Children's acquisition of new/given markers in English, Hindi, Mandinka and Spanish

Vishakha Shukla, Madeleine Long, Vrinda Bhatia & Paula Rubio-Fernandez (University of Oslo)
paula.rubio-fernandez@ifikk.uio.no

Languages have different ways of marking new and given referents, and these markings can be obligatory or optional. **Here we studied four typologically diverse languages (English, Hindi, Mandinka and Spanish) to test the *Optionality Hypothesis*, according to which the acquisition of optional markers is protracted relative to obligatory ones**. In Hindi, the numeral 'one' ('ek') can be used optionally to introduce new referents [1-2]. Based on diachronic and semantic evidence [3-6], we hypothesized that 'ek' is in the process of grammaticalization into an indefinite article, and that due to its current stage of optionality, the acquisition of this marker would be protracted relative to the other languages. English and Spanish require indefinite articles for character introduction, but bare nouns are more permissible in English, making the use of articles less consistent. Mandinka lacks an article system and employs a default lexical morpheme for all nouns, thus being the most consistent. As such, we predicted the following order of acquisition of discourse introduction markers: Mandinka > Spanish > English > Hindi. Given that each of these languages has obligatory markers of givenness which are mastered early [7-9], we predicted no cross-linguistic developmental differences for givenness markers.

　　　　EXP 1 employed a narrative elicitation task with a series of 14 pictures featuring 1 or 2 animal characters carrying out actions (see Fig. 1; [10]). 20 children (aged 5) and 15 adult controls were tested in each language. Of interest was the way in which children marked new and given referents relative to adults. As these languages vary in markers, we created a coding system based on the most frequently used markers in each language (see Table 1): 'A' responses are appropriate for introducing new referents, and 'B' and 'C' for marking given. For new referents, an LMER model of Marking (A=1, B & C=0) with Age Group and Language (reference level: Hindi) as FE and maximal RE structure revealed a main effect of Spanish and English relative to Hindi (p<.0001), with more A responses in Spanish and English than in Hindi. There was also an Age x Spanish interaction relative to Hindi (p=.021), driven by a difference in A responses for Hindi-speaking children and adults (p=.001). Responses from Mandinka-speaking children and adults were uniform; thus, no effects were found relative to Hindi. As predicted, Hindi-speaking children and adults differed the most, followed by English-speaking children and adults, then Spanish-speaking children and adults, and finally Mandinka-speaking children and adults (see Fig. 2). For given referents, the same LMER model revealed no Age x Language differences in the way familiar referents were marked (all p's>.05), confirming that by the age of 5 years, cross-linguistic developmental differences are not pronounced for obligatory givenness markers [7-9].

　　　　EXP 2 employed the same task, this time testing Hindi-speaking children and adults from outside of Delhi (Gorakhpur) to assess whether 'ek' is in the process of grammaticalization, or its use is simply a dialectal feature from Delhi. We tested 5-year-olds and adults from Gorakhpur, plus 10-year-olds from Gorakhpur and Delhi (because 10yos are known to have mastered even the harder discourse functions [7]). An LMER model of 'Ek' use for new referents ('Ek'=1, B & C=0) with Age Group (reference level: Adults) and Region (reference level: Delhi) as FE and maximal RE structure revealed a difference between 5-year-olds and adults (p<.0001), which disappeared by the age of 10, as 10-year-olds did not differ from adults (p=4.11) (see Fig. 3). Crucially, while Delhi adults used 'ek' more frequently than Gorakhpur 5-year-olds (p=.016), there was no difference between Delhi adults and Gorakhpur 10-year-olds (p=.287). The same pattern holds when Gorakhpur Adults is the reference level. These results suggest that the use of 'ek' for new referents is not simply dialectal variation, as similar patterns emerged in both regions.

　　　　Overall, our findings show that discourse markers emerge earlier in languages that use them consistently, and that 'ek' seems to be undergoing grammaticalization into an indefinite article. Future work should further investigate the use of 'ek' to introduce new referents to shed light on the very process of language change and its implications for common ground marking and Theory of Mind development (for discussion, see [11]).
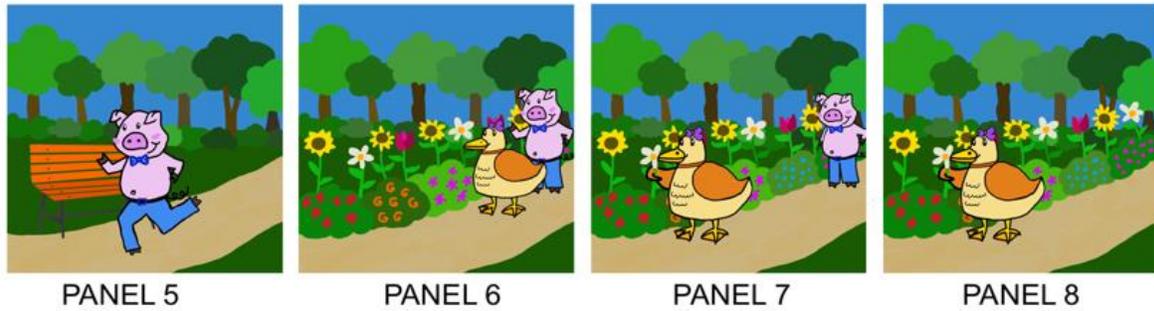
Figure 1. Panels from the narrative elicitation task used in the study, showing new and given characters. These materials were adapted from Long et al. (under review) [10].

Table 1. Coding of new and given markers in each of the languages.

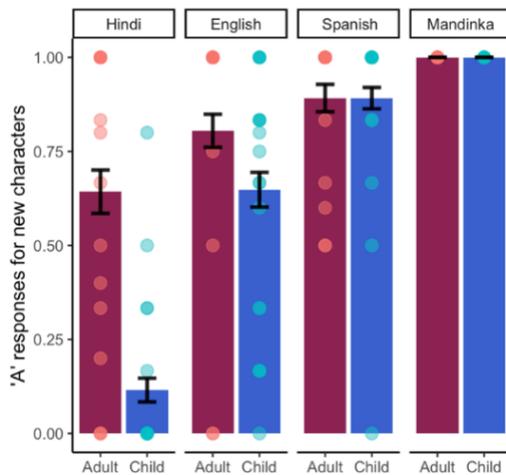| Coding	Mandinka	Hindi	English	Spanish
A	Bare noun	Numeral 'one'	Indefinite	Indefinite
B	Demonstrative + noun	Bare noun	Definite	Definite
C	Pronoun	Pronoun (overt or not)	Pronoun	Pronoun (overt or not)

Figure 2. Mean proportions of 'A' markers for new characters across ages in each of the four languages. Error bars represent 95% confidence intervals and points reflect participant means.
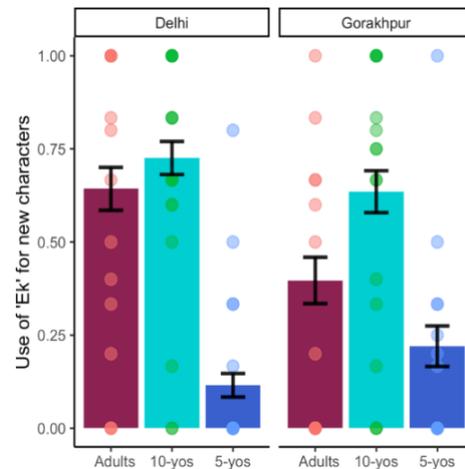

Figure 3. Proportion of 'ek' uses to introduce new discourse characters across three age groups and two regional varieties of Hindi. Error bars represent 95% confidence intervals and points reflect participant means.

References:
[1] Kachru, 2006. *Hindi.* John Benjamins **[2]** Dayal, 2018. *Trends in Hindi Linguistics.* Mouton de Gruyter **[3]** Givón, 1981. *Folia Linguistica Historica* **[4]** Heine, 1997. *Cognitive Foundations of Grammar*. OUP **[5]** Chierchia, 1998. *Natural Language Semantics* **[6]** Heine & Kuteva, 2006. *The changing languages of Europe.* OUP **[7]** Hickmann et al., 2015. *The Acquisition of Reference.* John Benjamins **[8]** Vion & Colas, 1999. *Journal of Psycholinguistic Research* **[9]** Wong & Johnston, 2004. *Journal of Child Language* **[10]** Long, Rohde, Oraa Ali & Rubio-Fernandez (under review), Speaker-internal biases in referential choice remain stable over the adult lifespan **[11]** Rubio-Fernandez, 2020. *Synthese,* SI *The Cultural Evolution of Human Social Cognition.*