergarten, and Turkish because the family was living in a settlement where a Turkish-speaking minority was living as well. Several publications document how typically developing Roma children grow up as bilinguals/multilinguals:
Kyuchukov (2014) in Bulgaria, Kubanik (2016) in Czech Republic, Kyuchukov, Kaleja, and Samko (2016) in Slovakia, and Reger and Gleason (1991) in Hungary. The Roma children grow up in a culture that has a rich oral history. The children learn Romani from communication with the members of an extended family. Raising children is taken as a responsibility of the whole community. In Roma settlements in East Europe, which are isolated from the rest of the society, the children are exposed to a rich oral history and they learn the language through songs, fairy tales, language games, teasing and jokes, taking an active role in the community life, and participating in different activities. The Western way of learning the language partly through books is not familiar in the traditional Roma communities. However, the Roma ways are similar to the ways that children in Indian minority groups learn the language as well. In a visit to India, the first author conducted research among Rajasthani migrant workers living along the highways outside of New Delhi. A mother was cooking a meal while involving her 2-year-old daughter in the cooking activity, asking her to bring her different cutlery or products that she needed for cooking. The mother was naming the cutlery or the product and the task of the child was to find it among the many other objects and to bring it to her. When she failed to bring the right object the mother simply sent her back to look for it. The mother did not show the object and name it. She simply said the name of an object and the child had to find it. This is also a strategy used by Roma mothers and a common way that the Roma children learn new words – through doing, not through naming (Kubanik, 2016; Kyuchukov, 2014).

Reger and Berko-Gleason (1991) describe how a Roma boy learns to speak Romani from his father who is singing a song to him, using the rich oral tradition. Reger (1999) describes other strategies such as teasing and language games for language learning among Roma children in Hungary. The adults play with the children and tease them or make jokes with them, and this is the way the children acquire the pragmatic aspects of the language. In another study, Kyuchukov (2011) shows how Roma children learn Romani from lullabies in Romani. For many parents, teaching Romani as a mother tongue in their household is carried out through the oral history – songs, fairytales, language jokes and games, and teasing (Kyuchukov, 2009). Roma children are thus richly engaged in language learning though their home experiences, though these may be at variance with the style of interaction and language learning encountered by other European children (Teasley, 2013). Heath (1975) documented divergent language learning experiences of different races and classes in the US. In her ethnographic work, the style of learning and the language skills learned by low-income African American children were mismatched with the linguistic and communicative demands of mainstream US schools (Hoff, 2013).

Recent studies on language acquisition have shown that Roma children follow the path of any other typically developing children. Between 3-6 years of age they acquire theory of mind and mental state verbs and their complements,
as well as evidentiality markers (Kyuchukov & de Villiers, 2009). A study in Bulgaria, Czech Republic, and Croatia showed that regardless of the dialect they speak, Roma children know different grammatical categories in their mother tongue. Between 3-6 years they acquire categories such as: wh- questions, wh- complements, passive verbs, possessives, tense, aspect, as well as the ability to “fast-map” novel nouns and adjectives from their linguistic context (Kyuchukov & de Villiers, 2014a; Kyuchukov & de Villiers, 2014b). Although the children have good knowledge of their mother tongue, this skill is not seen as an asset when it comes to school placement in several European countries.

It is common in many European countries (such as Bulgaria, Czech Republic, Slovakia, Poland and Hungary) to test Roma children in the official language of the country to make decisions about whether the children are sent to mainstream classes or “special schools.” These “special schools” in Europe are actually schools for children with mental disabilities. After WWII in 1950-1960, Vygotsky's 1929 work on "defectology" spread around Eastern Europe (Rieber & Carton, 1993). Unfortunately, the lack of knowledge by children of the official language of the country is still considered to be a “defect” in the development of the child. Roma children who do not attend kindergartens before attending primary schools and do not speak any of the official languages of the countries they live in are too often considered to be cognitively deficient. Since the testing, even if it is not language testing, is done in the official language of the country, the children might fail on the tests because they have not yet mastered that language, but nevertheless they are sent to special schools. The curriculum in these schools is well below par compared to the curriculum in regular schools. There is often training in practical trades that are not appropriate for modern life, such as shoe repairing (Simeonova & Kurkova, 2002). When they finish their education at the special schools, the children cannot continue in higher educational levels and risk being stigmatized for the rest of their lives as “mentally retarded.” They have a low likelihood of finding a job and being integrated into the society. Thousands of children in East European countries have such a destiny. Although some countries were ordered to close the “special schools” by the European court for Human Rights in Strasbourg (such as Czech Republic, Slovakia and Hungary), the special schools still exist and testing in the official languages continues to take place (New & Merry, 2014). New (2014) brings up some of the arguments and prejudices of the majority societies such as “Roma do not have a real language” and contests them, in order to advocate for the human right of Roma children to have education in their mother tongue.

**Is Testing Roma Children Important and to Whom?**

In the last two decades, international NGOs such as the Open Society Institute have brought attention to the issue of Roma children’s access to quality education because the number of Roma children in segregated and special schools has
greatly increased. In some countries such as in Slovakia, for example, up to 70% of Roma children are placed in special schools (Rona & Lee, 2001).

According to Rorke and Wilkens (2006):

“So long as governments deny Romani children equal access to a quality education, any talk of integration is empty. Too many Romani children today are classified as “mentally handicapped” and relegated to “special schools”; too many other Romani children attend substandard “Gypsy schools” situated in Romani ghettos; and far too many others still are simply allowed by teachers and administrators to drift away from the classroom ill-prepared to manage lives fraught with problems. These are all segregation by different names; and so long as segregation is allowed to continue, it will have disastrous effects on Romani communities. The countries participating in the Decade of Roma Inclusion have endorsed its education goals. In order to raise the academic achievement of Romani children to the same levels as those attained by their peers outside the Romani community, these goals stress full and equal access to schools and the need to focus on keeping Romani pupils in school until they complete vocational training or receive high school diplomas.”

A study on the desegregation of Roma schools in Bulgaria demonstrated that Roma children attending segregated schools in Roma ghettos have worse educational outcomes than Roma children who study in mainstream schools (Kyuchukov, 2006). These negative outcomes were observed when Roma children’s reading comprehension and writing skills were tested in Bulgarian and when the children did not receive support from Romani teaching assistants in the classroom. Hence, the author recommended that the segregated schools be closed down and that Romani teachers be introduced in the classrooms in order to help Roma children with the learning process. Another recommendation by the author is that in the schools predominantly attended by Roma children, the use of Romani as the mother tongue should be acknowledged and respected.

Of course, another factor for Roma children’s school attendance and success is considered to be the poverty Roma experience. Studies in Slovakia (Rusnakova et al., 2015) and in Romania (Sandru, 2015) show how poverty and societal exclusion are obstacles for Roma children to have access to the educational system.

However, recognition, respect and support of the native language are important for children to succeed in school. A study with Roma children from the Czech Republic and Slovakia in UK public schools (Fremlova & Ureche, 2011) demonstrates the negative impact of labeling children and segregating them from mainstream education programs to place them in less challenging programs. In this study, Roma children who were initially diagnosed with cognitive deficits in their home countries and placed in special schools were later placed in mainstream schools alongside English children when their families
relocated to the UK. After spending a year in regular English school system, these Roma children were studying successfully alongside native English children. Another study with Roma children from special schools in Bulgaria, the Czech Republic and Slovakia (Kyuchukov, 2008a) showed that Roma first graders in special schools that are provided support by Roma teaching assistants in the form of intensive bilingual after-school activities achieve the same level of skill as non-Roma children from mainstream schools after one year of education. These findings demonstrate the importance of respecting children’s native language (especially if different from the language of schooling) and of providing support to help children transition to schooling in a new language.

A conference organized by the Council of Europe in 2008 in Bratislava discussed educational issues, particularly those faced by Roma communities and families regarding the role of their mother tongue in the educational process of Roma children (Kyuchukov, 2008b). One of the recommendations of the Conference was that Roma children be tested in their mother tongue before being placed in special schools. However, in many European countries – not only East European but also in West European countries, e.g. Germany (Greiner, 2014) – special schools still exist and Roma children are placed there because they do not have enough mastery of the official language. As previously discussed, Romani-speaking children who are placed in regular education programs can succeed if given appropriate support in the classroom while their acquisition of the official language comes with exposure to the new language in school. Therefore, children should not be placed in special programs unless they have a real language or intellectual difficulty, which we will later show can only be accurately determined by assessment in their native language.

The justifications of the teachers, psychologists and policy makers for placing Roma children in special programs include the arguments that there is no a standardized/unified Romani language, that there are no tests in Romani, and that there are no Roma psychologists to test the children. These are not valid reasons to keep thousands of children from having access to quality education and to limit their opportunities to succeed in life.

An additional complicating factor is that in many European countries the parents of the children who are in special schools get some financial support for them, such as free meals and free textbooks, and money for tickets for public transport for the children. Living in extreme poverty and not being informed about the consequences for the children who attend the special schools, some Roma parents prefer their children to study in special schools, because they think their children will receive these free benefits (Grzymała-Moszczyńska et al., 2011). The negative consequences are faced later in life when the children who studied in special schools try to find jobs, and no one will accept them because it is clear what kind of school they attended from their school certificate, and with such a certificate they cannot continue in any secondary or high school. The certificates from special schools do not have the status of a diploma.
Even if the educational practices cannot be changed in the short term, we argue that minimally, Roma children need to be assessed in their mother tongue of Romani as well as the mainstream language of the country. Not only is Romani widely acknowledged as an independent language, but also dialectal variations need not be a barrier to testing. There are cases in Europe in which dialects are used not only for testing children’s language skills but also for teaching in schools. Switzerland is the best example, because the German, Italian and French spoken in Switzerland actually are dialects of the three mentioned languages and are used in schools. In the United States, African-American English is another example where dialect/variety is considered when testing in order to have a correct picture of the skills of the children from poor, disadvantaged communities. Without such consideration, the low socio-economic and societal status of the families are the basis for prejudices towards the children and are projected on the children’s performances on language tests (Seymour, Roeper, & de Villiers, 2005).

Testing Roma children in Romani is very important for them, for their parents, and for Roma communities as a whole. The development of tests in Romani is important in order to make the children feel comfortable with the language of testing. And, most importantly, Roma children’s language ability can be accurately captured with a language assessment in their native language. Testing Roma children in Romani will also provide educators with information they can use to decide how to place Roma children in schools. Ultimately, the assessment of Roma children in Romani will help avoid their incorrect placement in special schools and to eradicate the prejudice prevalent in Europe that Roma children are “mentally retarded.”

**Research on bilingual children's language development**

Assessing children’s early language skills is essential to detect potential delays, to intervene in a timely manner, and to ensure that children have the strong linguistic skills needed to succeed in school. Though children do not begin formal literacy education until first grade, the building blocks of literacy undergo development years prior to school entry (Scarborough, 2001). Children’s language skills at school entry predict linguistic and academic performance in grade school (Catts, Fey, Tomblin, & Zhang, 2002; Duncan et al., 2007; Goswami, 2001; National Institute for Literacy, 2010; NICHD Early Child Care Research Network, 2005). Early assessment and intervention for children with language problems are incredibly important because if language-delayed children are left unsupported, their language difficulties usually persist over time (Baydar, Brooks-Gunn, & Furserberg, 1993; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996).

Assessing bilingual children is complex because it entails capturing uneven linguistic knowledge in the two languages. Bilinguals use their languages in different contexts and with different interlocutors, and thus their knowledge
in each language reflects those contextual differences. For example, research has shown that the vocabulary known in each language is mostly unique to each language rather than known in both languages, especially in the early stages (Pearson, Fernández, & Oller, 1993; Peña, Bedore, & Zlatic-Giunta, 2002). This pattern of uneven knowledge in each language is called distributed knowledge.

A consequence of distributed knowledge in bilinguals is relatively slower growth in each language but comparable overall knowledge. While bilinguals know fewer words in each of their two languages relative to monolinguals of each of their languages (Bialystok & Luk, 2012; Bialystok, Luk, Peets, & Yang, 2010), their overall vocabulary is equivalent to that of monolinguals (Holowka, Brosseau-Lapré, & Petitto, 2002; Patterson, 1998; Pearson et al., 1993). This pattern of relatively lower competency in each language compared to monolinguals can also be seen in other areas of language (Fiolas & Peña, 2004; Gutiérrez-Clellen, 2002; Ivanova & Costa, 2008; Oller, Pearson, & Cobo-Lewis, 2007). Because learning two languages entails slower language development in each language relative to monolinguals, it is important to look for language skill across languages when assessing bilinguals, rather than to assess only one language to determine if a child has language problems. Indeed, researchers have adopted different methods of counting bilingual vocabulary (e.g., conceptual vocabulary, total vocabulary) that give children credit for knowing a concept in one language or the other (Oller et al., 2007; Pearson et al., 1993) or add together the two language vocabularies (Core, Hoff, Rumiche, & Señor, 2013). Thus, capturing bilinguals’ overall linguistic knowledge requires the assessment of both languages.

Second language proficiency alone is not an accurate indicator of language problems. Assessing Romani is necessary to detect language problems in Roma bilingual children because genuine language delays manifest themselves in the two languages (Paradis, 2010). Assessing the second language alone confounds language experience with true linguistic knowledge. For example, if children are only assessed in the language they have recently become exposed to, those children will have gaps in their second language that will seem to be signs of language delay. However, the child may really be at the early stages of learning the second language and merely need more exposure in that language to reach a normal level of language skill for their age. To distinguish limited language experience from a genuine language problem, it is necessary to look for delays in the two languages (Kohnert, 2010). Assessing both languages will reveal if children have difficulties in their first language, a phenomenon that is indicative of Specific Language Impairment (SLI) in bilinguals (Restrepo & Kruth, 2000; Salameh, Håkansson, & Nettlebladt, 2004). Dual assessments have already been achieved and normed for Spanish/English bilinguals in the US (Iglesias, de Villiers, Golinkoff, Hirsh-Pasek, & Wilson, in press; Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2014). For more than 20 other languages, Armon-Lotem, de Jong, and Meir (2015) describe tools for assessing language
in the light of the children's bilingualism. The process is beginning in Europe, but norms have not yet been developed.

Taken altogether, these findings show that assessing second language skills alone is not sufficient to determine if a bilingual child has language problems and underscores the importance of assessing Romani-speaking children in their home language.

Many bilingual children acquire their two languages sequentially, and therefore their early language skills are developed in their L1 well before they encounter their L2 in school. It has been proposed that for bilingual children to develop the level of language skill needed to succeed academically in their L2, they must have strong L1 skills as a prerequisite (Cummins, 1979). Indeed, a large body of research shows that there are many ways in which first language skills support second language acquisition (for a review see August & Shanahan, 2006). Here we review a robust set of findings that illustrates the importance of first language skills in second language acquisition.

**Phonological awareness** or the awareness that words are made up of smaller sound units, is a skill that supports children’s literacy acquisition (Cisero & Roger, 1995). Research on L2 learners of English with a variety of home languages (e.g. Spanish, Arabic, Mandarin, Italian, Hebrew, Cantonese) has shown that phonological awareness in the first language predicts literacy skills in L2 when taking second language skills into account (Cisero et al., 1995; Chen, Xu, Nguuyen, Hong, & Wang, 2010; D’Angiulli, Siegel, & Serra, 2001; Dickinson, McCabe, Clark-Chiarelli, & Wolf, 2004; Durgunoğlu, Nagy, & Hancin-Bhatt, 1993; Gottardo, Chiapepe, Yan, Siegel, & Gu, 2006; Gottardo, Yan, Siegel, & Wade-Woolley, 2001; Manis, Lindsey, & Bailey, 2004; Quiroga, Lemos-Britton, Mostafapour, Abbott, & Berninger, 2002; Wang, Perfetti, & Liu, 2005; Wang, Yang, & Cheng, 2009). These studies show that strong phonological skills in the first language contribute to second language learning, even if the two languages are very different from each other.

At the morphological level, there is also ample evidence that first language skills support L2 literacy development. **Morphological awareness**, the ability to perceive and to manipulate the morphemic structure of words (Carlisle, 1995), is an ability that contributes to reading comprehension (Kuo & Anderson, 2006; Nagy, Berninger, & Abbott, 2006; Nagy, Berninger, Abbott, Vaughan, & Vermeulen, 2003). Having solid morphological awareness allows readers to decipher the meanings of unfamiliar words more by examining morpheme constituents (Nagy & Anderson, 1984). Research on bilingual children with a variety of home languages (e.g. Spanish, Chinese, Arabic) learning to read and write in English (L2) has shown that bilinguals’ awareness of different types of morphology (compound, derivation, inflection) in their home language contributes to their reading skill in L2 when controlling for L2 language skills (Lam, Chen, Geva, Luo, & Li, 2012; Ramirez, Chen, & Pasquarella, 2013; Saiegh-Haddad & Geva, 2008; Wang et al., 2009; Zhang & Koda, 2014).
L1 morphological awareness is an aspect of the first language that contributes to the development of L2, and its transfer from L1 to L2 underscores the importance of assessing the L1 skills of bilingual children, since those skills will support subsequent L2 acquisition.

In summary, though Roma children may have been traditionally referred to special education programs on the basis of their second language proficiency, the previously cited research demonstrates that the assessment of the second language alone is inaccurate in detecting language impairment (Clark & Kamhi, 2010; Paradis, 2010). By assessing both languages, educators will be able to accurately determine if children have a language problem, which will likely manifest itself in the two languages. Moreover, the assessment of bilingual children’s Romani skills should not be overlooked because first language skills contribute to second language outcomes (August & Shanahan, 2006). Bilingual children’s proficiency in their native language of Romani should thus be regarded as an asset for second language learning and academic success, not a liability. It is therefore imperative to test Roma children’s first language skills to ensure that schools are able to make informed decisions about how best to serve Roma children.

Our Romani Test: ROMČHIT (Romano Čhibjako Testo) - ROMLAT (Romani Language Assessment Test)

Considerations in Test Development

When designing a psycholinguistic assessment in a language like Romani, for which no language assessments exist, it is important to consider several approaches (de Villiers & de Villiers, 2010). The first approach might be to look at standardized language tests from other languages, and to choose the kinds of items that are well represented on those tests. However, translating items from other language tests is not ideal because it is quite likely that the structures and even what gets lexicalized are different in the other language, and what might be trivial in one language may be very difficult, more elaborated, or even absent in the other (Peña, 2007). A second approach might be to look carefully at some linguistic domains (e.g. syntax, lexicon etc.) in the untried language and to use the research literature about language acquisition and methodology to generate appropriate subtasks for those structures (See Armon-Lotem, de Jong, & Meir, 2015). We opted for the second approach and tested these subtasks on the target age group to ensure that these new subtasks capture language skill.

We took several factors into account to ensure that we developed appropriate measures of Romani for the target population. First, we did not translate items from other language tests. Second, we considered the variability within Romani dialects. Third, we designed subtasks that reflect properties important to communicative functions related to early school success. Lastly, we chose skills and knowledge that are likely to be under mastery by age 6.
The resulting subtasks are described briefly below, together with the justifications for their inclusion.

The test is composed of nine subtasks (Table 1). In line with recent research on assessment for language impairment, we considered measuring not only acquired vocabulary and grammatical structures (language products), but also the processes of learning (language processes) (Hirsh-Pasek, Kochanoff, Newcombe, & de Villiers, 2005; Kyuchukov & de Villiers, 2014). These processes include the inferences by which a child identifies the meaning of a new word in context, and the use of a grammatical morpheme or structure with a novel lexical item. These kinds of tests reveal whether the child has appropriately generalized the language skills to allow creative use and growth.

The possessive and tense subtasks were based on the “wug” test developed by Jean Berko Gleason (1958). The wug test was designed to see if a child can apply a rule of the language, such as plural morphology, to a new lexical item: “Here is a wug, here is another one, there are two … (wugs)” using a cloze-test. This research revealed that the plural forms are not learned as new words, but as a productive rule. In the Romani subtasks, we use this kind of procedure to elicit possessive forms from the children, which together reflect a complex matrix wherein the gender and plurality of both subject and object influence the morphology of the possessive. This complexity is rather unique and impressive. Using the same principle, the tense subtask requires children to apply regular tense to a novel verb.

The fast mapping nouns, fast mapping adjectives, and aspect subtasks were inspired by research on children’s use of linguistic context in word learning (Alishahi, Fazly, & Stevenson, 2008; Heiback & Markman, 1987). This research stressed that children do not learn the majority of their words by ostension, i.e. by having the referent pointed out, but instead, made inferences based on linguistic and nonlinguistic context in the course of everyday conversation. In fact, there is no alternative for words that have no obvious concrete referent. Hence grammatical context is used to figure out word meaning, and a child’s ability to learn a new word is tied to their grammar skills and inferential ability (Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005). Assessing known words can be fraught because words vary with experience; the playing field can be leveled if all children are taught a new word (de Villiers & Johnson, 2007). Hence, the tests reveal whether a child can learn new nouns and adjectives by exploiting the morphological cues (such as gender) in Romani. In a test of aspect morphology, the child had to pick out which picture depicted an ongoing versus completed novel action, a stringent test to see if the aspect morpheme was properly analyzed in terms of its meaning. If the child can transfer the morphology to a newly introduced novel verb, then the meaning it carries has been understood.

The remaining subtasks reflect general principles that early grammar must reflect. For example, all languages have movement rules, in which there are displacements of items from their standard order in a sentence, as in questions in
many languages (Roeper & de Villiers, 2011; Sauerland et al., 2016; de Villiers, Roeper, Bland-Stewart, & Pearson, 2008), or passive sentences (Armon-Lotem et al., 2016; Deen, 2011). Hence, we chose sentences in Romani that reflect these structures.

For the *Wh-complements* subtask, we chose sentences with two clauses. An important progression in the preschool years is to learn embedded forms, i.e. recursive structures in which one sentence is embedded in another. These come in two varieties: tensed and non-tensed or infinitival. Children can be tested by asking questions to see if they can remember the content of an embedded clause that does not coincide with reality (de Villiers & Pyers, 2002). In English, it has been found that children have a much easier time with infinitives than with tensed clauses (de Villiers, Harrington, Gadilauskas, & Roeper, 2012). Mastery of these types of clauses is also linked with children’s ability to succeed at other tasks involving understanding others’ minds, so-called false belief reasoning, which is an important mastery in the area of social cognition prior to school (de Villiers & Pyers, 2002; de Villiers, 2005; de Villiers, 2007).

The *Wh-questions* subtask entails more of the logic of grammar, and we chose it because it has been used successfully across many European languages to reveal deficits in grammatical skill (Schulz & Roeper, 2011; Schulz, 2015). The child is asked a very simple question such as “who ate what?”, but the answer requires pairing two sets with one another: the set of subjects must be lined up with the set of objects, and any lesser answer is inadequate.

The *passives* subtask was designed to reflect the finding that children at younger ages and lesser ability reverse the argument roles and treat a passive sentence as if it were active. It also tests a difference that was consistently found in previous research, i.e. that children have more difficulty with passives if the verbs being passivized are about states rather than actions, such as “seeing” or “linking” (Maratsos, Fox, Becker, & Chalkley, 1985).

Finally, we have a complex *sentence repetition* subtask. Other studies in different languages have shown the value of asking the child to repeat sentences that tax their grammar skills, because children can best remember sentences that they have the capacity to produce spontaneously (Lust, Flynn, & Foley, 1996). Moreover, deficits in sentence repetition have been shown to be good indicators of language impairment (Conti-Ramsden, Botting, & Faragher, 2003; Klem et al., 2015; Stokes, Wong, Fletcher, & Leonard, 2006). Using sentence repetition, unlike the other specific subtasks, allows more general assessment of a variety of sentence forms of increasing complexity that may be less easily captured for a comprehension or elicitation task. The sentence repetition types and their English glosses are provided in Table 1.
Table 1. Sentence Repetition Types and English Gloss

<table>
<thead>
<tr>
<th>Sentence repetition type</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intransitive</td>
<td>The baby played</td>
</tr>
<tr>
<td>Imperative</td>
<td>Ride the horse</td>
</tr>
<tr>
<td>Intransitive with locative</td>
<td>The boy goes to the market</td>
</tr>
<tr>
<td>V S O</td>
<td>Sang Maria at school</td>
</tr>
<tr>
<td>S O V</td>
<td>The boy the chocolate eats</td>
</tr>
<tr>
<td>VP conjunction</td>
<td>The father saw the child and gave him a chocolate</td>
</tr>
<tr>
<td>Temporal adjunct clause</td>
<td>Before to read the book the boy at the chocolate</td>
</tr>
<tr>
<td>Temporal adjunct clause</td>
<td>After the school the boy played football with another boy</td>
</tr>
<tr>
<td>VP conjunction and relative clause with complement</td>
<td>The woman saw the boy and gave him a cat who knows how to sing</td>
</tr>
<tr>
<td>Subject relative clause</td>
<td>The boy who found the ant in the forest lays golden eggs</td>
</tr>
</tbody>
</table>

Note: Romani allows several word orders. In the sentence repetition subtask we gave the children the following types of sentence, chosen to increase in length and syntactic complexity.

Table 2 shows the subtasks on which we have tested 150 children in several countries: Bulgaria, Macedonia, Serbia, Croatia, and Sweden. The children – between the ages of 3-6 years old – spoke Romani as their dominant language and at the same time attended kindergartens and were learning the official language of the country they lived in. The results are reported in several recent papers and conference posters (Kyuchukov, 2014; Kyuchukov & de Villiers, 2014a; Kyuchukov & de Villiers, 2014b). There are interesting differences so far depending on the place/role of Romani and its quality in the parents’ lives. For example, in the Czech Republic the 30 children tested there showed weak control of the possessive morpheme paradigm, even though their average age was older than the 30 younger children tested in Bulgaria. The latter group for whom Romani was a strong L1 showed significant mastery (over 80%) on all subtasks except the Aspect test by the age of five years (Kyuchukov & de Villiers, 2014b).

Table 2. Components and Subtasks of Romani Test

<table>
<thead>
<tr>
<th>Test Components</th>
<th>Subtask</th>
<th>Number of Items</th>
<th>Examples in Romani</th>
<th>English Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Product</td>
<td>Wh Questions</td>
<td>8</td>
<td>Kon jkaj sovel?</td>
<td>Who where sleeps?</td>
</tr>
<tr>
<td></td>
<td>Wh Complements</td>
<td>8</td>
<td>O čhavoro phendas e dajake te kinel šokoladi, aj oj kindas sladoledi. So phendas o čhavoro pe dajake te kinel?</td>
<td>The boy told to his mother to buy him a chocolate, but she did buy ice cream. What did the boy say to his mother to buy?</td>
</tr>
<tr>
<td>Test Components</td>
<td>Subtask</td>
<td>Number of Items</td>
<td>Examples in Romani</td>
<td>English Gloss</td>
</tr>
<tr>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------------------------------------------------</td>
<td>----------------------------------------------------</td>
</tr>
<tr>
<td>Passive</td>
<td>O grast sas čalavdo e džuklestar</td>
<td>The horse was kicked from the dog.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentence Repetition</td>
<td>O čhavo i šokolata xal</td>
<td>The boy the chocolate eats.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Process</td>
<td>Possessive</td>
<td>26</td>
<td>Akaja si i mačka. Si la jekh cita. Akaja si .......(i mačkakeri cita)</td>
<td>This is the cat. Shi has a cita (novel object). Ths is …...(the cat’s cita)</td>
</tr>
<tr>
<td>Aspect</td>
<td>I daj sar kerla zumi, i čhajori travinela pe papuske bala.</td>
<td>While the mother cooks the soup, the girl (incomplete novel verb) the hair of the grandfather.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tense</td>
<td>Kada murš džanel te zitivinel. Tehara pale kan kerel akes. Tehara ov ....(ka zitivinel)</td>
<td>This men knows how to (novel verb). He is going to do that tomorrow again. Tomorrow he …....(the novel verb in future tense)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Mapping</td>
<td>Kate dikhes parni blepi. Sikav mange i parni blepi</td>
<td>Where do you see (novel noun)? Show me the (novel noun).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Noun</td>
<td>Dikh so si amen kate: her , mačka thaj džukel. Kaja boja si patravalí Kaj si o patravalo.?</td>
<td>Look what we have here: donkey, cat and the dog. This color is (novel adjective). Where is the (novel adjective with masculine ending)?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Here we report similar new results from Romani speaking children in Slovakia. All children received the full set of ROMČHIT-ROMLAT (Romani Language Assessment Test) subtasks, as well as two cognitive tests: forward digit span and two standard tasks making up a Theory of Mind test. These tasks
were chosen to tap memory and social/cognitive reasoning skills, and we expect both to correlate with language. Digit span is an index of verbal memory, likely to relate to sentence repetition, though sentence repetition involves the additional factor of morphosyntax understanding. Theory of Mind, especially understanding another’s false belief, has been found to be linked to language skills, whether the measure is general vocabulary (Astington & Baird, 2005), general syntax (Farrar, Benigno, Tompkins & Gage, 2017), or specifically sentence complementation (San Juan & Astington, 2012; de Villiers & de Villiers, 2013).

The digit span test was presented by asking the child to repeat a succession of strings of digits, two at each length until the child failed both of that length. The longest string at which the child got both examples correct was then considered their digit span (Ostrosky-Solis & Lozano, 2006).

The Theory of Mind tasks were the Unseen Displacement task (2 examples) and the Unexpected Contents task (asking about own previous and other's beliefs), both widely used (Wellman, Cross, & Watson, 1990) across the world and in our own previous work with Roma and Turkish children (Kyuchukov & de Villiers, 2009). Each task is worth two points for a total of 4.

Participants

29 Romani-speaking children were tested for this study. All were residents in a very poor part of the town of Spišska Nova Ves in Slovakia, a Roma settlement of extreme poverty. The living conditions were very poor even for Roma: the homes had no running water and no inside toilets. The children all attended a kindergarten with very little in the way of toys or material supplies.

<table>
<thead>
<tr>
<th>Table 3. Participants by Age Group and Gender</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age range 3;5-3;11</td>
</tr>
<tr>
<td>Age range 4;2-4;11</td>
</tr>
<tr>
<td>Age range 5;2-6;1</td>
</tr>
</tbody>
</table>

Method

The children were tested by the first author and a Roma teaching assistant working in the kindergarten, in the Romani dialect the children speak in this part of Slovakia. The testing was done in a separate room alone with the researchers. Before testing, the parents of the children were informed about the research and written permission was obtained from them.

The children were shown colorful pictures created by the second author. The answers of the children were written in protocols, which were later coded and statistically analyzed.
Results

The first question is whether the subtasks all showed improvement over age, as any language assessment needs to index developmental change. The results of a multivariate ANOVA confirm that all subtasks showed statistically significant change \((p < .001)\) with age group, with the exception of fast mapping nouns (see Table 4). That subtask showed less change because the 3-year olds were already highly proficient (80%), so the change between age 3 and 4 versus 5 years was only at level of \(p = .09\) in post-hoc analysis.

Table 4. Multivariate ANOVA on Subtask Performance by Age Group

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent Variable</th>
<th>df</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age group</td>
<td>wh</td>
<td>2</td>
<td>6.169</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>complements</td>
<td>2</td>
<td>10.412</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>passives</td>
<td>2</td>
<td>36.177</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>sentence repetition</td>
<td>2</td>
<td>10.817</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>possessives</td>
<td>2</td>
<td>44.714</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>tense</td>
<td>2</td>
<td>29.031</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>aspect</td>
<td>2</td>
<td>21.470</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>fastmapN</td>
<td>2</td>
<td>2.208</td>
<td>.130</td>
</tr>
<tr>
<td></td>
<td>fastmapadj</td>
<td>2</td>
<td>11.336</td>
<td>.000</td>
</tr>
</tbody>
</table>

Table 5. Comparison of probability of children's performance against conservative estimates of chance

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chance estimate:</td>
<td>33.3</td>
<td>50</td>
<td>33.3</td>
<td>33.3</td>
<td>50</td>
<td>25</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>3 year olds</td>
<td>.001</td>
<td>.035</td>
<td>.001</td>
<td>.001</td>
<td>.397</td>
<td>.195</td>
<td>.003</td>
<td>.169</td>
<td></td>
</tr>
<tr>
<td>4 year olds</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>5 year olds</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
<td>.002</td>
<td>.001</td>
<td>.001</td>
<td>.001</td>
</tr>
</tbody>
</table>

Secondly, one can ask how well the different items intercorrelate, i.e. are they tapping a common underlying construct? The usual test of this is Cronbach's alpha, which had a highly satisfactory value of .85 across 82 items. Total percentages correct on the subtasks themselves intercorrelated very substantially (see Table 5). Only some pairs did not reach significance, and close inspection suggests they may represent differently developing skills. For example, exhaustive
answers to paired wh questions failed to correlate with fast mapping nouns, and except for both being language skills, it is hard to make a theoretical argument for their relatedness. Of course, the lack of a significant correlation could also be an artifact of a small sample.

The language indices show substantial skills in these children, even at age 3, with all indices being well above chance. For comprehension indices, chance can be estimated by the number of explicit choices provided. Of course, chance is harder to estimate for production. On a highly conservative guess, chance could be considered 50%, i.e. the children could either mark or not mark a distinction. For example, there are 16 possible inflections in the possessive test among which to choose, and the three-year-olds were almost 80% correct. We estimated chance at 50% except for the wh task, where to get scored as correct the children had to use both exhaustive subjects (50%) and exhaustive objects (50%), so the chance of a complete answer would be 25%. Table 6 shows the results of one-sample t-tests against chance level for each subtask at each age. The performance of 3-year olds exceeded chance for 6 of the 9 subtasks, with all the "failures" in production, with its highly conservative estimates of chance.

### Table 6. Intercorrelations Among Language Subtasks on ROMLAT Pearson Correlations

<table>
<thead>
<tr>
<th></th>
<th>Wh</th>
<th>Comps</th>
<th>Passives</th>
<th>Sent Rep</th>
<th>Possessive</th>
<th>Tense</th>
<th>Aspect</th>
<th>Fast-Map Noun</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complements</td>
<td>.289</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Passives</td>
<td>.611**</td>
<td>.512**</td>
<td>.560**</td>
<td>.605**</td>
<td>.612**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sent Rep</td>
<td>.291</td>
<td>.495**</td>
<td>.564**</td>
<td>.772**</td>
<td>.706**</td>
<td>.759**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Possessives</td>
<td>.423*</td>
<td>.564**</td>
<td>.805**</td>
<td>.772**</td>
<td>.706**</td>
<td>.759**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tense</td>
<td>.560**</td>
<td>.626**</td>
<td>.772**</td>
<td>.706**</td>
<td>.759**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aspect</td>
<td>.512**</td>
<td>.543**</td>
<td>.656**</td>
<td>.710**</td>
<td>.694**</td>
<td>.818**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fast Map Noun</td>
<td>.245</td>
<td>.357</td>
<td>.432*</td>
<td>.259</td>
<td>.498**</td>
<td>.477**</td>
<td>.396*</td>
<td></td>
</tr>
<tr>
<td>Fast MapAdj</td>
<td>.260</td>
<td>.425*</td>
<td>.612**</td>
<td>.613**</td>
<td>.574**</td>
<td>.562**</td>
<td>.639**</td>
<td>.477**</td>
</tr>
</tbody>
</table>

** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

The digit span results showed sensible growth with age, and at levels compatible with general results from the US (see Table 7) (Chen & Stevenson, 1988).

### Table 7. Forward Digit Span by Age Group

<table>
<thead>
<tr>
<th>Age Group</th>
<th>3 year olds</th>
<th>4 year olds</th>
<th>5 year olds</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>9</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td>3.0</td>
<td>3.7</td>
<td>4.6</td>
</tr>
<tr>
<td>SD</td>
<td>.87</td>
<td>.67</td>
<td>.70</td>
</tr>
</tbody>
</table>
As an index of verbal memory, digit span unsurprisingly correlated highly significantly \((p < .005)\) with all but two language subtasks, slightly less for the fast mapping \((p < .05)\) and not quite for the exhaustive wh test (which was only \(p < .08\)).

The Theory of Mind skill, or more precisely, false belief understanding, showed growth over the age range and also correlated significantly \((p < .02)\) with all of the language subtasks except fast mapping nouns \((p = .06)\), probably because of the ceiling effect for that subtask. The major change was from 3 to 4, after which 4 and 5 year olds behaved equivalently though not particularly well, with the 5 year olds averaging only 2.5 of 4 possible points.

**Discussion**

The digit span results are firmly in the range typical for their age groups (see comparison data from US children in Chen and Stevenson (1988)), giving us confidence that these children are on average typically developing in this respect. But why are their ToM results relatively weak? Wellman et al.'s (2000) review suggests that the average age of passing the standard ToM tasks shows some variance across the world, hovering around 4.5 years, though others have found delays beyond that age even in typically developing children (Tardif, So, & Kaciroti, 2007; Vinden, 1996). It remains possible that the task materials were not ideally matched to the children's normal experience. But more likely, these particular children do not have a rich social experience. They know only the settlement and the kindergarten. At home they have TV but no internet, no access to children's books, nor any toys. In our other work on Romani- and Turkish-speaking Bulgarian children in Bulgaria, also living in settlements but with a higher living standard, the children performed at the same level as in other studies in the West (Kyuchukov & de Villiers, 2009).

We have the situation then that children who are growing up in extreme material and social poverty by European standards are well versed y the preschool years on the complex grammar and vocabulary of their native language, not just in spontaneous conversation but on tightly controlled test materials. These are skills that should not be dismissed.

**Conclusion**

Roma children are in double jeopardy. They are bilingual and assessed in only one language, but because that language is considered deficient and sub-standard, the situation with Romani is not like the situation with say, Spanish-English bilinguals in the US. The situation is akin to the problem of African American children in the US who grow up speaking African American Vernacular English. Yet it is worse, because Romani is not a dialect of the state language but historically remote from European state languages.
We argue that a child who shows considerable skill and versatility with their mother tongue by age 5 years is not likely to be a candidate for placement in “special” education. Of course, Roma children may not have all the needed social skills by this age because of the isolated life in the ghetto type of Roma settlements, where they grow up. They may also be "unschooled" in their preparation, since by the age of 6 when the testing is done, these children may never have seen a book or a pen or pencil. In deciding on placement, the children may be given other cognitive or social tests, but never via their home language. Usually the Roma children are tested in the official language of the country and very often these tests have been translated from English or German, so they cannot be considered "culture free" tests.

Testing in Romani as a mother tongue is very important for Roma children in Europe. If the children know the complex grammatical categories in their mother tongue by the age 5, that shows that the children are probably typically developing and do not have any form of cognitive deficit. Their placement in special schools is a function of government attitudes to Roma, not the needs of the children themselves.

The implications must be faced: schools need to accommodate the needs of children who speak minority languages and who are still learning the language of the state. And respect for Romani as a language is necessary so that the linguistic strengths children bring to school are not swept to the gutter. But linguistic prejudice is fed by deeper cultural prejudice. In 1979, James Baldwin wrote an essay entitled: "If Black English Isn't a Language, Then Tell Me, What Is?" To quote this essay:

"A child cannot be taught by anyone who despises him, and a child cannot afford to be fooled. A child cannot be taught by anyone whose demand, essentially, is that the child repudiate his experience, and all that gives him sustenance, and enter a limbo..."

The preliminary results clearly show that testing Romani as an L1 with the ROMČHIT (Romano Čhibjako Testo) - ROMLAT (Romani Language Assessment Test) is feasible and revealing of children's existing strengths. A child with low proficiency in L2 who cannot do these tasks in the L1 is in need of further testing and special support. However, a child who shows competence in these elaborate linguistic abilities in their L1 does not have an intrinsic language deficit.
References


Baldwin, J. (1979, July 29). If black English isn't a language, then tell me, what is? *New York Times*.


Kyuchukov, H. (2011) Роль колыбельных песен в овладении родным языком на примере циганской культуры. [The role of the sleeping songs in Romani language in the process of socialization of Roma children]. Sovremenyi detskiy sad [Contemporary kindergarten], 7, 50–53.


