Supplemental Material S1. Additional analysis.

Calibration of IPSyn items

Because each item in the IPSyn is scored as either 0, 1, or 2 (the maximum types scored for a given structure), we fit a graded response model (GRM; Samejima, 1969) that can handle ordered polytomous item responses. The full information maximum likelihood estimation with the Expectation-Maximization algorithm (Bock & Aitkin, 1981) implemented in the IRT software package flexMIRT (Cai, 2017) was used to obtain the model parameter estimates and cross-product standard errors. The model equation defines the conditional probability of obtaining score k given the latent ability using cumulative logit as follows:

$$P(X_{ij} = k \mid \theta_i = \theta) = P(X_{ij} \geq k \mid \theta_i = \theta) - P(X_{ij} \geq k+1 \mid \theta_i = \theta), \text{ where}$$
$$P(X_{ij} \geq k \mid \theta_i = \theta) = \frac{1}{1+\exp[-(a_j \theta_i + c_{jk})]}.$$

In the equation above, θ_i indicates the latent ability of an individual i. a_j is the item slope or discrimination parameter for the jth item, and c_{jk} is the kth intercept for item j. Note that the distribution of latent variable θ in a population is assumed to be normally distributed with mean equal to 0 and variance equal to 1 for model identification purpose. More specifically, for IPSyn analysis, the unidimensional θ represents the latent trait that explains the observed item responses for each subscale or for the overall scale. If we use item responses that pertain only to Noun Phrase items, θ is defined as the latent trait (we can call it Noun Phrase Assembly Ability) that explains the observed item responses for those items. Similarly, if we use all of the item responses from all four subscales, the latent trait is defined as the Overall Productive Syntax score.

Turning to the item parameters, a larger a_j means that the item is more sensitive to the levels of latent ability, and is, therefore, more highly associated with the latent ability. Note that a positive, larger a_j is considered to be a desired property of an item; however, when this estimate is too large (i.e., larger than 5), the item slope is too steep, and the item behaves like a Guttman-like-phenomenon (e.g., Guttman, 1944). In the IRT analysis, this phenomenon happens when the item response can be almost completely explained by the latent trait. In other words, the slope is close to infinite and this phenomenon is called as quasi-complete separation (Zeng & Zeng, 2019). In such cases, the standard error estimate is not trustworthy as the model estimation itself is unstable. We go to some effort to explain this phenomenon here, as we observed that some IPSyn items exhibited this feature due to the redundant information they provided. In lay terms, some scoring categories on the IPSyn, while they can be described in unique phrase structure formulas, are almost never observed in the absence of other structures. Thus, they do not add to the informativeness of a child's score, while adding to scoring time (and associated expertise). Relevant to the internal organization of IPSyn structures, in the GRM, the category (specific score 0, 1, and 2 for IPSyn) intercepts c_{jk} are strictly ordered, so that the higher latent trait level is required to reach the threshold (0.5 probability) in obtaining each score category or items ranked above them.

References

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika, 46*(4), 443–459. https://doi.org/10.1007/BF02293801

Cai, L. (2017*). flexMIRT®: Flexible multilevel multidimensional item analysis and test scoring* (Version 3.51) [Computer software]. Vector Psychometric Group.

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*(2), 139–150. https://doi.org/10.2307/2086306

Samejima, F. (1969). *Estimation of latent ability using a response pattern of graded score* (Psychometric Monograph No. 17). Psychometric Society. http://www.psychometrika.org/journal/online/MN17.pdf

Zeng, G., & Zeng, E. On the relationship between multicollinearity and separation in logistic regression. *Communications in Statistics - Simulation and Computation, 50*(7), 1989 –1997. https://doi.org/10.1080/03610918.2019.1589511