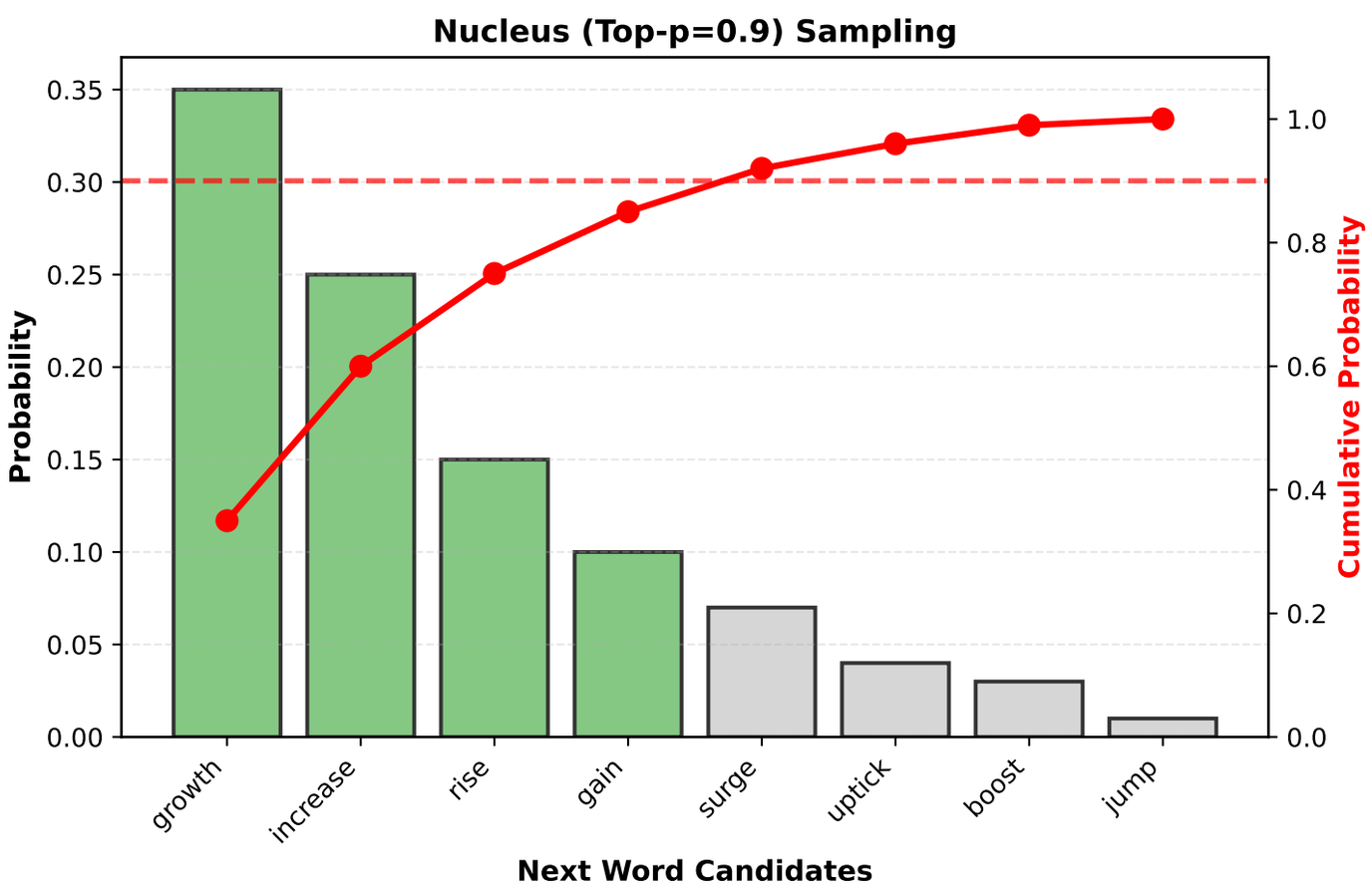

Top-p Algorithm:

1. Sort words by probability
2. Compute cumulative sum
3. Include words until sum ≥ p (0.9)
4. Sample from included set

Result: Dynamic vocabulary size
- Peaked distribution → few words
- Flat distribution → many words

Included (green): Top 90% probability
Excluded (gray): Bottom 10% (too unlikely)