Normal run with 10000 steps:

Iter= 1000, Average Loss= 3.275797, Average Accuracy= 21.30%

['she', 'decides', 'to'] - [keep] vs [decides]

Iter= 2000, Average Loss= 1.368491, Average Accuracy= 57.60%

['enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 3000, Average Loss= 1.108723, Average Accuracy= 66.00%

['home,', 'it', 'keeps'] - [changing.] vs [it]

Iter= 4000, Average Loss= 0.895450, Average Accuracy= 70.40%

['luck,', 'she', 'decides'] - [to] vs [to]

Iter= 5000, Average Loss= 0.874068, Average Accuracy= 70.20%

['taking', 'it', 'home,'] - [it] vs [changing.]

Iter= 6000, Average Loss= 0.768973, Average Accuracy= 73.90%

['to', 'keep', 'it.'] - [As] vs [As]

Iter= 7000, Average Loss= 0.656851, Average Accuracy= 77.60%

['is', 'taking', 'it'] - [home,] vs [home,]

Iter= 8000, Average Loss= 0.584527, Average Accuracy= 79.30%

['she', 'decides', 'to'] - [keep] vs [keep]

Iter= 9000, Average Loss= 0.522173, Average Accuracy= 80.70%

['of', 'treasure', 'on'] - [the] vs [on]

Iter= 10000, Average Loss= 0.488772, Average Accuracy= 80.80%

['it', 'keeps', 'changing.'] - [However,] vs [However,]

Optimization Finished!

Elapsed time: 1.839347251256307 min

Run on command line.

3 words: on the road

on the road while she is returning from work. her luck, she decides to keep it. As she is taking it home, it keeps it keeps it keeps it keeps it keeps it keeps it

**n\_input=4 and n\_hidden=256**

Iter= 1000, Average Loss= 2.310862, Average Accuracy= 36.70%

['keep', 'it.', 'As', 'she'] - [is] vs [is]

Iter= 2000, Average Loss= 1.132206, Average Accuracy= 59.70%

['work.', 'Delighted', 'with', 'her'] - [luck,] vs [is]

Iter= 3000, Average Loss= 0.703610, Average Accuracy= 72.80%

['A', 'woman', 'finds', 'a'] - [pot] vs [pot]

Iter= 4000, Average Loss= 0.537745, Average Accuracy= 77.00%

['keeps', 'changing.', 'However,', 'her'] - [enthusiasm] vs [enthusiasm]

Iter= 5000, Average Loss= 0.422844, Average Accuracy= 83.30%

['changing.', 'However,', 'her', 'enthusiasm'] - [refuses] vs [away.]

Iter= 6000, Average Loss= 0.314930, Average Accuracy= 87.50%

['However,', 'her', 'enthusiasm', 'refuses'] - [to] vs [to]

Iter= 7000, Average Loss= 0.275661, Average Accuracy= 88.60%

['her', 'enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 8000, Average Loss= 0.251069, Average Accuracy= 90.50%

['her', 'enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 9000, Average Loss= 0.238427, Average Accuracy= 90.00%

['pot', 'of', 'treasure', 'on'] - [the] vs [the]

Iter= 10000, Average Loss= 0.174094, Average Accuracy= 94.40%

['a', 'pot', 'of', 'treasure'] - [on] vs [on]

Optimization Finished!

Elapsed time: 33.47564363479614 sec

Run on command line.

4 words: on the road while

on the road while while she is returning from work. Delighted with with with with with with with with with with with with with with with with with with with with with with with with with

**n\_input=2 n\_hidden=512**

Iter= 1000, Average Loss= 3.405565, Average Accuracy= 5.30%

['the', 'road'] - [while] vs [work.]

Iter= 2000, Average Loss= 2.147785, Average Accuracy= 15.40%

['treasure', 'on'] - [the] vs [of]

Iter= 3000, Average Loss= 1.790865, Average Accuracy= 31.80%

['changing.', 'However,'] - [her] vs [to]

Iter= 4000, Average Loss= 1.698858, Average Accuracy= 35.00%

['her', 'luck,'] - [she] vs [she]

Iter= 5000, Average Loss= 1.554584, Average Accuracy= 42.40%

['As', 'she'] - [is] vs [decides]

Iter= 6000, Average Loss= 1.354635, Average Accuracy= 50.60%

['refuses', 'to'] - [fade] vs [enthusiasm]

Iter= 7000, Average Loss= 1.230746, Average Accuracy= 53.10%

['to', 'keep'] - [it.] vs [it.]

Iter= 8000, Average Loss= 1.132010, Average Accuracy= 58.20%

['enthusiasm', 'refuses'] - [to] vs [to]

Iter= 9000, Average Loss= 1.073296, Average Accuracy= 60.70%

['finds', 'a'] - [pot] vs [pot]

Iter= 10000, Average Loss= 1.006150, Average Accuracy= 64.00%

['her', 'luck,'] - [she] vs [she]

Optimization Finished!

Elapsed time: 1.4351339181264242 min

Run on command line.

2 words: the road

the road she is taking it keeps changing. her keeps she is taking it keeps changing. her keeps she is taking it keeps changing. her keeps she is taking it keeps changing. her keeps

**n\_input=4 n\_hidden=512**

Iter= 1000, Average Loss= 3.068345, Average Accuracy= 36.70%

['from', 'work.', 'Delighted', 'with'] - [her] vs [she]

Iter= 2000, Average Loss= 1.041878, Average Accuracy= 66.30%

['it', 'home,', 'it', 'keeps'] - [changing.] vs [changing.]

Iter= 3000, Average Loss= 0.737017, Average Accuracy= 73.80%

['her', 'enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 4000, Average Loss= 0.521668, Average Accuracy= 82.20%

['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]

Iter= 5000, Average Loss= 0.479517, Average Accuracy= 83.30%

['As', 'she', 'is', 'taking'] - [it] vs [it]

Iter= 6000, Average Loss= 0.393388, Average Accuracy= 86.60%

['she', 'is', 'taking', 'it'] - [home,] vs [home,]

Iter= 7000, Average Loss= 0.365448, Average Accuracy= 88.20%

['enthusiasm', 'refuses', 'to', 'fade'] - [away.] vs [away.]

Iter= 8000, Average Loss= 0.326776, Average Accuracy= 88.80%

['it.', 'As', 'she', 'is'] - [taking] vs [taking]

Iter= 9000, Average Loss= 0.264050, Average Accuracy= 89.90%

['she', 'decides', 'to', 'keep'] - [it.] vs [it.]

Iter= 10000, Average Loss= 0.201014, Average Accuracy= 93.10%

['taking', 'it', 'home,', 'it'] - [keeps] vs [keeps]

Optimization Finished!

Elapsed time: 2.211258355776469 min

Run on command line.

4 words: on the road while

on the road while while she is returning from work. Delighted with with her luck, she decides to keep it. As she is taking it home, it keeps changing. However, her enthusiasm refuses to fade away.

n\_input=3 and n\_hidden=1024

Iter= 1000, Average Loss= 4.350561, Average Accuracy= 22.40%

['is', 'taking', 'it'] - [home,] vs [keep]

Iter= 2000, Average Loss= 1.499837, Average Accuracy= 55.40%

['road', 'while', 'she'] - [is] vs [is]

Iter= 3000, Average Loss= 1.099754, Average Accuracy= 66.10%

['she', 'is', 'taking'] - [it] vs [it]

Iter= 4000, Average Loss= 1.012846, Average Accuracy= 69.50%

['luck,', 'she', 'decides'] - [to] vs [to]

Iter= 5000, Average Loss= 0.928940, Average Accuracy= 70.50%

['while', 'she', 'is'] - [returning] vs [returning]

Iter= 6000, Average Loss= 0.789392, Average Accuracy= 73.80%

['it.', 'As', 'she'] - [is] vs [is]

Iter= 7000, Average Loss= 0.757151, Average Accuracy= 74.00%

['A', 'woman', 'finds'] - [a] vs [a]

Iter= 8000, Average Loss= 0.596692, Average Accuracy= 79.20%

['enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 9000, Average Loss= 0.561144, Average Accuracy= 79.60%

['luck,', 'she', 'decides'] - [to] vs [to]

Iter= 10000, Average Loss= 0.416345, Average Accuracy= 85.60%

['keeps', 'changing.', 'However,'] - [her] vs [her]

Optimization Finished!

Elapsed time: 8.135157306989035 min

Run on command line.

3 words: on the road

on the road she is returning from work. with with with her luck, she decides to keep it. As her enthusiasm refuses to fade away. to fade away. to fade away. to fade away. to

**TWO STACK LSTM**

Iter= 1000, Average Loss= 2.599633, Average Accuracy= 24.70%

['woman', 'finds', 'a'] - [pot] vs [treasure]

Iter= 2000, Average Loss= 1.439252, Average Accuracy= 51.00%

['it.', 'As', 'she'] - [is] vs [is]

Iter= 3000, Average Loss= 1.045898, Average Accuracy= 62.40%

['changing.', 'However,', 'her'] - [enthusiasm] vs [enthusiasm]

Iter= 4000, Average Loss= 0.872506, Average Accuracy= 68.50%

['Delighted', 'with', 'her'] - [luck,] vs [luck,]

Iter= 5000, Average Loss= 0.702272, Average Accuracy= 75.30%

['decides', 'to', 'keep'] - [it.] vs [it.]

Iter= 6000, Average Loss= 0.652505, Average Accuracy= 77.20%

['A', 'woman', 'finds'] - [a] vs [a]

Iter= 7000, Average Loss= 0.593401, Average Accuracy= 79.90%

['work.', 'Delighted', 'with'] - [her] vs [her]

Iter= 8000, Average Loss= 0.536841, Average Accuracy= 82.70%

['while', 'she', 'is'] - [returning] vs [returning]

Iter= 9000, Average Loss= 0.542417, Average Accuracy= 81.90%

['it', 'home,', 'it'] - [keeps] vs [keeps]

Iter= 10000, Average Loss= 0.415996, Average Accuracy= 83.60%

['decides', 'to', 'keep'] - [it.] vs [it.]

Optimization Finished!

Elapsed time: 5.426478584607442 min

Run on command line.

3 words: on the road

on the road she is returning from work. Delighted Delighted her luck, she decides to keep it. As her enthusiasm refuses to fade away. As her enthusiasm refuses to fade away. As her enthusiasm refuses.

n\_input-4 and n\_hidden=512

Iter= 1000, Average Loss= 2.376754, Average Accuracy= 40.90%

['of', 'treasure', 'on', 'the'] - [road] vs [road]

Iter= 2000, Average Loss= 1.053653, Average Accuracy= 66.90%

['work.', 'Delighted', 'with', 'her'] - [luck,] vs [is]

Iter= 3000, Average Loss= 0.658652, Average Accuracy= 76.70%

['her', 'enthusiasm', 'refuses', 'to'] - [fade] vs [fade]

Iter= 4000, Average Loss= 0.488245, Average Accuracy= 80.80%

['is', 'taking', 'it', 'home,'] - [it] vs [it]

Iter= 5000, Average Loss= 0.400565, Average Accuracy= 85.10%

['keep', 'it.', 'As', 'she'] - [is] vs [is]

Iter= 6000, Average Loss= 0.355035, Average Accuracy= 87.80%

['Delighted', 'with', 'her', 'luck,'] - [she] vs [she]

Iter= 7000, Average Loss= 0.293030, Average Accuracy= 89.00%

['decides', 'to', 'keep', 'it.'] - [As] vs [As]

Iter= 8000, Average Loss= 0.281811, Average Accuracy= 90.70%

['keep', 'it.', 'As', 'she'] - [is] vs [is]

Iter= 9000, Average Loss= 0.210149, Average Accuracy= 92.90%

['changing.', 'However,', 'her', 'enthusiasm'] - [refuses] vs [refuses]

Iter= 10000, Average Loss= 0.167770, Average Accuracy= 94.60%

['enthusiasm', 'refuses', 'to', 'fade'] - [away.] vs [away.]

Optimization Finished!

Elapsed time: 6.469173622131348 min

Run on command line.

4 words: on the road while

on the road while while she is returning from work. Delighted her luck, she decides to keep it. As she is taking it home, it keeps changing. However, her enthusiasm refuses to fade away. However, her

**RMS optimizer:**

**Iter= 1000, Average Loss= 2.997751, Average Accuracy= 36.90%**

**['decides', 'to', 'keep', 'it.'] - [As] vs [As]**

**Iter= 2000, Average Loss= 1.208357, Average Accuracy= 59.30%**

**['changing.', 'However,', 'her', 'enthusiasm'] - [refuses] vs [refuses]**

**Iter= 3000, Average Loss= 0.763443, Average Accuracy= 72.30%**

**['it', 'keeps', 'changing.', 'However,'] - [her] vs [her]**

**Iter= 4000, Average Loss= 0.565441, Average Accuracy= 80.60%**

**['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]**

**Iter= 5000, Average Loss= 0.474172, Average Accuracy= 82.90%**

**['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]**

**Iter= 6000, Average Loss= 0.428448, Average Accuracy= 84.80%**

**['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]**

**Iter= 7000, Average Loss= 0.327455, Average Accuracy= 88.90%**

**['luck,', 'she', 'decides', 'to'] - [keep] vs [keep]**

**Iter= 8000, Average Loss= 0.284442, Average Accuracy= 89.10%**

**['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]**

**Iter= 9000, Average Loss= 0.254244, Average Accuracy= 91.20%**

**['taking', 'it', 'home,', 'it'] - [keeps] vs [keeps]**

**Iter= 10000, Average Loss= 0.202557, Average Accuracy= 92.40%**

**['work.', 'Delighted', 'with', 'her'] - [luck,] vs [luck,]**

**Optimization Finished!**

**Elapsed time: 2.2733872294425965 min**

**Run on command line.**

**4 words: on the road while**

**on the road while she is returning from work. Delighted her luck, she decides to keep it. As she is taking it home, it keeps changing. However, her enthusiasm refuses to fade away. However, her enthusiasm**

**ADAGRAd:**

**Iter= 10000, Average Loss= 1.025945, Average Accuracy= 67.70%**

**['she', 'is', 'returning', 'from'] - [work.] vs [work.]**

**Optimization Finished!**

**Elapsed time: 2.130667726198832 min**

**Run on command line.**

**4 words: on the road while**

**on the road while road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road**

**adam:**

**Iter= 10000, Average Loss= 0.108103, Average Accuracy= 96.10%**

**['home,', 'it', 'keeps', 'changing.'] - [However,] vs [However,]**

**Optimization Finished!**

**Elapsed time: 2.303109681606293 min**

**Run on command line.**

**4 words: on the road while**

**on the road while she is returning from work. Delighted with with with with her luck, she decides to keep it. As she is taking it home, it keeps changing. However, her enthusiasm refuses to fade**

**Gradient:**

**Iter= 1000, Average Loss= 3.735601, Average Accuracy= 24.40%**

**['decides', 'to', 'keep', 'it.'] - [As] vs [However,]**

**Iter= 2000, Average Loss= 1.688878, Average Accuracy= 48.10%**

**['with', 'her', 'luck,', 'she'] - [decides] vs [keep]**

**Iter= 3000, Average Loss= 1.302090, Average Accuracy= 56.10%**

**['she', 'decides', 'to', 'keep'] - [it.] vs [to]**

**Iter= 4000, Average Loss= 0.993967, Average Accuracy= 67.70%**

**['her', 'enthusiasm', 'refuses', 'to'] - [fade] vs [fade]**

**Iter= 5000, Average Loss= 0.967114, Average Accuracy= 67.90%**

**['taking', 'it', 'home,', 'it'] - [keeps] vs [keeps]**

**Iter= 6000, Average Loss= 0.892071, Average Accuracy= 68.60%**

**['work.', 'Delighted', 'with', 'her'] - [luck,] vs [luck,]**

**Iter= 7000, Average Loss= 0.831983, Average Accuracy= 72.10%**

**['she', 'decides', 'to', 'keep'] - [it.] vs [it.]**

**Iter= 8000, Average Loss= 0.780017, Average Accuracy= 72.00%**

**['while', 'she', 'is', 'returning'] - [from] vs [from]**

**Iter= 9000, Average Loss= 0.671636, Average Accuracy= 77.20%**

**['is', 'taking', 'it', 'home,'] - [it] vs [it.]**

**Iter= 10000, Average Loss= 0.653625, Average Accuracy= 76.30%**

**['As', 'she', 'is', 'taking'] - [it] vs [it]**

**Optimization Finished!**

**Elapsed time: 1.8860342502593994 min**

**Run on command line.**

**4 words: on the road while**

**on the road while road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road road**