Microstructural properties in the narrative retellings of young English learners in EMI schools in India: the role of L2 literacy, minority languages and English input in the classroom

Microstructural properties in the narrative retellings of young English learners in EMI schools in India: the role of L2 literacy, minority languages and English input in the classroom

Ianthi M. Tsimpli[1], Lina Mukhopadhyay[2], Anusha Balasubramanian[1] & Jeanine Treffers-Daller[3]

[1] Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK, [2]Department of Training and Development, EFL University, Hyderabad, Telangana, India, and [3]Department of English Language and Applied Linguistics, University of Reading, Berkshire, UK

Address for correspondence: Ianthi Maria Tsimpli, Theoretical and Applied Linguistics, University of Cambridge, Cambridge, UK. Email: imt20@cam.ac.uk

Abstract

The present study investigates narrative microstructure in the English oral retellings of primary school children from disadvantaged backgrounds in India. For these children, the combination of rich linguistic diversity and English as the medium of instruction is a challenge since exposure to English is almost exclusively in the school context. Measures of microstructure include syntactic complexity and finiteness marking as well-known indices of English language ability. Microstructural properties are examined in relation to literacy performance in order to identify possible links between the two modalities. Participants' language background was included to investigate a possible (dis)advantage associated with minority language speakers exposed to English as the medium of instruction. Our findings suggest that finiteness errors and number of function words are associated with performance on reading comprehension. English input in the classroom selectively predicts performance on complex syntax but not other aspects of microstructure measures. Children speaking minority languages at home are not disadvantaged in their English performance compared to children speaking the majority language (Telugu). Finally, error analysis suggests similarities between types of errors found in other studies of child L2 English. This study sheds light on English L2 narrative skills in a multilingual and underprivileged context with learners exposed to low levels of English language input.

Keywords: narrative retelling, microstructure skills, multilingual children, literacy, oral language input, minority language, English medium instruction (EMI)

0. Introduction

Multilingual children worldwide often receive education in settings where the medium of instruction (MoI) is not their first language (L1) or one of the languages spoken at home (Brutt-Griffler, 2017). In this context, learning the language of education often begins in school and in the print form of the language. In these educational settings, learners do not enjoy the linguistic advantages of oracy skills that monolingual or bilingual children instructed in a familiar language have, nor the positive repercussions that familiar languages used as MoI have on the development of literacy skills (Gathercole, 2013; Kapalkova et al., 2016; Ouane & Glanz, 2010; Romaine, 2013). Large cohorts of multilingual learners in Asia and Africa coming from low SES background and minimal home literacy support learn in school settings where the MoI is unfamiliar to them. This is either because it is a regional language not used in the children's homes or the school's official language is English as the medium of instruction. For these children, there is a dearth of literary practices at home coupled with economic challenges such as lack of nutrition and home care (Erling et al., 2016). In such contexts, the school and by extension the MoI play a crucial role in children's educational development. Lack of threshold level competence in the MoI becomes a linguistic barrier instead of a scaffold to learn through the language (Ucelli & Paez, 2007; Marphatia, Reid & Yasnik, 2019).

In spite of these adversities affecting multilingual learners from low SES, a highly diverse and rich multilingual society like India confers certain advantages to learners' cognitive skills (Tsimpli et al., 2020a). This is particularly true of the urban poor learners living and studying in slum and non-slum areas. Sociolinguistic diversity and the inevitable reality of multilingualism for the average individual in India can have a positive impact on the resilience children develop because they are often compelled to negotiate complex circumstances to earn a livelihood (Tsimpli et al., 2020b). This typically translates into an ability to learn and use new languages in school, home and communities relatively effortlessly and as a matter of shared habit. Furthermore, children are open to using multilingual resources to bootstrap into the lexis and syntax of any new language(s) they are exposed to. It is then quite natural for

learners growing up in multilingual societies to accommodate and develop knowledge of semantic and grammatical features of new languages in their natural environment. However, a totally unfamiliar language presented only in school contexts as the MoI has been shown to impede progress in school skills (literacy and numeracy) as well as children's language comprehension in oral and written form (Treffers-Daller et al 2022; Marphatia, Reid & Yasnik 2019; NEP 2020: pp. 9-10). When multilingual practices find support in schools it adds to learners' effort and scaffold their early learning experiences (Mohanty 2010; Meganathan 2018; Durairajan 2018).

Learners' proficiency in a given language, L1 or L2, can be assessed in a number of ways. Narratives provide an ecological way of eliciting linguistic data in a relatively naturalistic setting. Oral narratives have been examined in different languages and across different cultures (e.g. Berman & Slobin, 1994). Oral storytelling is a central part of human life and forms a strong tradition in many cultures like India, where storytelling is a source of regular and rich linguistic input, often in a variety of languages, in Indian households (Amritavalli, 2013). To a linguist, oral stories provide a wealth of linguistic data that speak to the narrator's lexical and morphosyntactic skills. Narratives constitute a type of connected discourse which includes utterances presented in a temporal order describing an event or an experience (e.g. Boudreau, 2008). The ability to produce structured narratives develops over time and is linked to later academic outcomes too (e.g., Beals, 2001; Dickinson & McCabe, 2001). Analyzing a story produced by a monolingual or a multilingual child is a valuable resource for the investigation of the child's linguistic abilities (Fiestas & Peña, 2004; Gagarina, 2016; Gutiérrez-Clellen, 2002; Pearson, 2002; Rojas *et al*., 2016; Uccelli & Paéz, 2007). Specifically, narratives provide information on the child's macrostructure and microstructure levels of linguistic and cognitive abilities. Macrostructure refers to the overall organization and content of the story. In particular, the coherence of the story, its episodes and the links between them expressed linguistically through complete goal-attempt-outcome (GAO) sequences, the setting of the story as well as internal state terms depicting the characters' emotions and cognitive states leading to the initiating of new actions and their response to story events (Stein & Glenn, 1979; Iluz-Cohen & Walters, 2012). Microstructure refers to linguistic measures of productivity and complexity (e.g., Baixauli, Colomer, Roselló & Miranda, 2016; Justice *et al*., 2006). Productivity involves the length of the narrative which can be measured in different ways, such as total number of words produced, lexical diversity, i.e. number of different words produced, or by number of clausal or utterance units produced referred to as Communication units or C-

units (Justice *et al*., 2006; Mäkinen *et al*., 2014). Complexity measures refer to the mean length of these units, the frequency and diversity of clauses used, as well as morphological errors (Altman, Armon-Lotem, Fichman & Walters, 2016; Baixauli *et al*., 2016; Justice *et al*., 2006; Mäkinen *et al*., 2014). The present study focuses on the investigation of microstructure properties with special attention paid to morphosyntactic errors in school children' L2 English used as the MoI in India.

Although several recent studies have investigated narratives in monolingual and bilingual children growing up in western societies (Gagarina, Klop, Tsimpli & Walters; 2016; Pearson, 2002; Pearson & de Villiers, 2005; Uccelli & Paz, 2007), documentation of the development of morpho-syntactic features in the L2 English of young multilingual learners in India from low SES background has been hitherto unexplored. In this study, we examine the oral linguistic competence as found in the presence of microstructural elements in narrative retellings of a group of 90 multilingual Indian learners from low SES background. Focusing on government schools in the city of Hyderabad, in the southern part of India, we also examine the nature of classroom input learners receive in English-medium schools in order to investigate the effects that input quantity in the target language, English, has on the use of microstructural properties in the English narrative retellings of the learners. Our study's novelty is twofold. First, it focuses on an under-studied linguistically diverse learner population in the Global South educated in a language unfamiliar to the learners from either home or societal surroundings. Second, it investigates the role of language input in schools with English as MoI and how these factors influence the use of morphosyntactic features in the target school language.

1. Microstructure skills in bilingual children's narratives and the role of input

Narratives are used as an ecological and child-friendly way of eliciting language samples from bilingual children. Apart from the child's age, microstructure measures of bilingual children's narratives reflect the lexical and morpho-syntactic skills that children possess in the tested language. Narrative retellings in particular have been shown to enhance macro- and microstructural properties in bilingual children's narratives (Kunnari, Välimaa, and Laukkanen-Nevala 2016; Otwinowska et al, 2020). Recent studies with bilingual populations, with and without language impairment, have focused on microstructure properties including syntactic complexity measures, lexical diversity, morphological use and accuracy of nouns and

verbs in narrative productions (Gusewski & Rojas, 2017). Unlike macrostructure skills which have been claimed to be transferrable between the languages of bilingual children (Pearson, 2002; Uccelli & Paz, 2007), microstructure is language-specific and sensitive to the vulnerable areas of morphosyntax in a given language, as identified through the delays attested in monolingual neurotypical and atypical language development (Altman et al. 2016; Bohnacker 2016; Iluz-Cohen & Walters, 2012; Rodina 2016). For instance, Altman et al. (2016) compared typically developing 5-year-old English-Hebrew bilinguals to language impaired bilinguals. The study found microstructure differences expressed through MLU, lexical diversity, and error rates between the two languages of the neurotypical bilinguals. Similarly, Rodina (2016) compared the narratives of Russian-Norwegian bilingual 4 and 5-year olds with those of Russian and Norwegian monolinguals using various microstructure measures. Her findings demonstrate that the bilinguals performed worse than the monolinguals in L1 Russian but not in L2 Norwegian, indicating an effect of stronger language input in the L2 than in the L1. This study in particular highlights the role of language input on the language of the narrative and its contribution to the bilingual child's performance on microstructure.

Although the role of input on L2 proficiency development is uncontroversial (e.g., Blom & Paradis, 2015; Jia & Fuse, 2007; Paradis, 2011; Unsworth, 2013) there are only a few studies on narratives that have examined the role of input on L2 production and specifically on microstructure properties (Govindarajan & Paradis 2019; Tsimpli et al 2016). Several studies focusing on individual differences in L2 input have revealed contradictory results regarding the contribution of input quantity with some studies showing input as a significant predictor of morphosyntactic abilities (e.g. Unsworth 2016) while others failing to find a direct link between language exposure and L2 performance (Chondrogianni & Marinis, 2011; Golberg et al., 2008; Paradis, 2011). Such differences in findings may be due to quality of input asymmetries when considering the L2 input that children receive from home (i.e. from parents who are non-proficient L2 speakers). Other studies have also revealed the effects of L2 input quality measured as frequency of shared book reading and interaction with native speakers on the use of verb morphology, complex syntax and vocabulary size (Jia & Fuse, 2007; Paradis, 2011). Summing up, microstructure measures differentiate bilingual and monolingual children in terms of the child's lexical and morpho-syntactic abilities in a given language. Microstructure may also differ across a bilingual's two languages, reflecting variance in the development of each language.

One of the aims of the present study is to investigate the role of L2 input on the use of morphosyntax in English narrative retellings in children from very low SES with no exposure to English at home or in their immediate community. For these children, English in the classroom is the only source of L2 input given that the school's official MoI is English. The choice of MoI reflects an immersion environment in the second language with all school subjects delivered in the English MoI, all textbooks (except for other language subjects, i.e. Telugu in our sample) in English and all assessments in English too. Although this is the official expectation across schools with English as a MoI, the reality of the Indian classroom in which multilingual teachers and children interact where for many English may be a totally unfamiliar language, switching and mixing between languages during lesson delivery is hardly surprising. Indeed, several studies investigating language mixing or translanguaging in India's classrooms confirm that lessons are delivered in multiple languages and that English as the Medium of Instruction (EMI) schools present a higher degree of code-switching and translanguaging (Lightfoot et al., submitted; Pal & Panda, 2020; Tsimpli et al., 2020b). Against this background, the present study investigates the role of language use in the classroom in order to evaluate its potential role in the microstructural aspects of children's narratives.

2. Finiteness in English and Telugu

English marks finiteness morphology through Tense and Agreement albeit agreement marking is restricted to a single form *–s* which fuses present Tense and third person singular agreement. Free forms such as the copula and the auxiliary *be* mark finiteness distinctions too. Bound morphology expressed through the third person singular present *-s* (3SG) and the past tense suffix *–ed* participate in binary agreement [+/-3s] and tense [+/-past] oppositions, respectively. Copula and auxiliary *be* are irregular forms which mark subject-verb agreement as well as tense (e.g. *I am/you are/he-she-it is* or we/you/they *are vs. we/you/they were*). Aspectual morphology is also encoded on English verbs, either periphrastically or on the verb form. Specifically, the suffix *-ing* in combination with the auxiliary BE marks the verb for progressive aspect while perfective past is encoded on the simple past tense form *–ed* with complex tenses (present or past perfect) encoding aspectual and tense distinctions periphrastically with the auxiliary 'have'.

Telugu is a Dravidian language that has some morphological marking for tense and agreement although their co-occurrence is restricted to the indicative present, future-habitual, past positive

forms and the future negative form (Kissock 2014; Murti 1972). Tense without agreement marking appears in a number of forms referred to as 'non-finite' or 'NoAgree' forms. For instance, *tin-tunna:-nu* (eat-Pres-1s) includes both markings whereas *tin-aḍam-le:du* (eat-INF-Neg) uses the infinitival suffix '-aḍam' to produce a present tense 'NoAgree' form which contrasts with *tin-le:du* (eat-Neg) for the past negative form. Thus, finiteness distinctions in Telugu seem to be associated with agreement marking rather than Tense. Unlike standard assumptions about the ban of non-finite clauses as matrix ones (Shlonsky 1997), there are matrix clauses in Telugu which can appear without tense or agreement marking but which are marked for aspect or types of modality, e.g. permissive or obligative, or the verb appears in the infinitival form (Kissock 2014). Based on the above morphological features of Telugu, the lack of a direct correspondence with categories which mark English finiteness and the different conditions imposed on the occurrence of tense and agreement marking in Telugu vs. English verbs, negative L1 transfer may lead to an increase in bare forms in English in contexts where finiteness marking is obligatory (Nadimpalli & Vijaya Babu 2017)[1].

2.1.Finiteness errors in child L2 English

Finiteness marking in English as a second language or in atypical development of English as L1 or L2 is a vulnerable domain of the language showing either extensive delays or deficits in acquisition (Rice & Wexler 2001; Paradis 2011). Finiteness errors in regular past tense marking (-ed) and present tense 3s agreement (-s) are morphological markings which are frequently omitted in obligatory contexts across different populations of English learners, children and adults (Gavruseva, 2008; White 2008). As the focus of the present study is on child L2 learners of English, we will briefly present some findings from the relevant literature focusing on the omission and commission errors made and the use of verb forms which are unattested in the input.

Studies of L2 English in early stages of child L2 acquisition show an asymmetry between bound morphology and free morphemes marking finiteness features (Li 2012; Paradis et al

[1] Although other languages are relevant to our study as minority languages spoken by children in our sample, only seven out of the 90 children did not have some knowledge of Telugu too. We will thus restrict our presentation of finiteness properties to Telugu only, noting the limitation that this raises with respect to our discussion of possible transfer errors.

2008). While copula BE is used at very high percentages by early child L2 learners, BE Aux is omitted in obligatory contexts but is still used more frequently than bound morphology of tense or agreement. In Li's (2012) study, BE Aux is used in incorrect combinations with a bare main verb form such as '*is wait'* or with an inflected main verb such as '*are went'*. Paradis et al.'s (2008) study of a group of child L2 learners of English, a group of L1 typically developing children and a group of children with SLI with a mean age of 5;7 and a mean of 9.5 months of exposure to English showed that L2 learners performed quite poorly in inflectional morphology (3SG -s: 16% and Past -ed: 20% accuracy) and lower than L1 children (3SG -s: 42%, Past -ed: 47%) while the groups did not differ with respect to the production of unbound morphemes. In Ionin & Wexler's (2002) study, Russian child learners of English with different lengths of exposure to the language following arrival in the US were tested on the same finiteness features: copula and auxiliary *be*, third person singular *-s*, and past tense *-ed*. The data came from spontaneous production and exhibited a high number of omissions of inflection especially for 3SG and past tense (78% and 58% respectively). In contrast, omissions of inflected '*be'* forms were lower, namely 16% for copula BE and 33% for BE Aux, revealing the same pattern shown in the previous two studies. Children overused *'be'* in contexts where it was followed by a bare verb form although the meanings of this overgeneralized form varied. The 'BE+bare main verb' form was used in contexts appropriate for the progressive, the generic, the stative, the past and the future, indicating that it should not be analysed as an incorrect progressive form. In Ionin & Wexler's (2002) study as well as in Li's (2012) study children's use of nominative subjects and verb placement were accurate, indicating that their syntactic knowledge of finiteness was in place. The picture emerging from the above studies is that child L2 learners of English from different L1 backgrounds show low accuracy in verbal morphology, very high omission rates, and some commission errors (e.g. *be + verb – inflected or uninflected).*

In a study including learners with longer exposure to English (more than five years) and different ages of arrival to the English-speaking immersion environment, Jia & Fuse (2007) found that only a third of the children had mastered the 3SG *-s* by the end of the study, fewer than half had mastered past irregular and no child mastered past regular. The authors also showed that input/the environment could explain the data rather than the differences in the age of arrival to the US. In a more recent study, Paradis et al. (2016) also examined age of onset effects investigating the acquisition of English L2 verb morphology by Chinese speaking children living in Edmonton, Canada. After six years of exposure, 13 of 18 children reached

monolingual standards on 3SG -s and 15 children had acquired past tense, while individual variation was found in the children's acquisition of verb morphology for the relevant features. Factors including vocabulary size, richness of English input outside the school and allomorph type (-s, -z, -iz) were significant predictors of performance on third person singular marking. Vocabulary size, allomorphy and verbal short-term memory were predictors of past regular marking. Irregular past tense formation was sensitive to word frequency. The contrast between the problematic subject-verb agreement marking with the unproblematic acquisition of BE is particularly relevant to the present study.

In all, previous studies on the acquisition of finiteness morphology in L2 English converge on the high omission rates of bound morphemes for both agreement and tense when compared to the use of copula and auxiliary BE in early stages of development, albeit with copula BE being more target-like than BE Aux. Finally, one error attested in all of the above studies involves the overuse of 'be + main verb*'* structure with or without inflection on the main verb. It is worth pointing out however that this error occurs in the early stages of L2 English regardless of the properties of the L1. For instance, in Paradis et al's (2008) study with learners from different L1s, in Spanish learners (García Mayo et al., 2005), but also in learners of non-English as an L2, e.g. Dutch and German (Van de Craats, 2009). As the use of this form is not a case of incorrect production of the progressive, it raises the question of the role of BE in early stages of L2 English (see Hawkins & Casillas, 2008 for an analysis).

3. Literacy and oral language skills in L2 development

As discussed in the Introduction, oral narratives are a form of connected discourse which have been shown to be developmentally linked to literacy development in young monolingual children. Literacy refers to the ability to read and comprehend in a language. According to the Simple View of Reading (Gough & Tunmer, 1986), reading comprehension is a higher level skill which develops as a result of good decoding and listening comprehension skills. Scarborough's (2001) 'Reading Rope' illustrates how the principles of language comprehension, oral and written, are interconnected and become increasingly strategic with school development and how the independent contribution of word recognition becomes increasingly automatic with training and print exposure. Several studies have also shown that listening comprehension and reading comprehension are highly correlated (Curtis, 1980; Hedrick & Cunningham, 1995). Recent evidence for the link between oral and reading

comprehension is presented in Rydland & Grover's (2019) study. The authors investigated the effects of argumentative peer discussions in Year 5 primary school students, in Oslo. Most children had different L1 backgrounds and the majority language, Norwegian, was used as the school language. In the post peer discussion assessment, children's reading comprehension scores improved when the peer discussions included relevant arguments and counterarguments (in line with Uccelli et al, 2015), supporting the claim that such practices benefit multilingual classrooms (Vaughn et al, 2017). Crucially, Rydland & Grover's study demonstrates that improving critical thinking and verbal reasoning in guided oral discussion has significant effects on reading comprehension in multilingual classrooms (see also McNeill, 2011).

Turning to second language learners, several studies on reading development have shown transfer effects from L1 literacy skills to L2 literacy development (Bialystok, 2001; Dressler and Kamil, 2006; Durgunoğlu, 2002). In particular, good readers use similar comprehension strategies in both of their languages (Jiménez, 1997; Langer et al., 1990; van Gelderen et al., 2007). Strategies such as predicting what the text may be about, relating information in the text to background knowledge, asking questions while reading, monitoring the understanding of the text, and summarizing what is being read (Pressley, 2002), are skills transferrable from L1 to L2 reading. An independent contribution of vocabulary knowledge to reading comprehension however is specific to the language used (L1 or L2) (Cobo-Lewis et al., 2002; Nakamoto et al., 2008). These findings imply that building L2 vocabulary is necessary for listening and reading comprehension. The importance of research in the field of reading development in L1 and L2 is paramount when considering multilingual learners from disadvantaged socioeconomic backgrounds with minimal language and literacy skills. In the present study, participants attend EMI schools, and as a result, the development of literacy is targeted towards English. However, the regional language (Telugu) is also taught as a subject in the schools and we can therefore assume that literacy development is not restricted to English alone. Some degree of literacy skills transfer may thus be expected across languages although not directly at the level of decoding as the Telugu script is different from English albeit based on alphabetic conversion rules. Our focus in this study is the link between oral English language skills expressed through narratives and English literacy development.

Storytelling has been proposed as a pedagogical means to increase early literacy and promote reading comprehension skills (Haven & Ducey, 2007). Thus, apart from the link between oral

language and reading skills widely advocated by the theories presented above and the link between storytelling and literacy development, storytelling is positively associated with comprehension, critical listening, and thinking skills when questioning, inferencing and retelling are practiced (Craig et al., 2001; Washburn, 1983). Since our present study's focus is on microstructure properties, we will not be considering possible links between inferential abilities in listening and reading comprehension based on the literacy and the narrative data we will be presenting (Vogelzang et al, submitted). Instead, we will seek to investigate the link between different aspects of literacy skills - starting from word and letter decoding and moving to sentence and story reading and comprehension - with learners' oral abilities in handling morphosyntactic aspects of English L2, syntactic complexity and narrative length.

4. Research questions and hypotheses

The present study focuses on the following research questions:

RQ1: *How do Indian children from low SES background attending EMI primary schools perform on microstructure in English narrative retelling?*

Hypothesis 1: Based on previous studies on child L2 English studies and given that our participants have had four or five years of English as a school language albeit not in a full immersion context (Lightfoot et al., 2021; Pal & Panda, 2020), we expect that our learners will show a certain amount of variation in the microstructure measures but they will nevertheless be still at beginner levels of abilities. As such we hypothesize that syntactic complexity will be low and that finiteness marking will be omitted more often than not. Finally, we predict that the widely attested overuse of BE+verb (inflected or uninflected) error found in beginner child learners of English from other language backgrounds will also be attested in our participants.

RQ2: *How is microstructure in English narratives linked to oral language input in the classroom?*

Hypothesis 2: Child L2 development has been shown to depend on input measures which in studies carried out in western societies varies considerably out of school but not in school where English is the language used consistently and in abundance.

Participants in this study have had, in theory, four or five years of English medium instruction but no English input at home or in their community[2]. Our hypothesis is then that input differences measured in terms of the amount of English use in the classroom will lead to variation in the microstructure skills observed in children's narrative retellings.

RQ3: *How is microstructure in English narratives related to literacy in English and minority language use at home?*

Hypothesis 3: Narratives involve the presentation of a sequence of events temporally and causally linked in connected discourse. As such, narrative discourse is related to higher levels of literacy skills, i.e. sentence and story reading as well as reading comprehension skills. We hypothesize that children's encoding of finiteness, syntactic complexity and narrative length will be associated with more demanding components of literacy skills such as reading comprehension. We also predict that minority language users, who are more likely to be bilingual at home, will show an advantage in English microstructure properties compared to majority language speakers who are likely to also be monolingual at home. Our hypothesis is based on previous findings from a larger cohort showing that minority language speakers are not at a disadvantage in their English literacy skills although they seem to perform lower in literacy skills in the regional language (Telugu or Hindi) compared to children with the regional/majority language as their home language (Vogelzang et al, submitted).

RQ4: *How is microstructure in English narratives related to demographic variables such as age and gender?*

Hypothesis 4: Children attending the same grade of primary school in India do not always belong to the same chronological year typically found in schools in

[2] In India, English is one of the languages used officially and as a 'link' language in society for historical reasons. However, there is a clear class divide reflected in the use of English in informal contexts, e.g. at home, and this is the reason why in the present study lack of familiarity with English is the same across our participants as they come from low SES background.

Western societies. The question of how overage children perform in relation to their school peers has been addressed in previous studies although the focus has not been on linguistic performance but on school skills such numeracy (Alcott & Rose, 2017). As reported in A&R's study, overage children did not seem to perform differently from their classroom peers. On the other hand, age effects on narrative performance are often found in studies where chronological and school age correlate (e.g. Berman & Katzenberger 2004; Hickmann 2003). It is therefore possible to find some age-related differences in narrative microstructure skills in our participant cohort. Turning to possible gender differences in school performance and learning outcomes, these are frequently reported in studies focusing on the Global South and on India in particular (Alcott & Rose, 2017; Chandra 2019; Asadullah & Yalonetzky, 2012; UNESCO, 2014). This finding is particularly relevant to disadvantaged children where girls show slower development and poorer learning outcomes than boys. On the other hand, language and literacy performance often finds girls outperforming boys (OECD, 2020). We would therefore remain agnostic with respect to the possibility of a gender gap in narrative microstructure skills or any of the measures adopted in the study.

5. THE STUDY

5.1. *Participants*

Ninety children from English medium Instruction (EMI) government schools in Hyderabad in urban slum and non-slum areas were recruited for the purposes of the study. The children were in Year 4 and Year 5 of primary school and differed in the number of home languages they spoke and their age (range 8-12 years)[3].

A child questionnaire which was adapted from Kaltsa, Prentza, & Tsimpli (2019) and Rothou & Tsimpli (2017) was administered to collect information on child demographics, and

[3] It is quite common in classes of India's government schools to include overage children. This is usually due to school capacity, administrative issues, children who may have had a gap in their schooling and other factors. Over age children in the class may lead to an increase of inequalities and may affect the quality of education provision (Tsujita 2009). Nevertheless, both risks and benefits are associated with overage children in the class (Alcott and Rose 2017).

language experience at home. The questionnaire was orally administered to the children by the research assistants. The children were asked to report which languages their parents, siblings, relatives and best friends use while speaking to them and which languages the children use to respond in. The complete questionnaire is provided in Appendix A. In addition, a classroom observation tool (See Appendix B) which was developed as part of the MultiLila project (Tsimpli et al, 2019) was used to gather information on language/s used by teachers and children during math and English lessons separately in 5 minute intervals for a 30- minute duration. The research assistants were asked to indicate the language/s used independently or mixed within the 5-minute interval. (for more details see Lightfoot et al, 2021).

The following Table presents the demographic details of the participating children:

Table 1. Overview of characteristics of the sample

Site	**Grades**	**Boys**	**Girls**	**Total**	**Age**	**Age range**
Slum ***n=61***	Grade IV	14	24	**38**	9.93 (1.17)	8-10
	Grade V	09	14	**23**	10.05 (1.08)	9-12
Non-slum ***n=29***	Grade IV	14	15	**29**	9.6 (0.95)	8-10

Note that these 90 children were drawn from a larger cohort of 204 children in EMI schools in Hyderabad. These 90 children were selected because they were the only ones who completed the story-retelling task in English in contrast to the rest, who despite listening to the story in English chose to retell the stories in Telugu or Hindi. The 90 participants were enrolled in six schools in Hyderabad (three in slum and three in non-slum areas) with English as the medium of instruction (EMI).

5.2. *Language Experience*

In order to establish the use of different languages at home and at school we used the information collected through the child questionnaire (See Appendix A) and the classroom observation tool (See Appendix B), respectively.

Somewhat expectedly, many of the children had multiple and diverse language exposure in their households. The languages spoken in our participant cohort were Telugu (n=49) which is also the majority language of Hyderabad, followed by Hindi (n=11), Kannada (n=10), Lambadi (n=10): Urdu (n=6), Marathi (n=3) and Voddera (n=1).

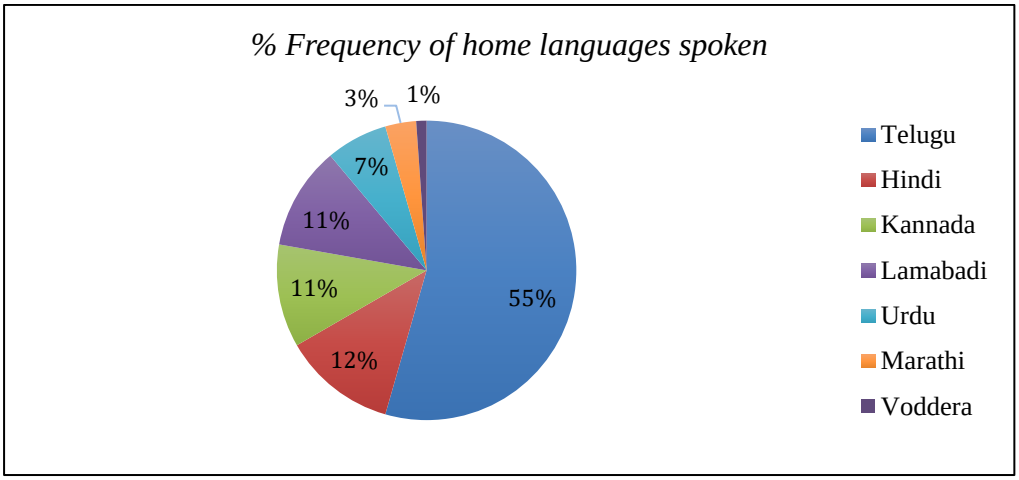

Figure 1. Percent frequency of home languages spoken by the children

We also considered the number of languages spoken at home. If a child had access to more than her/his mother tongue at home, the household was considered a bi/multilingual home in contrast to those who had only one mother tongue at home and were from monolingual households. As Hyderabad is a metropolitan area with a growing migrant population, children from these families have L1s that are minority languages in Telangana, the southern state, and its capital city Hyderabad.[4]. Figure 1 shows that 45% of our participating children are from non-Telugu households. 65 (72%) of the 90 children were from bi/multilingual households. Among the 90 children, data on minority languages was available only for 83 children, and 37 children i.e., 45%, had a minority language as a first language at home while the remaining 46 (55.42%) children had Telugu, i.e. the majority language, as a first language at home.

[4] Telugu is the state's official language and is also the majority language, spoken by 84.86% of the population. Owing to the presence of a Muslim population, Urdu speaking families are at 7.86% (Fatihi 2003)

Turning to language input in the classroom, a classroom observation tool was developed as part of the MultiLila project (Tsimpli et al, 2019) to obtain in-depth information about actual teaching practice, and in particular the use of languages in the classroom. This was important for obtaining further insights into any discrepancies between the school's official language of instruction and the actual languages spoken in class. In India, pedagogical practice in many schools is teacher and textbook centered. We therefore attempted to gauge to what extent learning was teacher-directed and/or student-led, using a 'good practice' table developed by the British Council India and adapted into the Classroom Observation tool. Thirty minutes each in an English language lesson and in a Maths lesson were observed during which teachers' and learners' activities were coded (e.g. reading aloud, asking questions, problem-solving exercises). Languages spoken during the 30 minutes observation were coded at five-minute intervals, to enable the researchers to obtain detailed information about the languages used and any language mixing practices, i.e. use of children's L1s in the classroom together with English or another language (for details refer to Tsimpli et al. 2019; Lightfoot et al. 2021). Language input in class was computed by summing and calculating the mean percent frequency of occurrence of one language (English only, Telugu only, Hindi only) versus language mixing (Telugu-English; Telugu-Hindi-English and so on).

The quantity of input in lesson delivery in the school's MoI, namely English, and in other languages was counted as the language of input in class. By means of a percentage analysis of the six classroom observations, we first present the amount of English-only input as opposed to language mixing input received in EMI classes.

In Figures 2 and 3 we see the extent of language mixing that takes place during teacher and learner talk in sessions of 30 minutes across all the schools. English-only use is attested in half the lessons, while in the rest of the classroom observation other languages are used which include a mix of Telugu with English or with other languages spoken or no language spoken by the children. 'No language spoken' refers to the observation time that the children remained quiet in the classroom. During this time teacher activity was primarily non-interactive, such as teacher reading a story or writing on board.

.

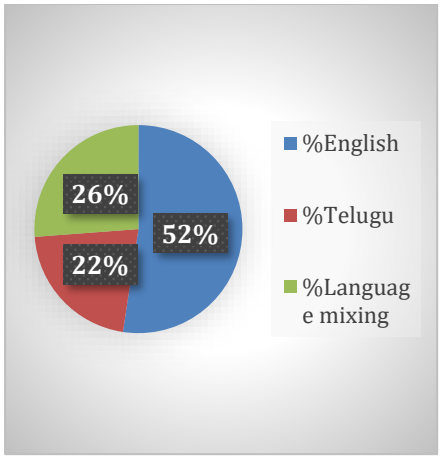

Figure 2. Teachers' use of languages

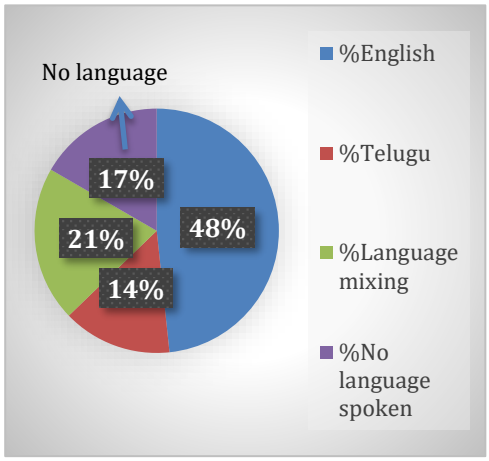

Figure 3. Learners' use of languages

From each school the language of input was calculated based on the percent of English used in a monolingual manner (English-only) to the total amount of language use within the 30-minute observations.

In Figures 4 and 5 we present the focus of the teaching during the English lessons and the types of activities teachers use during these sessions. While most of the teaching revolves around reading aloud from the textbook, a third of the lesson focuses on grammar and vocabulary development. Though teachers cover a wide spread of activities to teach English such as read aloud, narrations, demonstrations and asking questions, they do not spend much time on giving learners feedback that could make the target language input more comprehensible.

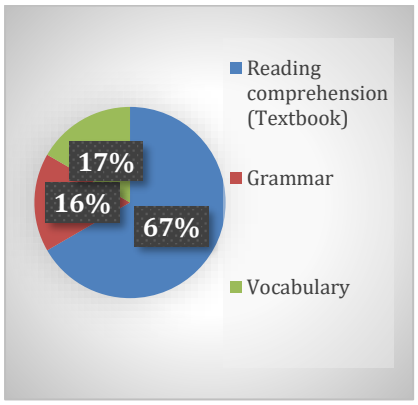

Figure 4. Focus of language lessons

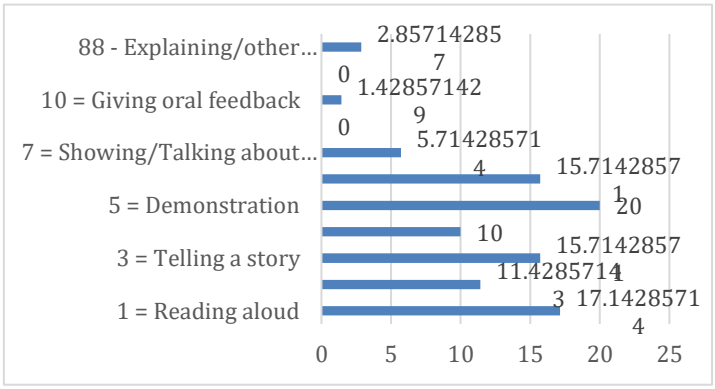

Figure 5. Types of teacher activity

Figures 2-5 give us a glimpse of the quantity of classroom input and the kind of pedagogical activities teachers engage in class. For children growing up in a non-English speaking home or community environment, the amount of English input available in the class and the method in which it is presented, i.e. mostly from a textbook, raises questions about the efficiency and feasibility of achieving appropriate learning outcomes both in terms of language proficiency and in comprehending content taught through English. Furthermore, children's low familiarity with English has an impact on teacher's possibility to teach the lesson's content and clarify concepts using an English-only mode. This is coupled with the problem of teachers' low proficiency in the target language and their lack of training in encouraging learners to use the language communicatively. Teachers then resort to multilingual practices in class to teach English as a subject or a Mathematics class in English. Although this is likely to be a better way for teachers to engage with multilingual children with minimal familiarity with English, it is unlikely to be a sufficient source of English input for children to develop a good level of understanding and production skills in the language.

5.3. Tasks and Method

ASER literacy test

The ASER tool is a widely used test developed in India to measure literacy of learners from grades two to eight enrolled in schools, both government and private, in rural areas (Pratham 2014, 2017). The ASER literacy tasks are available in many different Indian languages as well as in English and have been tested with over 600,000 children in India. We therefore chose the ASER English literacy tool to assess our participants' levels of literacy in their medium of instruction. The test includes items to assess reading for decoding (e.g., naming letters and single word reading), and reading for meaning with read aloud of sentences and short passages. As part of the MultiLila project's use of the ASER tool, we included two novel inferential questions to assess children's reading comprehension of the story (Tsimpli et al, 2020a).

ASER scoring and analysis

The ASER tool had 10 letter and 10 word recognition items, 4 discrete sentence-reading items, 8 sentences in a story and 2 comprehension questions. Children were awarded 1 point for each correctly read letter and word in the recognition tasks. In discrete sentences or sentences in the story, children got 1 point if the sentence was correctly read or one word was read incorrectly.

With two or more incorrectly read words, the response was marked as zero. One point was awarded for each correctly answered question, bringing the maximum score to 34 points.
As a general observation from the data collection process, participants often struggled to read aloud and comprehend the sentences (in isolation as well as in the story). They often had to ask for a translation of the comprehension questions in Hindi or Telugu and gave the answers in those languages only. In some cases, learners chose to point to the parts of the texts that contained the relevant information without articulating the answers.

Narrative retelling task

Narratives were elicited with the Multilingual Assessment Instrument for Narratives (MAIN; Gagarina et al, 2012, 2019). We opted for the retelling mode as in several studies retellings include longer and linguistically richer narratives (Otwinowska, Mieszkowska et al, 2018). Furthermore, the retelling task has been shown to elicit sufficiently rich narratives to investigate micro- and macrostructural properties in either the L1 or the L2 (Tsimpli et al, 2016). Two stories – 'The Cat story' and 'The Dog story' were presented. The stories were piloted with children before the main study began to ensure feasibility of the study in terms of cultural and socioeconomic attributes. As oral culture is very strong in India, storytelling was a task familiar to and enjoyable for the children.

Each child was administered either 'The cat story' or 'The dog story' as picture panels with audio input using a power point presentation. Each story consisted of 3 picture panels with 2 pictures in each panel. Children in EMI were administered the story in English and the narrative retelling was elicited in English[5]. Children in Grade IV were administered the cat story and children in Grade V were administered the dog story.

Data analysis

[5] Note that these 90 children are a sub-set of 204 children from EMI schools. Only these 90 children were selected because they were the only ones who completed the story-retelling task in English in contrast to the rest, who narrated the stories either in Telugu or Hindi.

All narratives were transcribed and then formatted in CHAT format (MacWhinney, 2016). For the purposes of this study, the narratives were analyzed as frequency counts of microstructural properties of two types: (a) syntactic complexity and (b) finiteness marking.

We measured syntactic complexity on children's narratives on the following sub-categories denoting an increasing order of complexity:

(i) Proportion of function words to text length
(ii) Number of verb clauses
(iii) Proportion of verb clauses to text length
(iv) Proportion of simple (mono) clauses to verb clauses
(v) Proportion of complex clauses (coordination & subordination) to verb clauses
(vi) Proportion of coordinate clauses to verb clauses
(vii) Proportion of subordinate clauses to verb clauses

Each of the seven sub-features provides information about the participants' basic syntactic knowledge of English L2. The child's ability to use function words is indicated through the ratio measure over text length. The use of function words was expected to be a good indicator of children's syntactic knowledge of English. Based on the limited English input children receive, we expected syntactic production to instantiate single clause sentences and low finiteness marking accuracy. If any of the utterances did not include a verb, they were considered instances of errors. Finally, we anticipated some instances of complex clausal utterances in the retellings partly because of participants' age and partly due to the English language in participants' textbooks.

We also measured finiteness marking in children's narrative production. Finiteness is an important index of morphosyntactic development particularly at early stages of acquisition. Since the children who retold the narratives were at relatively early stages of English language development the emergence of finiteness was considered to be a relevant and useful area of examination. Moreover, the differences between Telugu and English in terms of finiteness marking presented in Section 2.1 above, may trigger transfer effects from Telugu into English expressed through omission of morphological marking of tense and agreement on the verb forms used in the narratives. Furthermore, the retelling task was expected to trigger the use of reporting the story's events as either with the 'there and then' feature and use of the past tense or as the 'here and now' feature with the use present tense (Krashen, 1985: 86, 107).

The contexts where finiteness was coded are presented in Table 2 below:

Table 2: Contexts for accuracy/error in finiteness marking

Use of correct finiteness	Errors
Use of auxiliary [BE]	(i) [+BE], [-prog] e.g. 'he is catch the butterfly'
Use of copula [BE]	(ii) [-BE], [+prog] e.g. 'he eating fish'
Correct finiteness i. [+Copula BE, + Complement] ii.[+Aux BE], [+ prog] iii. [+other verb, + correct tense morphology]	(iii) [+Other verb, +wrong morpheme /ed/] e.g. 'the dog was stucked in the tree'
	(iv) Bare verbs used (BV) e.g. 'the dog run'

The verb BE was considered in its uses as copula or auxiliary; other light verbs as GET, SEEM etc. and auxiliary verbs as HAVE, DO, or modal auxiliaries were not expected to be frequently occurring because of the nature of the task and the low English proficiency of participants. So any occurrence of these verbs would be counted under the category of 'other verbs'.

Correct finiteness was measured as a sum of (i) to (iii) in the column 'use of correct finiteness' in Table 2 above and then converted to percent frequency. Likewise, errors in finiteness were calculated by the sum of (i) to (iv) in the column 'errors in finiteness' and converted to percent frequency.

Table 3. Examples of errors in finiteness marking

Finiteness error types	Participant: 211154AFEB Gender: Female Year: 4	Instances of errors	Percent frequency of errors
[+BE, -prog]	boy is ball, fish there bucket ball is water go	2	40
[-BE, +prog]	one day cat butterfly looking. cat looking	2	40
bare verbs	-	0	0
[other verb,+wrong tense]	cat is going to fish eating.	1	20
Finiteness errors (Total)		5	100

The microstructural properties were evaluated in relation to (i) text length measured as number of words in the narrative and (ii) the frequency of verb clauses appropriate for each of the sub-categories of syntactic complexity and finiteness. Both of these measures served as denominators to calculate complexity and finiteness.

Analyses

We considered five independent variables: proportion of English language input in the classroom, literacy in English, minority language used as the first home language, age and gender. The dependent variables are the microstructural properties of *syntactic complexity* and *finiteness* errors. The total score of finiteness errors is also expressed in percent proportion. We used Generalized Linear Model (GLM) implemented through statsmodels (Seabold & Perktold, 2010) in Python. In addition, we carried out correlational analyses to examine the relationship between the different dependent variables (i.e., use of function words and use of complex sentences and use of function words and use of simple sentences) using Spearman's rank correlation which was calculated using the SciPy 1.0 package in python (Virtanen et al., 2020). The ranks were calculated based on equation 14.19 given in Zwillinger and Kokoska (2000).

6. Results

6.1. How do children perform on microstructure in L2 English narratives?

The descriptive statistics for the 90 participants on a range of narrative microstructure properties (including proportion of function words, content words, types of sentences and finiteness) are presented in Table 4. In general, children produce a higher proportion of content words (mean=0.63) compared to function words (mean=0.37) when controlled for text length. The number of verb clauses that children produce range from 1 to 36. They also produce fewer complex sentences (mean=1.76) as compared to simple sentences (mean=5.73). Participants' production of correct finiteness ranged from 0 to 100 percent with a mean of 46.77%.

Table 4. Performance of children N=90 on narrative microstructure properties.

Microstructure properties	**Min-max**	**Mean (SD)**
Syntactic Complexity		
Number of Function words	0-114	22.68 (17.00)
Proportion of Function words/Text length	0-0.63	0.37 (0.11)
Number of Content words	10-127	35.99 (20.08)
Proportion of Content words/Text length	0.37-1	0.63(0.11)
Number of Verb Clauses (VC)	1-36	9.51 (6.37)
No. of Simple sentences (SS)	0-19	5.73 (3.32)
SS/VC	0-1	0.70 (0.29)
SS/VC (%)	0-100	70.44 (29.10)
No. of Complex sentences- coordinate (CSC)	0-10	1.28 (1.75)
CSC/VC	0-0.4	0.09 (0.11)
CSC/VC (%)	0-40	9.20 (10.92)
No. of Complex sentences-subordinate (CSS)	0-6	0.48 (1.05)
CSS/VC	0-0.33	0.04(0.07)
CSS/VC(%)	0-33.33	3.51(6.83)
Total Complex Sentences (CS)	0-12	1.76 (2.34)
CS/VC	0-0.5	0.13 (0.14)
[CS/VC]%	0-50	12.71 (13.78)
Finiteness		

Correct finiteness	0-23	4.80 (4.24)
Correct Finiteness%	0-100	46.77 (27.16)
Finiteness errors	0-20	4.68 (3.82)
Finiteness error (%)	0-100	51.85 (27.05)

Additionally, we examined if there was a correlation between different microstructure properties of the narratives. The results are presented in Table 5 below. The findings indicate a significant negative correlation between simple sentences and function words (r=-0.34, p<0.01) and a significant positive correlation of simple sentences with content words (r=0.34, p<0.01) showing that children who produced higher numbers of simple sentences produced fewer function words and more content words.

Table 5. Spearman's rank correlation between different microstructural properties of the narratives

	Simple sentences	Complex sentences	Function words	Content words
Simple sentences	1			
Complex sentences	-0.73***	1		
Function words	-0.34**	0.21	1	
Content words	0.34**	-0.21	-1	1

*p<0.05; **p<0.01; **p<0.001

6.2. How is microstructure in English narratives linked to oral language input in the classroom?

We examined whether the *level of oral language input* that children receive in the classroom affected any of the microstructure properties of their oral narratives. We used a General Linear model (GLM) to determine which of the independent variables significantly predicted the dependent variable. The independent variable was percentage of English input in the classroom while the dependent variables were Correct finiteness (%), Finiteness errors (%), Simple sentences/Verb clause, Complex sentences-coordinate/Verb clause, Complex sentences-subordinate/Verb clause, Proportion of function words and Proportion of content words. Separate GLM models were run for each of the dependent variable with the independent variable.

The percentage of English input in the classroom showed a significant positive effect on Complex sentences-coordinate/Verb clause ($\beta = 0.09$, $z = 4.05$, $p<0.001$) but no significant effects on any of the other microstructure properties of oral narratives.

6.3. How is microstructure in English narratives related to literacy in English and minority language use at home?

The third research question examined whether *Literacy scores (Letters-Words, Sentences-Paragraphs and Comprehension questions*) and *minority language use at home* had an effect on any of the microstructure properties of English narratives of children.

Separate GLM models were carried out, one with literacy scores *(Letters-Words, Sentences-Paragraphs and Comprehension questions*) as independent variables and each of the microstructure properties of English narratives as the dependent variables (one dependent variable per model) and the other with *minority language use at home* as an independent variable and the microstructure properties of English narratives of children as the dependent variables (one dependent variable per model).

Table 6 below presents participants' performance on the ASER literacy tool. On average, children showed relatively high performance on letter-word reading in English, with performance decreasing with the complexity of the task (i.e. poor performance on sentence-paragraph reading and lowest performance on comprehension questions).

Table 6. English Literacy scores per component and literacy composite score.

Literacy scores (English)	**Min-Max**	**Mean (SD)**
Letters-words reading	0-20	17.42 (3.36)
Letters-words (%)	0-100	87.11 (16.81)
Sentences-paragraph reading	0-12	8.62 (4.26)
Sentences-paragraph reading (%)	0-100	71.85 (35.52)
Comprehension questions	0-2	0.62 (0.84)
Comprehension questions (%)	0-100	31.11 (42.12)
Total Composite score	0-34	26.69 (7.69)
Total Composite score (%)	0-100	78.50 (22.61)

The statistical comparisons of the microstructure properties of the narratives and English literacy scores using general linear models revealed none of the independent variables had a significant effect on Simple sentences/Verb clause, Complex sentences-coordinate/Verb clause and Complex sentences-subordinate/Verb clause. However, performance on reading comprehension questions had a significant effect on finiteness errors (β = -0.17, z = -2.303, p=0.018), and proportion of function words (β =0.075, z = 2.53, p=0.011).

Turning to the role of minority language in English L2 performance, at first glance, the minority language speakers in our sample seemed to perform better on some of the microstructural properties of the narratives. For instance, minority language speakers produced fewer simple sentences, more complex-coordinate clauses and a higher proportion of function words as compared to majority language speakers (Figure 6 below). However, the statistical comparisons of the microstructure properties of the narratives and use of minority languages at home using general linear models revealed no statistically significant differences between majority and minority language speakers on any of the microstructural properties of narratives in children.

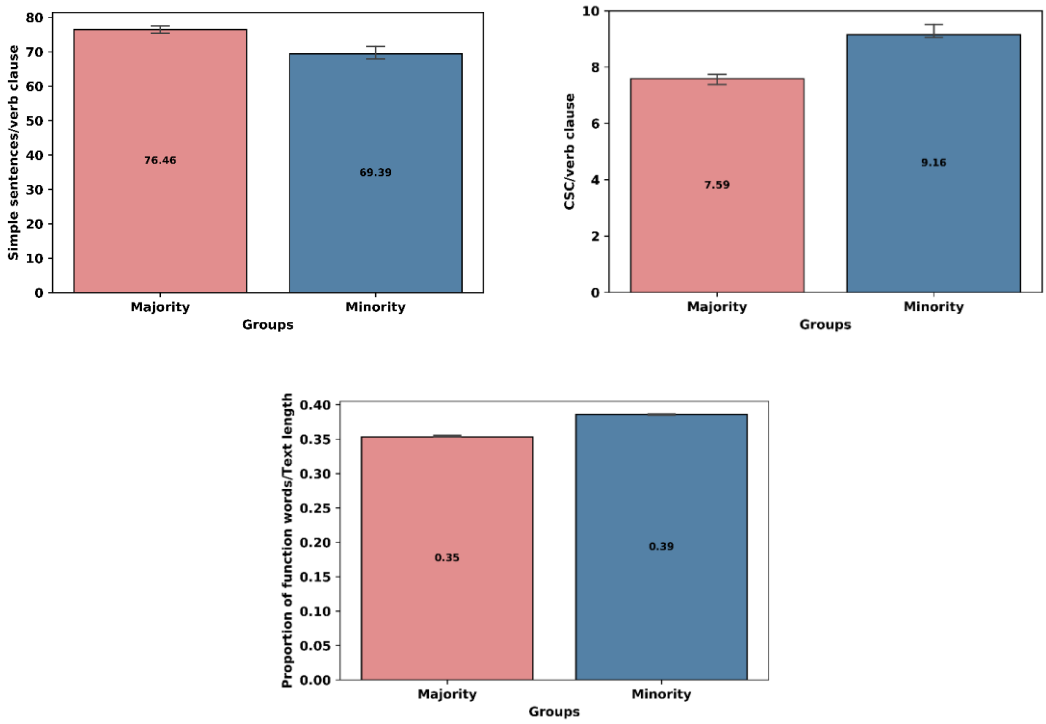

Figure 6. Performance of majority language speakers and minority language speakers on microstructural properties of narratives- simple sentences (top left), complex co-ordinate clauses (CSC) (top right) and proportion of function words (bottom).

6.4: *How is microstructure in English narratives related to demographic variables such as Age and Gender?*

To examine the fourth research question *Age* and *Gender* were considered as independent variables with microstructural properties of English as dependent variables. The results showed that age had a significant effect only on Simple sentences/Verb clause (β = -5.33, z = -2.05, p=0.04). Younger children (<10 years: Mean=76.63) produced more simple sentences compared to older children (> 10 years: Mean=66.69). Gender did not have a significant effect on any of the microstructural properties of the narratives, however there were some effects of gender on finiteness error types which are discussed below.

Finiteness marking and error analysis

Based on a frequency count analysis, it was found that the children had used a total of 856 verb clauses while 117 utterances had no verbs. Furthermore, in the verb clauses, 432 instances of correct finiteness occurred in three different categories: 118 instances of [+copula BE, +complement], 171 instances of [+aux BE,+prog] and 143 instances of [+other verb, + correct tense]. There were also 421 uses of incorrect finiteness spread across the following 4 categories: 78 instances of [+BE, -prog], 131 of [-aux, +prog], 188 of bare verbs, and 24 of [other verb, +wrong tense]. Table 7 below presents the frequency counts and mean percentage of categories of correct finiteness and finiteness error types.

Table 7. Frequency count and mean percentage of errors of Types of correct finiteness and finiteness errors

Correct finiteness types	**Frequency count of correct finiteness**	**Mean percentage of correct finiteness**
[+copula BE, +complement]	118	27.27
[+aux BE,+prog]	171	34.98

[+other verb, + correct tense]	143	26.62
Total	432	100
Finiteness error types	**Frequency count of errors**	**Mean percentage of errors**
[+BE, -prog]	78	20.97
[-BE, +prog]	131	34.53
bare verbs	188	35.80
[other verb, +wrong tense]	24	5.37
Total	421	100

The most frequent error type was the use of bare verbs closely followed by [-BE, +prog] verb forms. The least frequent error type was [+Other verb, +wrong tense]. A repeated measures ANOVA with a Greenhouse-Geisser correction indicated that there was a statistically significant difference among error types [F (3.22, 286.74) =16.12, p<0.001, η^2=.15]. Post-hoc tests using a Bonferroni correction revealed a significant difference between [+Other verb, +wrong tense] and all other error types (p<0.001); Figure 7:

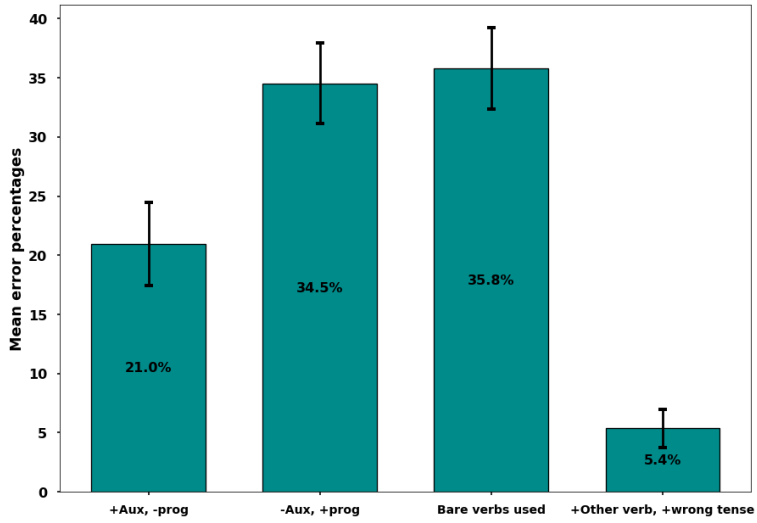

Figure 7: Percent of finiteness errors per type.

In addition, a repeated measures ANOVA was carried out to examine if there were differences in error types across boys and girls. There was no significant main effect of Gender [F (1,88) =0.07, p=0.78, η^2=.001]. There was a significant main effect of error type [F (2.4, 212.77) =17.86, p<0.001, η^2=.16]. There was a significant interaction between Gender and error types [F (2.41, 212.77) =4.74, p=0.006, η^2=0.05]. Both boys and girls produced a similar proportion of errors for three categories of error types; however, for use of bare verbs male participants

(47.81%) had a higher mean compared to females (27.41%). A breakdown of errors by gender is presented in Figure 8 below.

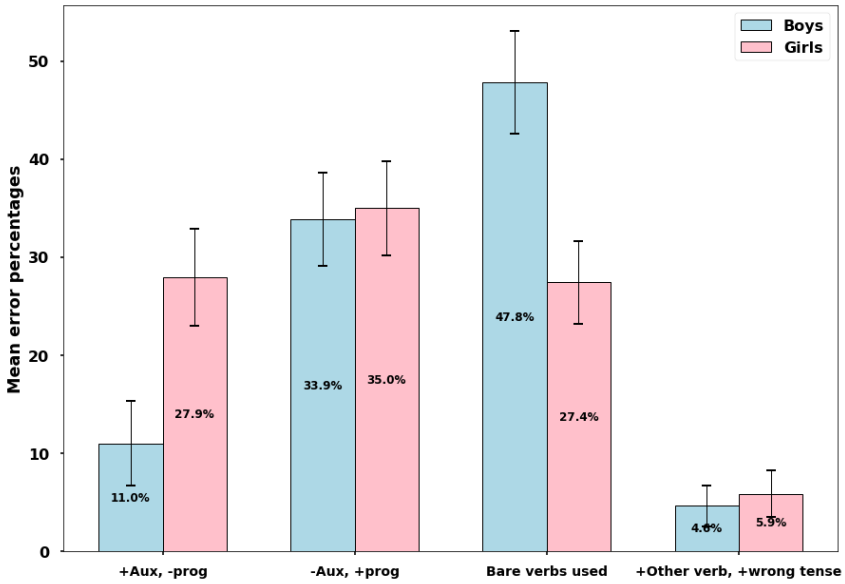

Figure 8: Percent of finiteness errors by participant gender.

7. Discussion

The present study aimed to examine narrative microstructure in the narrative retellings of primary school children in India coming from a disadvantaged background and educated in EMI government schools. The role of classroom input in the target language, literacy performance, language background as well as the age andgender of the children in the study were considered as potential factors affecting L2 English production. The children's language background identified as the number and nature of household languages (minority or majority language as first language of the participant) were important variables due to India's considerable linguistic diversity (Linguistic Diversity Index; UNESCO 2009) which entails that multilingualism in the household, the classroom and the society are the norm (Tsimpli et al, 2019). A further distinction in the participants' language experience concerns the status of their first language. As internal migration is a common phenomenon in urban sites in India a minority language is often the first language children are exposed to while the majority language (Telugu in our study) is likely to also be used at home but not with the same frequency as the minority language. Minority language speakers may be further disadvantaged in education in EMI schools due to their lower familiarity with the majority language which, as

shown in our study, is the language which is often mixed with English. This added disadvantage for minority language children is referred to as the 'linguistic double divide' (Mohanty 2013, 2019) and reflects the educational status and prestige assigned to English first, the vernaculars or the regional dominant languages second, and tribal minority languages featuring last in the list.

Socioeconomic status was not a factor considered in the present study mainly because all children attended government schools which nowadays indicate families' underprivileged status (Agarwal 2014; Alcott and Rose 2015; Pratham 2017), although some variation in SES can still be found among low SES children (Tsimpli et al, 2020a). Crucially for our study, children's level of deprivation at the SES level was informative with respect to the lack of English in the home environment or the community and the minimal home literacy support children were expected to receive. Many of the children, particularly from slum areas, are 'first generation learners', i.e. children whose parents are not literate in any language (Tsimpli et al. 2020b). We can therefore rather safely conclude that children's opportunities for additional support with school skill development is minimal or unavailable. The task of learning English and school skills in EMI contexts needs to take place only within the classroom. To this end, it was essential to measure the amount of English language input in the classroom with the use of our observation tool. The pattern of language use in the classroom showed a certain percentage of English-only input (teachers use English only for 52% while learners use it at 48% of the class time for a period of 30 minutes in a school day) to be available to the child. Language mixing and use of the regional language is used for the rest of the time and this makes up the use of languages for classroom interactions of the total observation time of 30 minutes per class.

Our focus on microstructure measures included syntactic complexity and the use of finiteness marking in the children's narrative production. The choice of finiteness measures is motivated by findings from previous studies indicating that this domain is an index of English language development and a cause of delay in different populations of learners, neurotypical and atypical. Furthermore, the language-specific means of encoding finiteness in English and Telugu could possibly trigger negative transfer effects showing up as omission errors in children's L2 English verb forms.

Accordingly, our first research question asked about participants' performance on the microstructure of English narratives. As expected, the children's narrative length varied considerably and a preference for simple over complex clauses was found. In addition, subordinate clauses were rare among complex clauses while coordination was more frequent in children's narratives. The proportion of function words was lower than that of content words, which is expected with learners of low language ability (Miller & Klee 1995; Tsimpli et al, 2016).

With respect to finiteness marking, just under half of the forms in the children's narratives were accurate in the use of inflectional morphology and in the use of BE as auxiliary or as a main verb (47%). For the remaining half of produced verb forms, the majority of errors (36% of the total errors) were omission errors, which are very common in early stages of English as L1 or L2. However, the error rate in our study seemed to be lower than that reported in Paradis et al (2008) or Ionin & Wexler's (2010) study, although the methodology of data elicitation did not include narratives in those studies. In Paradis et al.'s study, data were collected using the test of early grammatical impairment (TEGI) task (Rice & Wexler 2001), which is an elicitation task targeting individual verb forms. Ionin & Wexler's (2010) study used spontaneous production data. Interestingly, the high percentages of omission in these studies were found in data from immigrant children residing in the US and having had more cumulative exposure to the language than the children in our study.

In terms of accuracy rates of finiteness marking, the present findings seem to be more similar to the data in Ntalli (2020) from narrative production by Chinese 12-year-old learners of English who have had five years of exposure to English through instruction in 3hrs/week structured lessons out of school hours. Ntalli (2020) compared Chinese with Russian L2 learners of English using the TEGI task but also the MAIN narratives, used in the present study too. All of her participants had 5 years of exposure to English L2 in their home countries. Although the 12-year-old Russian children outperformed the Chinese age-peers in Ntalli's (2020) study, the data from our Indian children are similar to the group of 12-year old Chinese children who were 47% accurate with bound morphology on main verbs in English. These findings indicate that our participants have progressed in their use of English L2 marking of finiteness and that the narrative retell mode of data elicitation may offer them the opportunity to produce more target forms than they would otherwise. Furthermore, the influence of Telugu, a first language for two thirds of our participants, may have led to the most frequent error

observed in our data, namely omission of finiteness marking, as a result of negative transfer. This is consistent with Ntalli's (2020) L1-based explanation of the contrast between Russian and Chinese children in the use of bare forms, whereby Chinese children make more such errors despite having similar years of English L2 exposure as the Russian children in her study. Omission of the auxiliary BE with a main verb in the *-ing* form was the second most frequent type of error in our data. This error type has also been found in early L1 English production and in L2 child and adult learners of English (Hawkins & Casillas 2008).

The overuse of BE + bare verb form constituted roughly one in five errors in our results. Although the presence of this form in our data provides further support to the claim that it is a marker of an early stage of English interlanguage (Hawkins & Casillas 2008), it is interesting that its use was not as high as the use of bare verb forms or the omission of the auxiliary BE in finite contexts. Although the status of the BE + bare form has not been conclusively analyzed, it has been suggested that it is a form that encodes a finiteness feature / phrase in early L2 grammars but remains underspecified for tense and agreement features (cf. Ionin & Wexler 2002; Ntalli 2020). We suggest that the present study supports this view and that the overuse of BE + bare verb form, although lower in terms of frequency compared to other errors, may also be on a par with the apparently accurate uses of BE auxiliary + progressive verb form which constitute the majority of correct uses of finiteness in our data. In other words, we would like to propose that our participants' English grammar includes a finiteness category which is not as yet specified for the appropriate finiteness features in the target L2. As such, it is expressed mostly periphrastically with the use of BE as a placeholder selecting either a bare verb form or an *–ing* form as the second element of the periphrasis.

Turning to our second research question regarding the role of English classroom input on microstructure properties, our findings reveal a positive effect of input on syntactic complexity and in particular on the use of complex sentences. This result is clearly in the expected direction: the more the input the higher the number of complex sentences produced by the children. However, given that none of the other measures of microstructure are related to English input, it is worth considering whether the development of morphosyntactic skills, i.e. finiteness marking in our study, is not *directly* affected by input measures but instead involves the interaction between L1 effects and the expected developmental trajectory of finiteness in English L2. Our results appear to run counter to previous studies (Jia & Fuse 2007, Paradis et al 2016, Unsworth 2016, Armon-Lotem et al. 2011) where input, along with

other variables like vocabulary, predict performance on L2 English morphosyntax. In Paradis et al's (2016) study it is richness of input, i.e. a quality of input measure, rather than input quantity that predicted performance. The present study focused on input as a quantitative measure although one could argue that the variety of language input provided in the classroom and the as yet unexplored quality of the English input teachers in government primary schools in India provide may require a more sophisticated measure that combines quality and quantity. We leave this question open for future research.

Our third research question investigated the role of literacy on one hand and the demographics of the participants on narrative microstructure. Although the decoding components of the literacy task were not associated with microstructure, narrative comprehension was found to be a significant predictor of the use of function words as well as finiteness errors. These findings are in the right direction in separating decoding skills in literacy tasks from the higher level skills of reading comprehension which have been shown to connect better with oral language abilities (Nation et al 2010; Lervåg et al 2018). In this respect, our study shows that better oral grammar skills are associated with better reading comprehension measures. In a study focusing on lexical diversity measures within the same set of narratives from our participants, lexical diversity has been found to be a predictor of reading comprehension (Treffers-Daller et al, 2022). It therefore seems that oral skills measured through microstructural features including lexical diversity and grammar are good predictors of reading comprehension skills. Based on these findings, we would like to suggest that oral skills measured through narrative retelling offer the opportunity of a promising intervention method for improving reading comprehension skills. Narratives can then become a powerful pedagogical tool to provide comprehensible input and output for the benefit of oral and written language skills in instructional settings. In addition, oral narratives can be beneficial for a learner community with limited exposure or opportunities to use the target/school language outside the classroom which can in turn have a positive impact on literacy development.

Turning to participants' language background there was no interaction between minority/majority language and narrative microstructure measures. Nevertheless, when looking into the errors that minority and majority language speakers made on English finiteness we found that minority language speakers produced more complex sentences, albeit coordinate structures rather than subordination, as well as more function words, indicating higher syntactic complexity in their English narratives. For minority language children, English is – at least-

their third language, given that Telugu will also be familiar to them to different degrees, being the majority language. In a different study from the same project with 1200 participants (Vogelzang et al, under review) we found a similar advantage in English literacy scores for children coming from multilingual households, including children with a minority language as their first home language. Although minority language speakers in this larger study are disadvantaged in literacy skills when tested in the majority language compared to the performance of majority language speakers, the two groups do not differ when reading in English (Vogelzang et al, under review). Finally, error analysis by age revealed an unsurprising advantage of older children with complex sentences while an analysis by gender only revealed a difference in error types: boys' errors in finiteness marking are primarily omission errors whereas girls attempt to produce a periphrastic structure, albeit erroneously, more often than a bare verb form. It is not clear what could trigger the difference between omission errors and periphrasis although omissions may signal an avoidance strategy in view of low proficiency skills while periphrasis may indicate a transfer effect from Telugu.

8. Conclusion

To conclude, the microstructural properties in oral narrative retellings of Indian multilingual learners through an elicited production task gives us a glimpse into their acquisition of morpho-syntactic features in English which is not a language that the children have access to outside of the classroom. Their acquisition rate is not age appropriate; but when the years of exposure and (low) quantity of English input are considered, their learning seems to be going in the right direction. The findings from the study show they have made attempts to use English syntax through a meaningful activity and reveal their learning capacities in an environment that is not input rich and use of the language in a communicative manner is not often encouraged. Their performance provides evidence of acquiring lexical and syntactic features along the natural path of acquisition (White 2003). Although their rate of development is low, their path of acquisition is comparable to other second/foreign language learners of English.

References

Agrawal, T. (2014). Educational inequality in rural and urban India. International Journal of Educational Development, 34, 11-19.

Alcott, B., & P. Rose. (2017). Learning in India's primary schools: How do disparities widen across the grades? *International Journal of Educational Development* 56, 42-51.

Altman, C., Armon-Lotem,S., Fichman, S., and Walters,J. (2016). Macrostructure, Microstructure, and Mental State Terms in the Narratives of English–Hebrew Bilingual Preschool Children with and Without Specific Language Impairment. *Applied Psycholinguistics 37 (1),* 165–193. doi:10.1017/S0142716415000466.

Amritavalli, R. (2013). An English for every schoolchild in India. *English Impact Report*, 21.

Asadullah, M.N., Yalonetzky, G., 2012. Inequality of educational opportunity in India: changes over time and across states. World Dev. 40 (6), 1151–1163

Baixauli, I., Colomer, C., Roselló, B., & Miranda, A. (2016). Narratives of children with high-functioning autism spectrum disorder: A meta-analysis. *Research in Developmental Disabilities, 59*, 234-254.

Beals, D. E. (2001). Eating and reading: Links between family conversations with preschoolers and later language and literacy. *Beginning literacy with language: Young children learning at home and school*, 75-92.

Berman, R. & I. Katzenberger (2004) Form and function in introducing narrative and expository texts: A developmental perspective. *Discourse Processes,* 38(1), 57-94.

Berman, R., & Slobin, D. (1994). Relating events in narrative: A crosslinguistic develop mental study. Mahwah, NJ

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press. http://www.cambridge.org

Blom, E., & Paradis, J. (2015). Sources of individual differences in the acquisition of tense inflection by English second language learners with and without specific language impairment. *Applied Psycholinguistics*, *36*(4), 953.

Bohnacker, U. T. E. (2016). Tell me a story in English or Swedish: Narrative production and comprehension in bilingual preschoolers and first graders. *Applied Psycholinguistics*, *37*(1), 19-48.

Boudreau, D. (2008). Narrative abilities: Advances in research and implications for clinical practice. *Topics in Language Disorders*, *28*(2), 99-114.

Brutt-Griffler, J. (2017). English in the multilingual classroom: implications for research, policy and practice. *PSU Research Review*,1(3),216-228. doi: 10.1108/PRR-10-2017-0042.

Chondrogianni, V., & Marinis, T. (2011). Differential effects of internal and external factors on the development of vocabulary, tense morphology and morpho-syntax in successive bilingual children. *Linguistic Approaches to Bilingualism*, *1*(3), 318-345.

Cobo–Lewis, A. B., Pearson, P. Z., Eilers, R. E., & Umbel, V. C. (2002). Effects of bilingualism and bilingual education on oral and written Spanish skills: A multifactor study of standardized test outcomes. In D. K. Oller & R. E. Eilers (Eds.), Language and literacy in bilingual children (pp. 98–117). Clevedon, UK: Multilingual Matters.

Craig, S., Hull, K., Haggart, A. G., & Crowder, E. (2001). Storytelling addressing the literacy needs of diverse learners. *Teaching exceptional children*, *33*(5), 46-51.

Curtis, M. E. (1980). Development of components of reading skill. *Journal of Educational Psychology*, *72*(5), 656–669. https://doi.org/10.1037/0022-0663.72.5.656

Dickinson, D. K., & McCabe, A. (2001). Bringing it all together: The multiple origins, skills, and environmental supports of early literacy. *Learning Disabilities Research & Practice*, *16*(4), 186-202.

Dressler, C., & Kamil, M. L. (2006). *First- and Second-Language Literacy.* In D. August & T. Shanahan (Eds.), *Developing literacy in second-language learners: Report of the National Literacy Panel on Language-Minority Children and Youth* (p. 197–238). Lawrence Erlbaum Associates Publishers

Durairajan, G. (2018). 'Convivial cross-lingual evaluation in multilingual Indian classrooms: an exploration of possibilities.' *Multilingual Education in India: The Case for English.* New Delhi: Viva Books.

Durgunoğlu, A. Y. (2002). Cross-linguistic transfer in literacy development and implications for language learners. *Annals of Dyslexia, 52,* 189–204. doi:10.1007/s11881-002-0012-y

Erling, E. J., Adinolfi, L., Hultgren, A. K., Buckler, A., & Mukorera, M. (2016). Medium of instruction policies in Ghanaian and Indian primary schools: an overview of key issues and recommendations. *Comparative Education*, *52*(3), 294-310.

Fatihi, A. R. (2003). Urdu in Andhra Pradesh. *Language in India: strength for today and bright hope for tomorrow*, *3*.

Fiestas, Christine E., and Elizabeth D. Peña.(2004). Narrative Discourse in Bilingual Children: Language and Task Effects. *Language Speech and Hearing Services in Schools 35 (2),*155. doi:10.1044/0161-1461(2004/016.

Gagarina, N. V., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Balčiūnienė, I., Bohnacker, U, & Walters, J. (2012). MAIN: Multilingual assessment instrument for narratives. *ZAS papers in linguistics, 56,* 155-155.

Gagarina, N. (2016). Narratives of Russian-German preschool and primary school bilinguals: Rasskaz and Erzaehlung. *Applied Psycholinguistics*, *37*(1), 91.

Gagarina, N., Klop,D., Tsimpli,I.M., and Walters, J. (2016). Narrative Abilities in Bilingual Children. *Applied Psycholinguistics 37 (1)*,11–17. doi:10.1017/S0142716415000399.

Gagarina, N., Klop, D., Kunnari, S., Tantele, K., Välimaa, T., Bohnacker, U., & Walters, J. (2019). MAIN: Multilingual Assessment Instrument for Narratives–Revised. *ZAS Papers in Linguistics*, *63*, 20-20.

García Mayo, M. D. P., Ibarrola, A. L., & Liceras, J. M. (2005). Placeholders in the English interlanguage of bilingual (Basque/Spanish) children. *Language Learning*, *55*(3), 445-489.

Gathercole, V. C. M. (Ed.) (2013). *Issues in the assessment of bilinguals*. Bristol: Multilingual Matters.

Golberg, H., Paradis, J., & Crago, M. (2008). Lexical acquisition over time in minority first language children learning English as a second language. *Applied psycholinguistics*, *29*(1), 41-65.

Gough, P. B., & Tunmer, W. E. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, *7*(1), 6–10. doi:10.1177/074193258600700104

Govindarajan, K., & Paradis, J. (2019). Narrative abilities of bilingual children with and without developmental language disorder (SLI): Differentiation and the role of age and input factors. *Journal of communication disorders*, *77*, 1-16.

Gusewski, S., & Rojas, R. (2017). Tense marking in the English narrative retells of dual language preschoolers. *Language, Speech, and Hearing Services in Schools*, *48*(3), 183-196.

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

Haven, K. F., & Ducey, M. (2007). *Crash course in storytelling*. Greenwood Publishing Group.

Hawkins, R., & Casillas, G. (2008). Explaining frequency of verb morphology in early L2 speech. *Lingua*, *118*(4), 595-612.

Hedrick, W. B., & Cunningham, J. W. (1995). The relationship between wide reading and listening comprehension of written language. *Journal of Literacy Research*, *27*(3), 425–438. https://doi.org/10.1080/10862969509547890

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

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*. doi:10.1017/ S1366728911000538

Ionin, T., & Wexler, K. (2002). Why is 'is' easier than '-s'?: acquisition of tense/agreement morphology by child second language learners of English. *Second language research*, *18*(2), 95-136.

Jia, G., & Fuse, A. (2007). Acquisition of English grammatical morphology by native Mandarin-speaking children and adolescents: Age-related differences. *Journal of Speech, Language, and Hearing Research*, 50, 1280–1299. doi: 1092-4388/07/5005-1280

Jiménez, R. T. (1997). The strategic reading abilities and potential of five low-literacy Latina/o readers in middle school. *Reading Research Quarterly*, *32*(3), 224-243.

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

Kaltsa, M., Prentza, A. & Tsimpli, I. M. (2019). Input and literacy effects in simultaneous and sequential bilinguals: The performance of Albanian-Greek speaking children in sentence repetition. International Journal of Bilingualism. doi:10.1177/1367006918819867

Kapalková, S, Polišenská,K., Marková, L & Fenton,J.(2016). Narrative Abilities in Early Successive Bilingual Slovak–English Children: A Cross-Language Comparison. *Applied Psycholinguistics 37 (1)*: 145–164. doi:10.1017/S0142716415000454.

Kissock, M. J. (2014). Evidence for 'finiteness' in Telugu. *Natural Language & Linguistic Theory*, *32*(1), 29-58.

Krashen, S. (1985). *The Input Hypothesis: Issues and Implications*. London: Longman.

Kunnari, S, Välimaa,T and Laukkanen-Nevala,P. (2016). Macrostructure in the Narratives of Monolingual Finnish and Bilingual Finnish–Swedish Children. *Applied Psycholinguistics 37 (1),* 123–144. doi:10.1017/ S0142716415000442

Langer, J. A., Bartolome, L., Vasquez, O., & Lucas, T. (1990). Meaning construction in school literacy tasks: A study of bilingual students. *American Educational Research Journal*, *27*(3), 427-471.

Lervåg, A., Hulme, C., & Melby-Lervåg, M. (2018). Unpicking the developmental relationship between oral language skills and reading comprehension: It's simple, but complex. *Child Development*, *89*(5), 1821-1838.

Li, M.-C. (2012). *The acquisition of tense and agreement by early child second language learners*. Unpublished PhD Dissertation, University of Illinois, Urbana-Champaign.

Lightfoot, A., Balasubramanian, A., Tsimpli, I., Mukhopadhyay, L., & Treffers-Daller, J. (2021). Measuring the multilingual reality: lessons from classrooms in Delhi and Hyderabad. International Journal of Bilingual Education and Bilingualism, 1-21, DOI: 10.1080/13670050.2021.1899123. MacWhinney, B. (2016). The TalkBank Project tools for analyzing talk—electronic edition part 2: the CLAN programs. In.

Mäkinen, L., Loukusa, S., Nieminen, L., Leinonen, E., & Kunnari, S. (2014). The development of narrative productivity, syntactic complexity, referential cohesion and event content in four-to eight-year-old Finnish children. *First Language, 34*(1), 24-42.

Marphatia, A. A., Reid, A. M., & Yajnik, C. S. (2019). Developmental origins of secondary school dropout in rural India and its differential consequences by sex: a biosocial life-course analysis. *International Journal of Educational Development, 66,* 8-23.

Meganathan, R. (2018). 'An exploration into English language teaching in multilingual contexts.' *In* Tomlinson, B. and Keedwell, A. (Eds). *Explorations: Teaching and Learning English in India – Issue 10 Using inclusive practices and multilingual approaches (2).* New Delhi: British Council.

McNeill, K. L. (2011). Elementary Students' Views of Explanation, Argumentation, and Evidence, and Their Abilities to Construct Arguments over the School Year. *Journal of Research in Science Teaching* 48 (7): 793–823.

Miller, Jon F. & Thomas Klee. (1995). Computational approaches to the analysis of language impairment. In Paul Fletcher & Brian MacWhinney (eds.). The handbook of child language, 545-572. Oxford: Blackwell.

Ministry of Education (2020). *National Education Policy 2020.* Delhi: Ministry of Human Resource Development.

Mohanty, A. (2010). 'Languages, inequality and marginalisation: implications of the double divide in Indian multilingualism'. *International Journal of the Sociology of Language.* 205: 131–154. doi:10.1515/ijsl.2010.042

Mohanty, A. K. (2013). Multilingual education in India: Overcoming the language barrier and the burden. *Multilingualism and language diversity in urban areas: Acquisition, identities, space, education. Amsterdam: John Benjamins*, 305-326.

Mohanty, A. K. (2019). The Multilingual Reality: Living with Languages. Bristol, United Kingdom: Multilingual Matters. 272 pages. ISBN: 978-1788921954

Murti, k. (1972). Telugu syntax: a brief description. *Linguistics, 10*(81), 49-82.

Nadimpalli, s. K., & Babu, K. R. V (2017). Unraveling the intricacies of tense and aspect of Telugu for an efficacious learning of English. *International Journal of Linguistics and Literature*,6 (5),23-26.

Nakamoto, J., Lindsey, K. A., & Manis, F. R. (2008). A cross-linguistic investigation of English language learners' reading comprehension in English and Spanish. *Scientific Studies of Reading, 12*(4), 351–371. https://doi.org/10.1080/10888430802378526

Nation, K., Cocksey, J., Taylor, J. S., & Bishop, D. V. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of child Psychology and Psychiatry*, *51*(9), 1031-1039.

Ntalli, A. (2020) The acquisition of finiteness in English by child second language learners in instructed contexts: age of onset and L1 effects. Unpublished PhD dissertation, University of Cambridge.

OECD (2020) Economic Outlook for Southeast Asia, China and India 2020- Update: Meeting the challenges of COVID-19. OECD Publishing, Paris. https://doi.org/10.1787/e8c90b68-en

Otwinowska, A., Mieszkowska, K., Białecka-Pikul, M., Opacki, M., & Haman, E. (2020) Retelling a model story improves the narratives of Polish-English bilingual children, International Journal of Bilingual Education and Bilingualism, 23:9, 1083-1107, DOI: 10.1080/13670050.2018.1434124

Otwinowska, A., Opacki, M., Mieszkowska, K., Białecka-Pikul, M., Wodniecka, Z., & Haman, E. (2020). Polish–English bilingual children overuse referential markers: MLU inflation in Polish-language narratives. *First Language, Special issue*,1-25, https://doi.org/10.1177%2F0142723720933769.

Ouane, A., & Glanz, C. (2010). *How and why Africa should invest in African languages and multilingual education*. UNESCO Institute for Lifelong Learning.

Pal, R & Panda, M (2020). English-Vernacular Divide: teaching learning strategies in English medium school. *MLE Newsletter: Swara,Vol (14), No.1.*

Paradis, J. (2008). Tense as a clinical marker in English L2 acquisition with language delay/impairment. *Current trends in child second language acquisition: A generative perspective*, 337-356.

Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic approaches to bilingualism*, *1*(3), 213-237.

Paradis, J., Tulpar, Y., & Arppe, A. (2016). Chinese L1 children's English L2 verb morphology over time: Individual variation in long-term outcomes. *Journal of Child Language*, *43*(3), 553-580.

Pearson, B. Z. (2002). Narrative competence among monolingual and bilingual school children in Miami. In Oller & Eilers (eds.), pp. 135–174.

Pearson, B. Z., & de Villiers, P. A. (2005). Child language acquisition: Discourse, narrative and pragmatics. In K. Brown&E. Lieven (Eds.), *Encyclopedia of language and linguistics*. Oxford: Elsevier.

Pratham (2014). *Annual Status of Education Report (Rural) 2013*. New Delhi: Pratham. Available at http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202014/fullaser2014mainreport_1.pdf

Pratham (2017). *Annual Status of Education Report (Rural) 2016*. New Delhi: Pratham. Available at http://img.asercentre.org/docs/Publications/ASER%20Reports/ASER%202016/aser_2016.pdf

Pressley, M. (2002). Metacognition and self-regulated comprehension. In A. E. Farstrup & S. J. Samuels (Eds.), What research has to say about reading instruction (pp. 291–309). Newark, DE: International Reading Association

Rice, M.L. and Wexler, K. (2001) Rice/Wexler Test of Early Grammatical Impairment. San Antonio, TX: Harcourt Assessment

Rodina, Y. (2016). Narrative Abilities of Preschool Bilingual Norwegian-Russian Children. *International Journal of Bilingualism*. doi:10.1177/1367006916643528.

Rojas, R., Iglesias, A., Bunta, F., Goldstein, B., Goldenberg, C., & Reese, L. (2016). Interlocutor differential effects on the expressive language skills of Spanish-speaking English learners. *International journal of speech-language pathology*, *18*(2), 166-177.

Romaine, S. (2013). 18 The Bilingual and Multilingual Community. *The handbook of bilingualism and multilingualism*, 445.

Rothou, K. & Tsimpli, I. M. (2017) Biliteracy and reading ability in children who learn Greek as a second language. International Journal of Bilingual Education and Bilingualism. doi:10.1080/13670050.2017.1386614

Rydland,V., & Grover,V.(2019) Argumentative peer discussions following individual reading increase comprehension, Language and Education, 33:4, 379-394, DOI: 10.1080/09500782.2018.1545786

Scarborough, H. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. Pp. 97-110 in S. B. Neuman & D. K. Dickinson (Eds.) *Handbook of Early Literacy*. NY: Guilford Press.

Seabold, S, and Perktold ,J.(2010). "statsmodels: Econometric and statistical modeling with python." Proceedings of the 9th Python in Science Conference. 2010.

Shlonsky, U. (1997). *Clause structure and word order in Hebrew and Arabic: An essay in comparative Semitic syntax*. Oxford University Press.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary children, vol. 2. *Norwood, NJ: Ablex*.

Tsimpli I. M., Peristeri E., & Andreou M. (2016). Narrative production in monolingual and bilingual children with specific language impairment. *Applied Psycholinguistics, 37*, 195–216. https://doi.org/10.1017/S0142716415000478

Tsimpli, I.M., Mukhopadhyay,L., Treffers-Daller,J., Alladi,S., Marinis, T., Panda, M., Balasubramanian,A., and Sinha,P. (2019). Multilingualism and multiliteracy in primary education in India: A discussion of some methodological challenges of an interdisciplinary research project. *Comparative & International Education, 14*(1), 54-76.

Tsimpli, I.M., Balasubramanian, A., Marinis, T., Panda, M., Mukhopadhyay, L., Alladi S., & Treffers-Daller, J.(2020a). Research Report of a four-year study of Multilingualism, Literacy, Numeracy and Cognition in Delhi, Hyderabad and Patna. *University of Cambridge.*

Tsimpli, I. M., Vogelzang, M., Balasubramanian, A., Marinis, T., Alladi, S., Reddy, A., & Panda, M. (2020b). Linguistic diversity, multilingualism and cognitive skills: A study of disadvantaged children in India. *Languages, 5*(1), 10. doi: 10.3390/languages5010010

Tsujita, Y. (2009). Deprivation of Education in Urban Areas: A Basic Profile of Slum Children in Delhi, India.

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

Uccelli, P., Galloway, E. P., Barr, C. D., Meneses, A., & Dobbs, C. L. (2015). Beyond vocabulary: Exploring cross-disciplinary academic-language proficiency and its association with reading comprehension. *Reading Research Quarterly*, *50*(3), 337–356. https://doi.org/10.1002/rrq.104

UNESCO (2009). A methodology for assessing language vitality and endangerment. http://www.unesco.org/culture/ich/index.php?pg=00142

UNESCO, 2014. World Inequality Database on Education. Available at http://www.educationinequalities.org/.

Unsworth, S. (2013). Assessing age of onset effects in (early) child L2 acquisition. *Language Acquisition, 20*(2), 74-92.

Unsworth, S. (2016). Quantity and quality of language input in bilingual language development. In E. Nicoladis & S. Montanari (Eds.), *Language and the human lifespan series. Bilingualism across the lifespan: Factors moderating language proficiency* (p. 103–121). American Psychological Association. https://doi.org/10.1037/14939-007

Van de Craats, I. (2009). The role of IS in the acquisition of finiteness by adult Turkish learners of Dutch. *Studies in Second Language Acquisition*, 59-92.

Van Gelderen, A., Schoonen, R., Stoel, R. D., De Glopper, K., & Hulstijn, J. (2007). Development of adolescent reading comprehension in language 1 and language 2: A longitudinal analysis of constituent components. *Journal of Educational Psychology*, *99*(3), 477-491. https://doi/10.1037/0022-0663.99.3.477

Vaughn, S., L. R. Martinez, J. Wanzek, G. Roberts, E. Swanson, & A. M. Fall. (2017). Improving Content Knowledge and Comprehension for English Language Learners: Findings from a Randomized Control Trial. *Journal of Educational Psychology* 109 (1): 22–34.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... & van der Walt, S. J. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, *17*(3), 261-272.

Vogelzang, M., Balasubramanian, A., Tsimpli, I.M., Panda, M., Alladi, S., Reddy, A., Mukhopadhyay,L., Treffers-Daller,J., & Marinis, T.(Submitted). Multilingualism and Multiliteracy: Examining reading skills and reading comprehension in Indian primary education

Washburn, J. S. (1983). Literature, Storytelling and Creative Drama. Paper presented at the Annual Meeting of the National Council of Teachers of English (73rd, Denver, CO, November 18-23, 1983)

White, L. (2003). *Second language acquisition and Universal Grammar*. Cambridge University Press.

Zwillinger, D., and Kokoska, S. (2000). CRC standard probability and statistics tables and formulae. Chapman & Hall: New York. 2000. Section 14.7, Crc Press.