**Effect of Prior Factor on Training Perplexity** Prior Factor = 0.0625Prior Factor = 0.125Prior Factor = 0.25 Training Perplexity (Log Scale) Prior Factor = 0.510<sup>1</sup> 100 200 300 400 500 600 700 **Training Step**