

Merve Horuz

1801042651

CSE 654/484 hw3 report

**Word2Vec** is a technique for generating dense, continuous-valued vector representations of words, which can capture semantic and syntactic relationships between words in a dataset. The vectors can be used to compute the similarity between words, using a measure such as cosine similarity.

**The similarity function** is a measure of how similar two words are to each other based on their vector representation. There are various similarity functions that can be used, such as cosine similarity, Euclidean distance, and Pearson correlation coefficient.

**The most\_similar function** is a method that takes in a word and a list of words and returns the most similar word in the list based on the similarity function being used. For example, if the similarity function is cosine similarity and the input word is "cat", the most\_similar function might return "dog" as the most similar word in the list, because "cat" and "dog" are often used in similar contexts and have a high cosine similarity.

- Used 'ngrams' library from 'nltk' to create n-grams.

First, created word embeddings (word vectors). Some example is here:

##### syllable vectors for uni gram#####

```
##### Syntactic vectors for all grammar #####
'fil'
[ 5.0378526e-03 -1.6850602e-03 -3.9510429e-03  9.0912817e-04
 6.3125286e-03 -1.9788230e-03 -6.5968190e-03  6.6759526e-03
 5.9973858e-03  8.9686718e-03  5.3017940e-03  4.2754724e-03
 9.4780028e-03 -4.2331181e-03  7.1711410e-03 -7.0763435e-03
-2.7173616e-03 -2.0851993e-03  3.9772224e-03 -8.0750398e-03
 5.1702191e-03 -9.3283225e-03  1.5515447e-04 -2.6025652e-04
-9.0682963e-03 -8.7846890e-03  3.5692407e-03  6.2696827e-03
 7.4217079e-04  2.4604751e-03  5.4428647e-03 -7.3542343e-03
-6.5751686e-03 -1.3910532e-05 -1.6870057e-03 -2.2251583e-03
 3.0444635e-03  5.5322528e-04 -8.4515866e-03  6.4620459e-03
 1.0260439e-03 -9.6053556e-03 -3.3680045e-03 -3.1688667e-03
-7.2343326e-03  5.3124665e-04 -2.7976299e-03  2.2477687e-03
-4.4852290e-03 -8.6251856e-04  7.5097024e-03 -9.4790366e-03
-1.6173518e-03  1.5786863e-03 -9.9260323e-03  7.3003052e-03
-1.2203264e-03 -8.4454212e-03 -8.2941242e-03  8.7321308e-03
 3.7727796e-03  9.8000709e-03 -4.3108501e-03 -9.5664151e-03
-6.3341381e-03  1.4187288e-03 -2.3612738e-04 -7.6901910e-05
-1.3371133e-03  9.1843260e-03 -2.6578904e-04 -2.7981996e-05
-2.1420598e-04  8.6938404e-03  5.7432437e-03 -3.8001800e-03
-2.9666806e-03 -6.3951756e-03 -5.1868916e-04  5.7993354e-03
-9.5287142e-03 -8.4573147e-04 -4.3754019e-03 -3.5587740e-03
-2.5917888e-03 -4.5560584e-03  1.3219833e-03 -5.8313562e-03
-9.7628478e-03  8.8319024e-03  8.7499181e-03  3.7152385e-03
 5.9779952e-03 -6.8627298e-03 -6.3541653e-03  3.6373914e-03
-3.0190467e-03  5.6580319e-03  4.1870298e-03  9.0407161e-03]
```

```
'min'
[-2.68333941e-03  6.77283388e-03  7.18137203e-03  1.06962612e-02
 6.91156415e-03 -9.00309347e-03  3.93758016e-03  7.30899815e-03
-3.76902823e-03  5.17717376e-03  3.57770780e-03  7.31618609e-03
-7.16296118e-03 -3.37856007e-03 -4.98507498e-03 -9.04858671e-03
 2.28542811e-03  3.41216777e-03 -5.96935721e-03 -1.15477573e-02
 1.02352938e-02  6.37942459e-03  1.25894637e-03  7.29419710e-03
 1.08766621e-02 -4.25657537e-03  1.52154302e-03 -7.68658938e-03
 5.62928896e-03 -8.43717530e-03 -4.44186386e-03  6.46189507e-03
-1.05880490e-02 -5.91326738e-03  6.96768472e-03 -5.35813300e-03
 5.94695518e-03  5.24796022e-04 -1.14370259e-02 -1.39615331e-02
 1.00861294e-02 -8.82314611e-03 -3.86993610e-03 -5.57065476e-03
-9.96895600e-03  7.38819654e-04 -8.47607665e-03  3.71787464e-03
 3.97433434e-03 -6.11918373e-03  5.10485470e-03 -8.31968896e-03
-8.92826449e-03  5.85969118e-03 -1.10790050e-02  5.86431660e-03
-4.99016698e-03  6.53971685e-03  6.78177504e-03 -4.29617660e-03
 7.55167194e-03 -6.12620264e-03 -5.85963862e-05 -1.14333797e-02
-5.62232593e-03 -6.96725026e-03  1.27320942e-02 -2.53275782e-03
-1.18453419e-02 -5.87502448e-03  4.97272098e-03  5.08960383e-03
-1.90425629e-03 -8.25169403e-03  2.73387413e-04  9.58480127e-03
-1.01186205e-02  7.79502001e-03 -6.18447876e-03  9.94898472e-03
-4.62071458e-03  5.45501616e-03  7.74937263e-03  5.52632986e-03
-1.04173180e-02  3.99299944e-03 -2.15812982e-03  9.24237911e-03
 4.47627855e-03 -3.54458764e-03 -1.92186286e-04 -2.78357998e-03
 2.78043328e-04 -8.13608617e-03  1.20057473e-02  9.88732651e-03
 9.19656188e-04 -5.12638502e-03  8.75739940e-03 -6.20612968e-03]
```

'mü'

```
[ -3.47381714e-03  8.26063287e-03  1.33770695e-02  1.65748615e-02
  4.86517820e-05 -3.28813330e-03  9.68110841e-03  3.95050738e-03
 -1.85173787e-02  1.31896501e-02  1.13700507e-02  4.84996941e-03
 -8.82875919e-03  8.83083511e-03  1.09992800e-02  4.10977704e-03
  9.72628221e-03  3.95856390e-04  2.13503535e-03 -7.24693248e-03
  3.32237105e-03  4.80623962e-03 -6.63465937e-04  4.09364700e-03
 -2.30362033e-03  3.94687988e-03 -1.08774239e-02 -1.30822137e-02
  3.42518534e-03  9.01377574e-03  1.25582097e-02 -9.44863725e-03
 -2.22266489e-03  8.58810323e-04  5.64477127e-03  1.44489817e-02
  4.06944472e-03 -3.34438775e-03 -2.09343582e-02 -1.34671526e-02
  4.68941312e-03 -2.42467895e-02 -5.16869640e-03  1.25141684e-02
 -1.09983133e-02 -2.70020054e-03 -1.50627578e-02 -1.53843835e-02
 -5.21023350e-04  5.65941865e-03  6.35245536e-03 -1.24798305e-02
  6.04891870e-03  6.16233854e-04 -2.28021555e-02 -8.84875376e-03
 -5.69540728e-03 -2.46907701e-03 -2.49861553e-03  3.10326897e-04
 -8.99954815e-04 -1.89901039e-03 -9.00150277e-03 -1.24354055e-02
 -1.02212629e-03 -2.32684799e-03  1.77823957e-02  1.75440237e-02
 -1.85399223e-02  7.93252699e-03 -1.40303522e-02  4.61637136e-03
 -7.04889360e-04 -1.41539425e-03  1.00761317e-02  1.33105237e-02
 -1.21780597e-02  7.20581273e-03  7.46004982e-03  1.10622058e-02
 -1.49635049e-02 -5.87651785e-03 -7.93481246e-04  1.53833795e-02
 -9.64324363e-03 -2.75074970e-04  5.31807868e-03  8.92759301e-03
  9.85741802e-03  2.11093412e-03  1.93281658e-02  8.00563674e-03
  3.03963874e-03  9.38020274e-03  1.82991382e-02  6.99936179e-03
  1.49276550e-03 -7.19117746e-03  1.22558884e-03 -7.41179707e-03]
```

'zik'

```
[ -9.4428584e-03  4.4573452e-03 -3.0342482e-03  8.4774327e-03
 -1.2690068e-04  6.3967276e-03 -4.0253685e-03  9.8239898e-04
 -1.1122703e-04  2.4561023e-03  9.3161650e-03 -6.6320016e-03
  7.0750988e-03  2.9600870e-03  8.3875144e-03 -7.4688257e-03
 -5.1884828e-03  9.5645310e-03 -8.0704186e-03  6.3978028e-03
  6.2934039e-03  2.7734053e-03  5.8713341e-03  6.5580760e-03
  9.4512319e-03 -4.2901994e-03  1.8664480e-03 -3.8939023e-03
 -3.8864612e-04 -8.2299504e-03  8.1936214e-03 -6.0531404e-03
  7.1778237e-03 -7.5237309e-03 -3.2418894e-03 -2.2175873e-03
  8.0743348e-03 -2.6935542e-03 -1.8330455e-04  8.4495591e-03
  9.0996263e-04  8.3678570e-03 -4.6373606e-03 -1.0799586e-03
  3.9929366e-03 -1.8709612e-03 -4.3043494e-04 -7.7580211e-05
  1.1204160e-03  5.2519892e-03  6.6574500e-03  4.7798217e-03
 -4.0208469e-03 -4.3776989e-04  7.6228858e-04  3.8250780e-03
 -8.7549305e-03  3.4348464e-03  8.4799016e-03 -5.6824074e-03
 -3.7366832e-03  2.5834751e-03  2.9378510e-03 -9.9649122e-03
  6.2108384e-03 -3.8401450e-03  2.6292694e-03 -9.3298936e-03
  4.8818756e-03 -7.5629232e-03 -2.9198492e-03  7.6259007e-03
 -7.8152660e-03 -9.4074849e-03 -9.4593074e-03 -7.6040495e-03
  6.4997780e-03 -8.0134962e-03 -1.5307820e-03 -1.9195951e-03
  5.4195118e-03 -4.8397230e-03  4.4370140e-03 -6.7353011e-03
 -3.2026256e-03 -4.2165937e-03 -2.6423932e-04 -6.5023755e-03
 -2.1236909e-03  2.8221749e-03  8.8256430e-03 -6.5178405e-03
  2.6921201e-03 -4.8278226e-03  8.7657021e-03  2.1452273e-03
  6.3136588e-03 -3.5132624e-03  2.2422313e-03 -2.1226644e-03]
```



"##### syllable vectors for bi gram#####

```
lelody@lelody:~/Documents/Python/Projects/lelody_nlp/nlp$ python3
'e lekt'
[ 3.8256668e-03  9.6832626e-03 -1.4268542e-03 -9.3647288e-03
 6.9340980e-03  7.9287896e-03  6.5564383e-03  6.1862026e-03
 4.5674536e-03  8.2222931e-03 -2.2545445e-03  3.4797050e-03
 2.7797592e-03 -1.1863804e-03  5.7199681e-03  1.8080950e-03
 8.9477524e-03  6.8569016e-03  8.1716869e-03  4.8104096e-03
 -4.9125422e-03  7.4925171e-03  4.0231622e-03 -9.4495025e-03
 5.3485953e-03  1.0490108e-03  2.8889359e-03  7.4366285e-03
 -3.4594547e-03 -7.7020528e-04  6.7384900e-03 -9.5400978e-03
 -9.3453266e-03 -7.2950828e-03  1.5164852e-03  8.1528220e-03
 -7.2718691e-03 -4.3858243e-03  4.2307391e-03  2.0038057e-03
 8.0596730e-03 -8.6289933e-03 -2.0423126e-03 -2.0954954e-03
 -9.7546186e-03  6.2839617e-03  9.2913341e-03  1.5408122e-03
 -6.2772916e-03 -6.2272563e-03  6.8357587e-03 -3.8103224e-03
 9.5961224e-03  6.0323713e-04 -6.9936871e-04  6.4442791e-03
 -4.6999548e-03  3.3570146e-03 -6.4424695e-03  9.8381341e-03
 5.4238606e-03 -9.4407536e-03 -7.1742986e-03 -9.8457476e-03
 9.5216418e-03  7.6652528e-03 -1.5542328e-03  6.5616844e-03
 7.7385008e-03 -2.2981106e-03  2.7331519e-03  2.6643204e-03
 5.9150886e-03  5.4284991e-03  9.0994836e-05  1.9899727e-04
 -5.1329494e-04  6.5469206e-03  1.7160154e-03 -8.1583271e-03
 -8.1837121e-03  3.1463527e-03  5.3243281e-04 -6.9872402e-03
 1.9835746e-03  5.5175317e-03  7.3206928e-03  7.4034538e-03
 6.6654421e-03 -1.9810975e-03  9.2195878e-03  8.2291206e-03
 -7.1965908e-03  3.1469238e-03 -8.3189858e-03 -8.1567168e-03
 3.4938729e-03 -4.8846258e-03 -2.0657706e-03 -5.4090903e-03]
```

```
'ge len'
[-0.00525184 -0.00360843  0.00927302  0.00214908 -0.00151456 -0.00236027
 -0.00134047  0.00625269  0.0028843  -0.00382252 -0.00856215 -0.00144526
 -0.00600497  0.00554366 -0.00172096  0.00281822 -0.00722744  0.00633013
  0.00134124 -0.00869615  0.00110226  0.00991389 -0.00667128 -0.00742828
  0.00764084  0.00310373  0.00869802  0.00867926  0.00885236  0.00097672
 -0.00139281  0.00821084  0.006478  0.00944719  0.00946532 -0.00552207
  0.00289769  0.00212886  0.00390586  0.0077577  0.00953291  0.00276993
  0.00460209 -0.0022182  0.00677854 -0.00920016  0.00657752  0.0099577
 -0.00798871 -0.00559093 -0.00850869 -0.00327938 -0.00191037  0.00408867
 -0.00758352 -0.00130616 -0.00656428  0.00686749 -0.00867557 -0.00526148
  0.00810323 -0.00146535 -0.00032161  0.00205183  0.00470011 -0.00942035
  0.00365619  0.00855224  0.00340486  0.00262577 -0.00437042  0.00065664
 -0.00253686  0.00640386 -0.00294695  0.00814409 -0.00126742  0.0064801
 -0.00926746  0.00483508  0.00038146  0.00996467  0.0072539  -0.00513269
 -0.00841788  0.00472146 -0.00407305 -0.00814938  0.0042584  -0.00384067
 -0.00250383 -0.00706877 -0.00666022  0.00400841 -0.00237421 -0.00777652
 -0.00028711  0.00674684  0.00307943 -0.00971734]
```

```
'nen ce'
[-0.00426306  0.00648062 -0.00312174 -0.0043469  0.0094298 -0.00811222
 0.006479  0.00139503 -0.0069535  0.00514541  0.00127513 -0.00757256
 0.00899446  0.00955938 -0.00655886  0.00526296 -0.00909529  0.00922226
-0.00864458 -0.00474605  0.00489786  0.00048181  0.00356207  0.00224642
-0.00622795 -0.00848933 -0.00390632 -0.00393118  0.00297766 -0.00182534
-0.00333566 -0.00935472 -0.00855647 -0.00394788 -0.0010447 -0.0037553
 0.00483447 -0.00557921  0.00044492 -0.00242985  0.00668654  0.0051058
-0.00538627 -0.00874314  0.00039148  0.00013857  0.0080862 -0.00234781
 0.00383112  0.0072113  0.00464676 -0.00699011  0.00247818  0.0055567
-0.00471517  0.008778  0.00058283  0.00683827 -0.0032488  0.00722507
 0.00546758 -0.00308026  0.00250784  0.00549752  0.00629684  0.00907396
-0.00641031 -0.00606727  0.00211471  0.0003403  0.00527617 -0.00221127
-0.00409851 -0.00704258 -0.00103335  0.00039326  0.00239346 -0.00070286
-0.00348648 -0.00272986  0.00420162 -0.00186783  0.00422822  0.00854771
-0.00369753 -0.00230391  0.00579277 -0.00564218 -0.00222456  0.00020556
 0.00918382 -0.00642632 -0.00033492 -0.00269044  0.00238411 -0.00545122
-0.00611619  0.00033453  0.00693177 -0.00677497]
```

```
'al di'
[-0.00013627 -0.00103831  0.0049814  0.00223369 -0.00692837  0.00573997
-0.00349263  0.00256323  0.00746696  0.00175508  0.0091698 -0.00804317
 0.00966846  0.00359341  0.00878636  0.00276682 -0.00925543 -0.00155985
 0.00734962  0.00060072  0.0011229  0.00991393  0.00171794  0.002737
-0.00522434 -0.00039378 -0.00393984  0.0044719 -0.00982787  0.00369862
 0.00851404 -0.00283764  0.00322242  0.0095363 -0.00399416  0.00558475
-0.00283721  0.00681554 -0.00331993  0.00437804 -0.00735744  0.00064885
 0.00122797 -0.00923679  0.00608771  0.00733916 -0.00548859  0.00075685
 0.00360962  0.00061106  0.00252087 -0.00184728 -0.0064555 -0.00154123
 0.00237503 -0.00607008  0.00437149 -0.0069836 -0.00091415 -0.00846333
 0.00211604 -0.00434111  0.00529229  0.00462466 -0.00518399  0.00062356
 0.00760913 -0.00676192 -0.00456399 -0.00651479 -0.00328612 -0.00546739
-0.00904213 -0.00857 -0.00399882  0.00665192 -0.00185124  0.00994625
-0.00714183 -0.00969878 -0.00779061 -0.00950482 -0.00971833 -0.00959515
-0.00486043  0.00686048 -0.00948857  0.00185545  0.00566692  0.00409829
 0.00571389 -0.00493123  0.00740435  0.00423611  0.00710547 -0.00161269
-0.00723172  0.00329693  0.00892039 -0.00036149]
```

##### syllable vectors for tri gram#####

```
'çek ti la'
[-0.00638104 -0.00624818 -0.00762885  0.00367238  0.00595551 -0.00526036
-0.0077802  -0.0091204  -0.00086587  0.00204904 -0.00023188 -0.00268111
-0.00186203  0.00101773  0.00380701  0.0047857  -0.00381143 -0.00345892
-0.00244959  0.00311322 -0.00888826 -0.00687722  0.00909482  0.00317877
-0.00341676 -0.00513682  0.00985571  0.00757192 -0.006589  -0.00390394
-0.00023211  0.00195388  0.00309336 -0.00731657 -0.00411387  0.00950448
-0.00704119 -0.0062694  0.00374996  0.00404904  0.00445721  0.00882577
 0.00644936  0.00198033 -0.00865063  0.00847559  0.0098839  -0.00078713
-0.00623531  0.00512145  0.00915438 -0.00017621  0.00380307  0.00458936
-0.00450828  0.004988  -0.00940902  0.00716033  0.00652434  0.00063745
-0.00534689  0.00416827 -0.00681979 -0.00491756  0.00636296 -0.00718545
 0.00012575  0.00357919 -0.00447283 -0.00961754  0.00543799  0.00347632
 0.00100283 -0.00487181  0.00368774  0.00738197  0.00796305  0.00223167
 0.00255186  0.00082874 -0.00690132 -0.00541154  0.0037361  -0.00524238
-0.00994366 -0.00113805  0.00910035 -0.00116617 -0.00904121 -0.00509593
 0.00802371  0.00875085 -0.00207041  0.00885933  0.00879751 -0.00787951
 0.00706  0.00336076  0.00079128 -0.00939002]
```

```
ti la rin'
[ 5.0328062e-03 -6.8789544e-03 -7.9965638e-03 -7.6011969e-03
-6.6722272e-04 -7.9532322e-03 -8.6518526e-03  1.0169065e-03
-1.9648958e-03 -5.0039007e-03 -7.7607357e-03 -3.1291293e-03
-7.9276133e-03  5.4015936e-03  1.1909222e-03 -7.1944143e-03
-7.9997554e-03  3.0487084e-03 -7.6580979e-03 -7.5451494e-04
-9.1529991e-03  9.4989181e-04  6.1780941e-03 -5.2751005e-03
-2.3214221e-03  9.7091366e-03  7.1284999e-03 -6.1897994e-03
-1.5973854e-03 -3.2008325e-03  9.5805824e-03 -6.0325693e-03
-1.7932164e-03  3.2778860e-03 -5.3740456e-03  6.9975699e-03
-1.0598254e-03  1.2622416e-03 -6.3595679e-03 -4.5343409e-03
-6.7283404e-03  7.0191012e-03  9.4041061e-03  9.5847659e-03
-6.5550338e-03  3.7025809e-03  7.1989582e-03 -5.9572961e-03
-7.1779750e-03 -7.9876669e-03  2.8891109e-03 -2.0139623e-03
 2.3955763e-03  7.7700927e-03 -3.5568404e-03  4.6534631e-03
 8.5703209e-03 -3.2499230e-03 -4.3824697e-03  5.8245636e-03
 4.0883422e-04 -5.4694200e-03  1.8123209e-03  9.2665134e-03
 5.5445433e-03 -3.6122478e-03 -1.2186289e-04  7.3664500e-03
 5.6100069e-03 -4.1458057e-03 -6.7828428e-03 -5.0952397e-03
 3.6444603e-03  3.9865579e-03  6.5326705e-03 -7.3878262e-03
 2.6463151e-03 -5.0328290e-03 -6.8920315e-03  6.6269911e-03
 7.4465298e-03  5.8396650e-03  5.4657795e-03 -8.8354871e-03
 4.3583130e-03  7.6084174e-03 -2.7476312e-03 -7.4251126e-03
-8.6618904e-03  6.9283331e-03  6.5065408e-03 -4.4090878e-03
-8.8680210e-03 -9.9298870e-03 -1.1348986e-03 -2.9221727e-03
 3.7717818e-05 -3.8729047e-03 -3.4271812e-03  7.9607256e-03]
```

## Similarities

As can be seen, 'lar' and 'ar' are more similar than the other input, so the similarity is higher.

```
##### similarities for uni-grams #####  
  
'lar', 'ar' 0.592881  
'rak', 'sans' -0.027578644  
  
##### similarities for bi-grams #####  
  
'ge le', 'ge len' 0.20346063  
'ge le', 'nen ce' -0.088584624  
'e lekt', 'ro nik' -0.0051746266  
'al dı', 'zal dı' -0.033639576  
'al dı', 'ro nik' 0.09241031  
  
##### similarities for tri-grams #####  
  
'o la rak', 'la rak bir' 0.022072857  
'o la rak', 'ün ka zan' -0.068939336
```

## Most similarities

1)

```
print(my_model.wv.most_similar(positive=['cen', 'han'], negative=['giz']))  
  
[('ye', 0.7208379507064819),
```

2)

```
print(my_model.wv.most_similar(positive=['top', 'la'], negative=['dı']))
```

```
[('to san', 0.4331499934196472),  
(('la', 0.40073548289434914), (
```

3)

```
print(my_model.wv.most_similar(positive=['mo', 'lis'], negative=['ğö']))
```

```
[('laş', 0.4711923897266388),  
(('la', 0.46642860770225525), (
```

4)

```
print(my_model.wv.most_similar(positive=['hu kuk ve', 've ik ti'], negative=['kuk ve ik']))
```

```
[('ca, za man', 0.46642860770225525),  
(('la', 0.46642860770225525), (
```

5)

```
print(my_model.wv.most_similar(positive=['ya kın', 'liş'], negative=['i']))
```

```
[('ağ zın da', 0.428831547498703),  
(('la', 0.428831547498703), (
```



## All most similar results

```
# most similar
print(my_model.wv.most_similar(positive=['top', 'la'], negative=['d1']))
print("\n")
print(my_model.wv.most_similar(positive=['cen', 'han'], negative=['giz']))
print("\n")
print(my_model.wv.most_similar(positive=['mo', 'lis'], negative=['ğö']))
print("\n")
print(my_model.wv.most_similar(positive=['hu kuk ve', 've ik ti'], negative=['kuk ve ik']))
print("\n")
print(my_model.wv.most_similar(positive=['ya kın', 'liş'], negative=['i']))
```

```
zeroday@zeroday-Lenovo-V330-15IKB:~/Desktop/cse654 nlp/hw2$ python3 hw3.py
[('to san', 0.4331499934196472), ('ol', 0.4177238643169403), ('sil yo da', 0.41598740220069885), ('dan', 0.40726637840270996), ('dan ve ki', 0.40580177307128906), ('do', 0.40072548389434814), ('ri', 0.40061402320861816), ('nim se yen', 0.40002647042274475), ('yon tem de', 0.39891549944877625), ('ner bo ra', 0.3970712125301361)]

[('ye', 0.7208379507064819), ('ne', 0.7192790508270264), ('de', 0.7183389067649841), ('lar', 0.7101946473121643), ('ri', 0.7095198631286621), ('bir', 0.709324002659302), ('nin', 0.708918330535889), ('a', 0.7059782147407532), ('ge', 0.7051292657852173), ('nın', 0.7044931650161743)]

[('laş', 0.4711923897266388), ('bo', 0.45675280690193176), ('tek', 0.42720043659210205), ('ra ma ni', 0.4217362105846405), ('an cak ha', 0.4172513484954834), ('t i', 0.4156123399734497), ('bi', 0.4150746762752533), ('li ka çı', 0.4131702482700348), ('ye', 0.4075290560722351), ('ko', 0.40556588768959045)]

[('ca, za man', 0.46642860770225525), ('an ma kon', 0.4242631793022156), ('ce gö', 0.4197002947330475), ('ke si mill', 0.4150841236114502), ('u, fkf ge', 0.4110562801361084), ('hakk ri', 0.4079899787902832), ('to ya u', 0.40600964426994324), ('ka lı ol', 0.40276122093200684), ('rı ca gı', 0.40227288007736206), ('yaz ma y a', 0.39353084564208984)]

[('ağ zın da', 0.428831547498703), ('le rek yü', 0.3997102677822113), ('tav lık', 0.39695674180984497), ('lizm ne dir', 0.3948558270931244), ('lış ma', 0.39159682393074036), ('bi rer telg', 0.3883526921272278), ('cı ü', 0.3861619234085083), ('ge len ba', 0.3851616084575653), ('rak, ev ri', 0.3838247060775757), ('cak bü', 0.3775055706501007)]
zeroday@zeroday-Lenovo-V330-15IKB:~/Desktop/cse654 nlp/hw2$
```