a. Prequential code length of languages

$K(z_i | w_i; \theta_{i-1})$ vs Data index i

normal

iterated learning
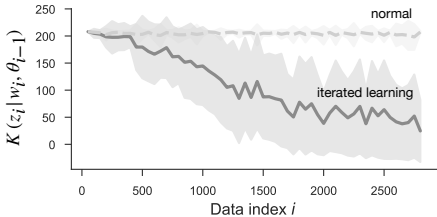
b. Compositionality of languages

Compositionality vs Training method (Normal, Iterated learning)

Topological similarity

$C(Z) = \dfrac{K(Z)}{K(Z \mid W)}$

$Topsim(Z, W)$