B: Train last speaker $\,\mathcal{S}_t$ with new meta-listener \mathcal{L}_t^m and dataset \mathcal{D} A:Train meta-listener \mathcal{L}_{t-1}^m with speakers in the buffer $\mathcal{B}_{\mathcal{S}}^t$ It is Red Sphere. S_{t-1} \mathcal{L}_{t-1}^{m} It looks like blue cylinder. This is a Green Cone. - It is a Blue Cylinder. - This is blue cylinder. - This seems like a Blue Cylinder. C: Add new speaker $\,\mathcal{S}_{t+1}$ to the buffer $\,\mathcal{B}^{t+1}_{\mathcal{S}}$ Self-play updates Supervised updates Meta updates Parameter copy Speaker-Input/Listener-Output Agent parameter frozen Speaker utterance