Vocabulary Growth Curves 1000 1200 1400 V (types) observed expected N (k tokens)