# LOST IN TRANSITION: LANGUAGE ACQUISITION IN CHILDREN AND MACHINES

Haerim Hwang (The Chinese University of Hong Kong)
haerimhwang@cuhk.edu.hk

**Background:** The current study tested whether children and a large language model can acquire syntactically complex sentences, as shown in (1)–(4), whose grammaticality contrast involve a learnability challenge in language acquisition: The contrast at issue is not learnable solely through input, due to the scarcity of ungrammatical sentences like (2) and (3) in the input data (e.g., Zukowski & Larsen, 2011). Using a psycholinguistic experiment and computational modeling, this study provides new insights into the persistent discussions regarding the role of innate versus input-driven mechanisms in language acquisition.

(1)   Who + Gap:        *I wonder who you wanna work with*.
(2)   *Who + No gap: **I wonder who you wanna work*.
(3)   *If + Gap:          **I wonder if you wanna work with.*
(4)   If + No gap:        *I wonder if you wanna work*.

**Method:** The stimuli comprised 20 target sentences (e.g., (1)–(4)), which were distributed in a 2 x 2 Latin square design with the factors *Clause* (Who vs. If) and *Gap* (Gap vs. No gap), alongside 48 fillers. Human ratings for these sentences were obtained through an acceptability judgment task, where 70 adults (as a control group) and 39 children (age: 3–7) rated each sentence on a four-point Likert scale. For the large language model's acceptability ratings, log probability calculations were performed using the API of Generative Pre-trained Transformer (GPT) 3.5 (i.e., text-davinci-003).

**Results (Figure 1):** Three linear mixed-effects regression models were fitted to each of the adults' *z*-transformed acceptability scores, the children's *z*-transformed acceptability scores, and the GPT 3.5's *z*-transformed log probability scores, with *Clause* and *Gap* as fixed effects as well as *participant* (adult and child data only) and *item* as random effects. All three models showed a significant *Clause*-by-*Gap* interaction (all $p$s < .001). Post-hoc analyses for the three datasets displayed a higher acceptance rate for If + No gap than *If + Gap (all $p$s < .001). However, only in the adult and child datasets, did the analyses reveal a higher acceptance rate for Who + Gap than for *Who + No gap (all $p$s < .001); and such a difference was absent in GPT 3.5 ($p$ > .1). A further individual analysis showed that children as young as the age of 3;11, like adults, had the target grammaticality contrast.
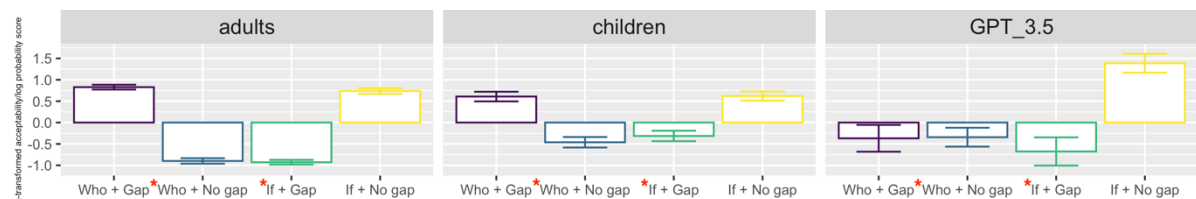


**Figure 1.** Mean *z*-transformed acceptability/log probability scores per condition and dataset

**Discussion:** A clear divide between the acceptability ratings of adults and children, on the one hand, and the log probability scores of GPT 3.5, on the other, indicates that whereas children can successfully acquire the subtle grammaticality contrast, up-to-date large language models have difficulties in inducing human-like grammatical knowledge. In light of the debates on language acquisition mechanisms, our findings seem to align with the position that the acquisition of sophisticated grammatical knowledge is not possible through statistical learning from input data alone (e.g., Chomsky, 1965).

## References

Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.

Hwang, H. (accepted). *Wanna* contraction in first language acquisition, child second language acquisition, and adult second language acquisition. *Bilingualism: Language and Cognition*.

Zukowski, A., & Larsen, J. (2011). Wanna contraction in children: Retesting and revising the developmental facts. *Language Acquisition*, *18*, 211–241.