graph_bipartite

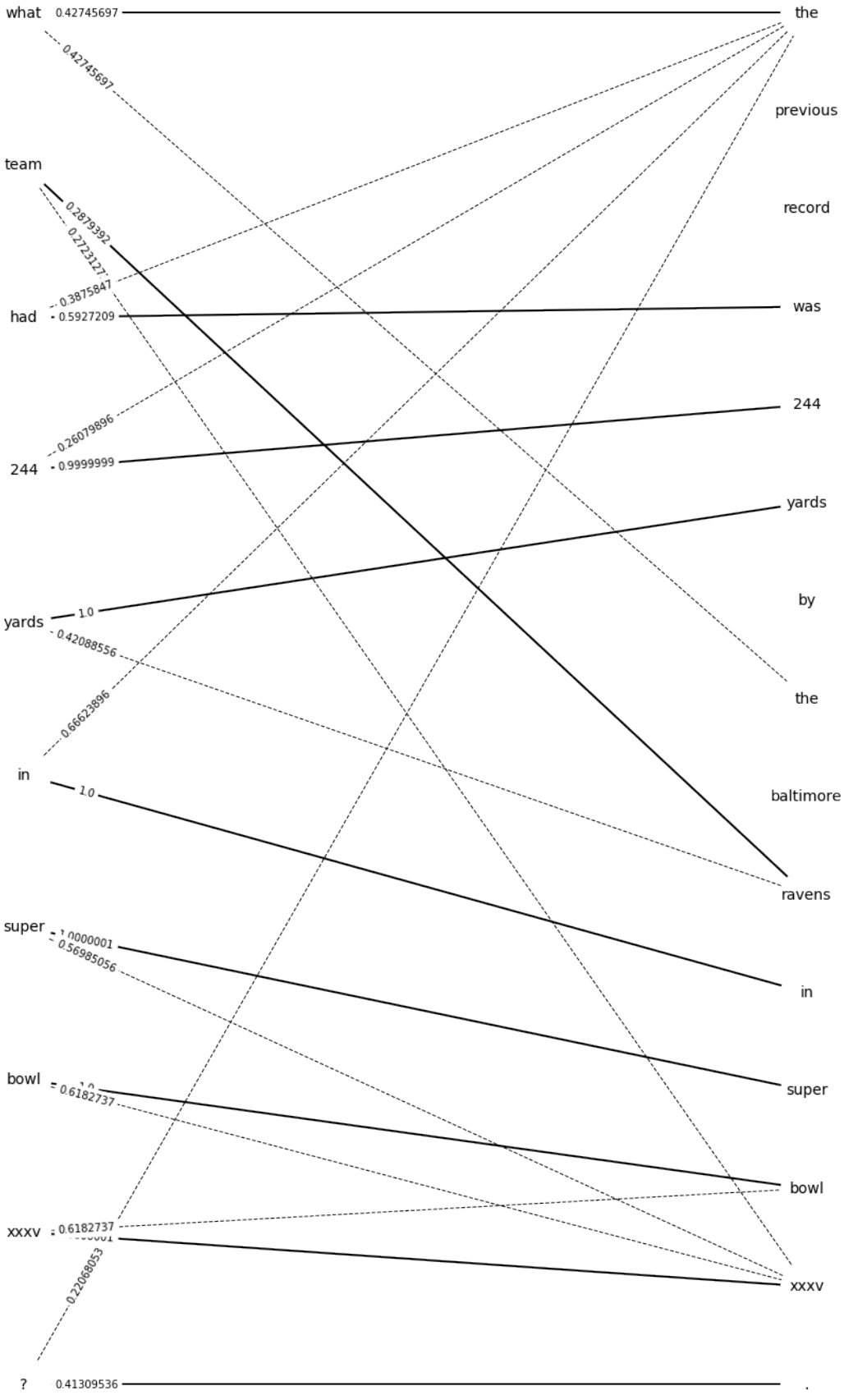
October 10, 2018

In [7]:
```
import networkx as nx
from networkx import *
import matplotlib.pyplot as plt
from networkx.algorithms import bipartite
from nltk import sent_tokenize,word_tokenize
import numpy as np

import sys
sys.path.append('../..')
sys.path.append('../../utils/')
from utils import *
import operator


import fastText
model = fastText.load_model('../../../Divers_Data_Maitrise/wiki.simple/wiki.simple.bin')

#ignorer toutes les depreciations de fonctiona
import warnings
warnings.simplefilter('ignore')
```

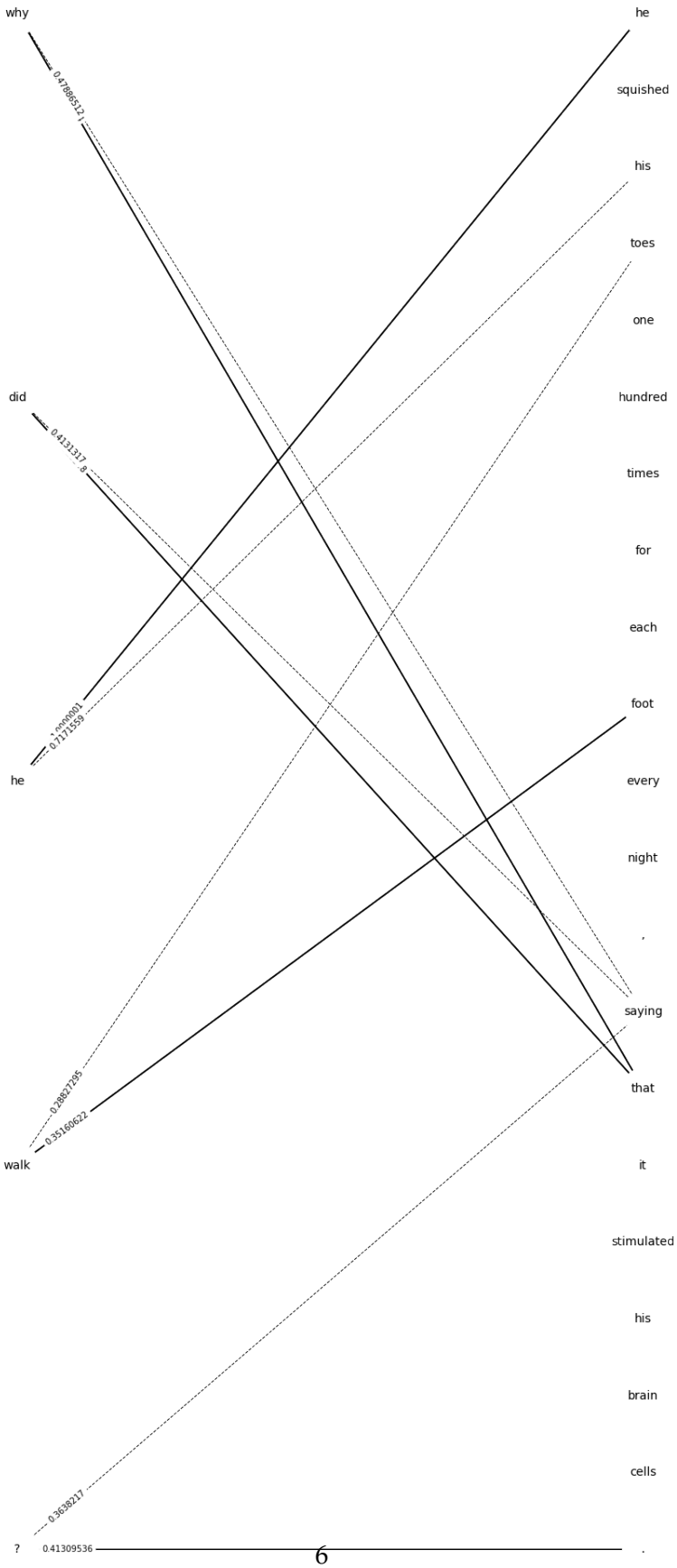
In [17]:
```
def Alignement_graph_bipartite(question,sequence, nb_alignement_a_afficher = 1, lower_c
    '''
    Fonction qui crée un graphe bipartit (question, sentence) et qui affecte la similar
    cosine comme poids de chaque arete. L'arete en trait plein represente la plus forte
    les autres similarite seront affichées avec des traits discontinus.
    :param nb_alignement_a_afficher: nombre de similaritées a afficher
    :param lower_case_bool: True pour mettre tout en minuscule, False sinon
    '''

    nb_alignement_a_afficher = max(1,nb_alignement_a_afficher)
    G=nx.Graph()

    list_words_question = word_tokenize(question.lower() if lower_case_bool else questi
    list_words_sequence = word_tokenize(sequence.lower() if lower_case_bool else sequen
```

```
nb_words_question = len(list_words_question)
nb_words_sequence = len(list_words_sequence)
height_colum = max(nb_words_question,nb_words_sequence)
y_linspace_question = np.linspace(0,height_colum,nb_words_question)
y_linspace_sentence = np.linspace(0,height_colum,nb_words_sequence)

fixed_positions={}
labels={}


list_major_edges = []
list_minor_edges = []
edge_labels = {}
for i in range(0,nb_words_sequence): # mots de la phrase
    num_node = nb_words_question + i
    G.add_node(num_node)
    labels[num_node]=list_words_sequence[i]

    fixed_positions[num_node] = (2,y_linspace_sentence[nb_words_sequence-i-1])

for i in range(0,nb_words_question): # mots de la question
    G.add_node(i)
    labels[i]=list_words_question[i]
    fixed_positions[i] = (0,y_linspace_question[nb_words_question-i-1])

    max_sim = 0.0
    major_edge = None
    edges_list = []
    edges_sim_dict ={}
    vect_word_question = model.get_word_vector(list_words_question[i])
    for j in range(0,nb_words_sequence):
        num_node = nb_words_question + j
        vect_word_sentence = model.get_word_vector(list_words_sequence[j])
        sim=cosine_similarity(vect_word_sentence, vect_word_question)

        edges_sim_dict[len(edges_list)]=sim
        edges_list.append((i,num_node, sim))
        G.add_edge(i,num_node,sim=sim)

    edges_sorted = sorted(edges_sim_dict.items(), key=operator.itemgetter(1), rever
    list_major_edges.append(edges_list[edges_sorted[0][0]])
    edge_labels[(edges_list[edges_sorted[0][0]][1],edges_list[edges_sorted[0][0]][0
    for k in range(1,len(edges_sorted)):
        list_minor_edges.append(edges_list[edges_sorted[k][0]])
        edge_labels[(edges_list[edges_sorted[k][0]][1],edges_list[edges_sorted[k][0

pos = nx.spring_layout(G,pos=fixed_positions, fixed=fixed_positions.keys())
plt.figure(3,figsize=(15,height_colum*1.75))
```

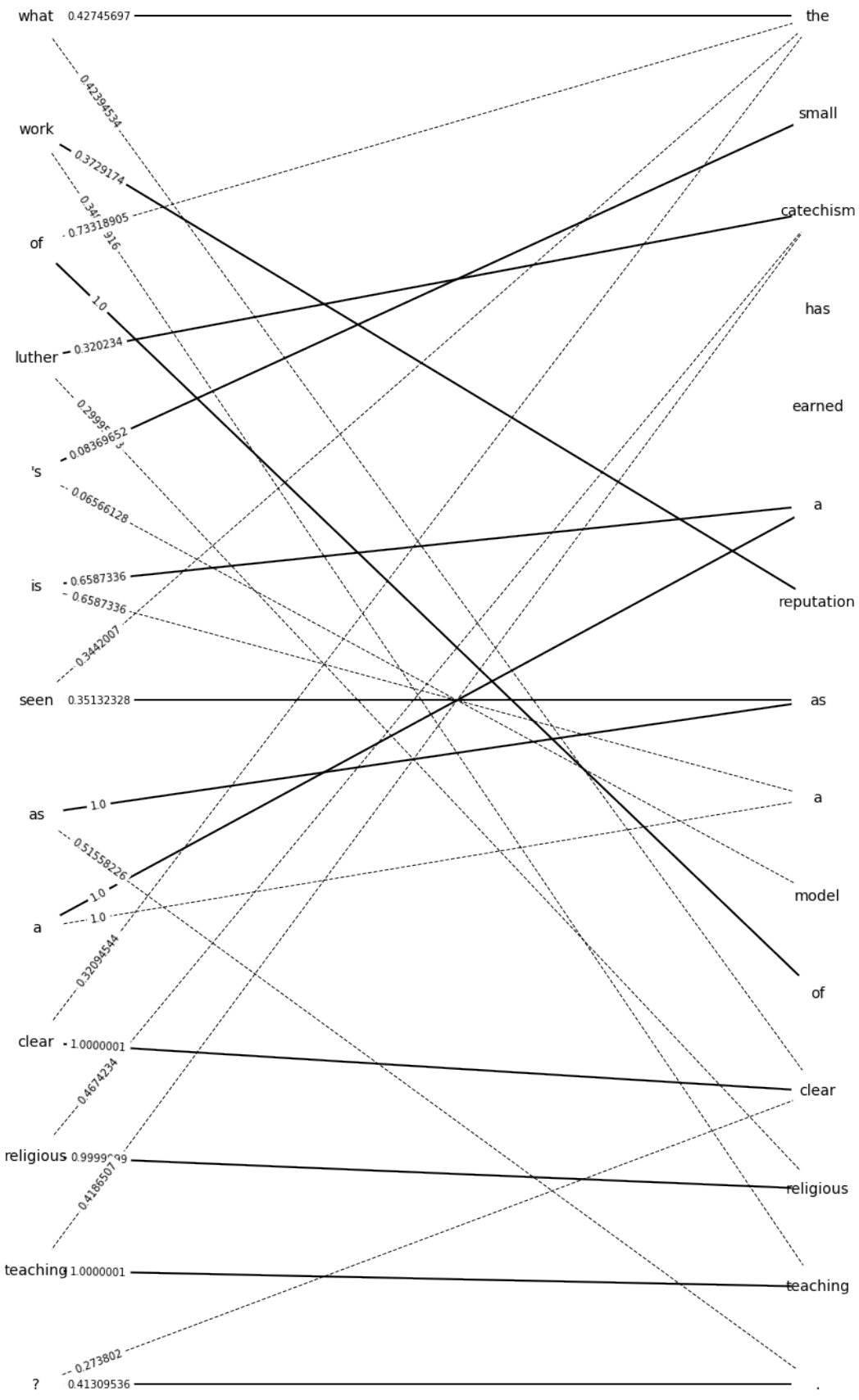
```
        nx.draw_networkx_nodes(G,pos,node_color='w',node_size=2500)
        nx.draw_networkx_labels(G,pos,labels,font_size=14)
        nx.draw_networkx_edge_labels(G,pos,node_color='b',edge_labels=edge_labels, label_po
        nx.draw_networkx_edges(G,pos,node_color='b', edgelist =list_major_edges,  width = 2
        nx.draw_networkx_edges(G,pos,node_color='b', edgelist =list_minor_edges,style = 'da

        plt.axis('off')
    #     plt.savefig("Graph.png", format="PNG") # pour enregistrer l'image dans un fichier
        plt.show()

In [15]: Alignement_graph_bipartite(" What team had 244 yards in Super Bowl XXXV? "," The previo
```


what
team
had
244
yards
in
super
bowl
xxxv
?
the
previous
record
was
244
yards
by
the
baltimore
ravens
in
super
bowl
xxxv
.
0.42745697
0.42745697
0.2879392
0.27231127
0.3875847
0.5927209
0.26079896
0.9999999
1.0
0.42088556
0.66623896
1.0
0.56985056
0.6182737
0.6182737
0.22068853
0.41309536

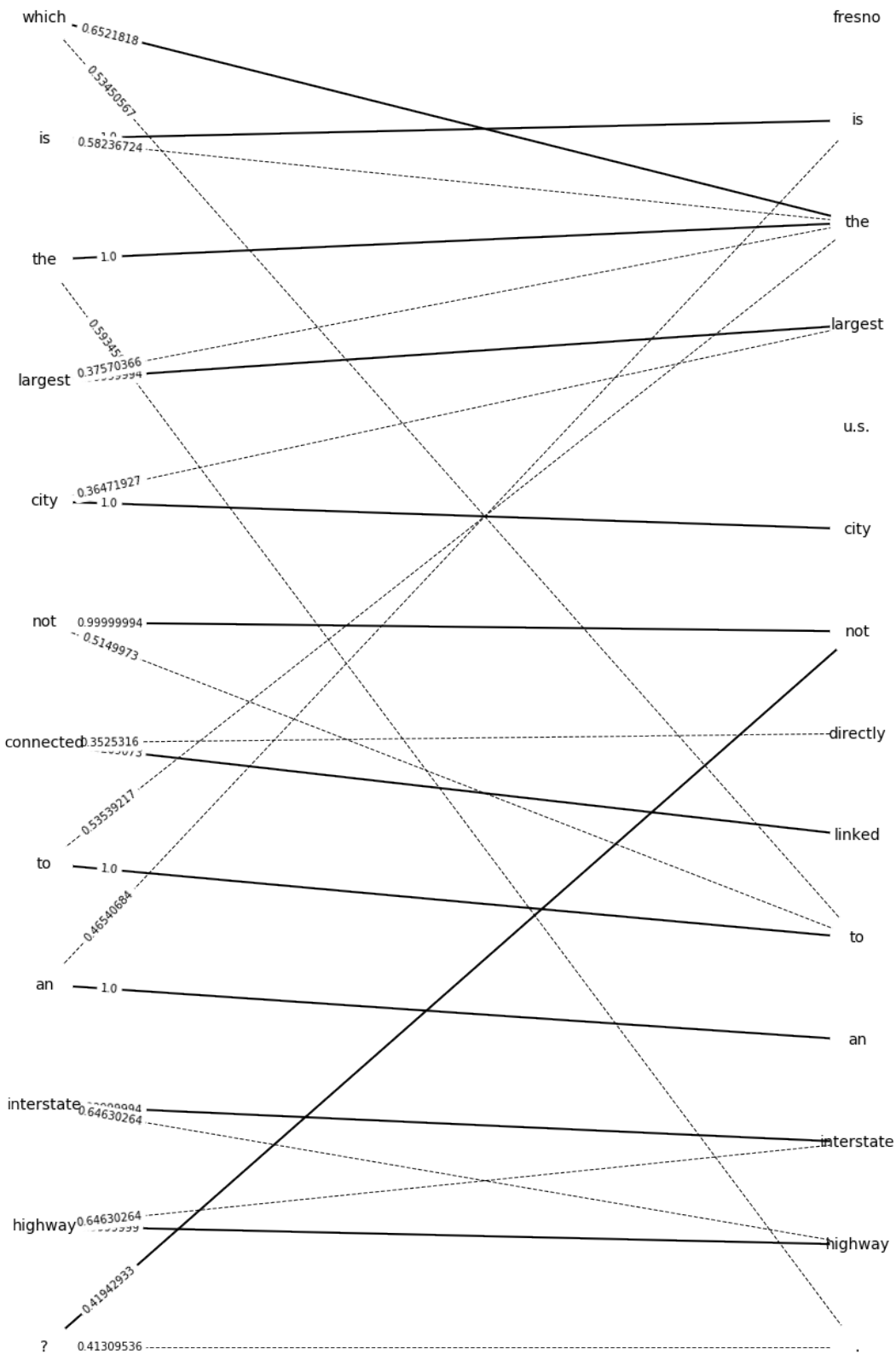
```
In [18]: Alignement_graph_bipartite(" Why did he walk? "," He squished his toes one hundred time
```


why
did
he
walk
?
he
squished
his
toes
one
hundred
times
for
each
foot
every
night
,
saying
that
it
stimulated
his
brain
cells
.
0.47886512
0.4131317
0.71171559
0.28827295
0.35160622
0.3638217
0.41309536

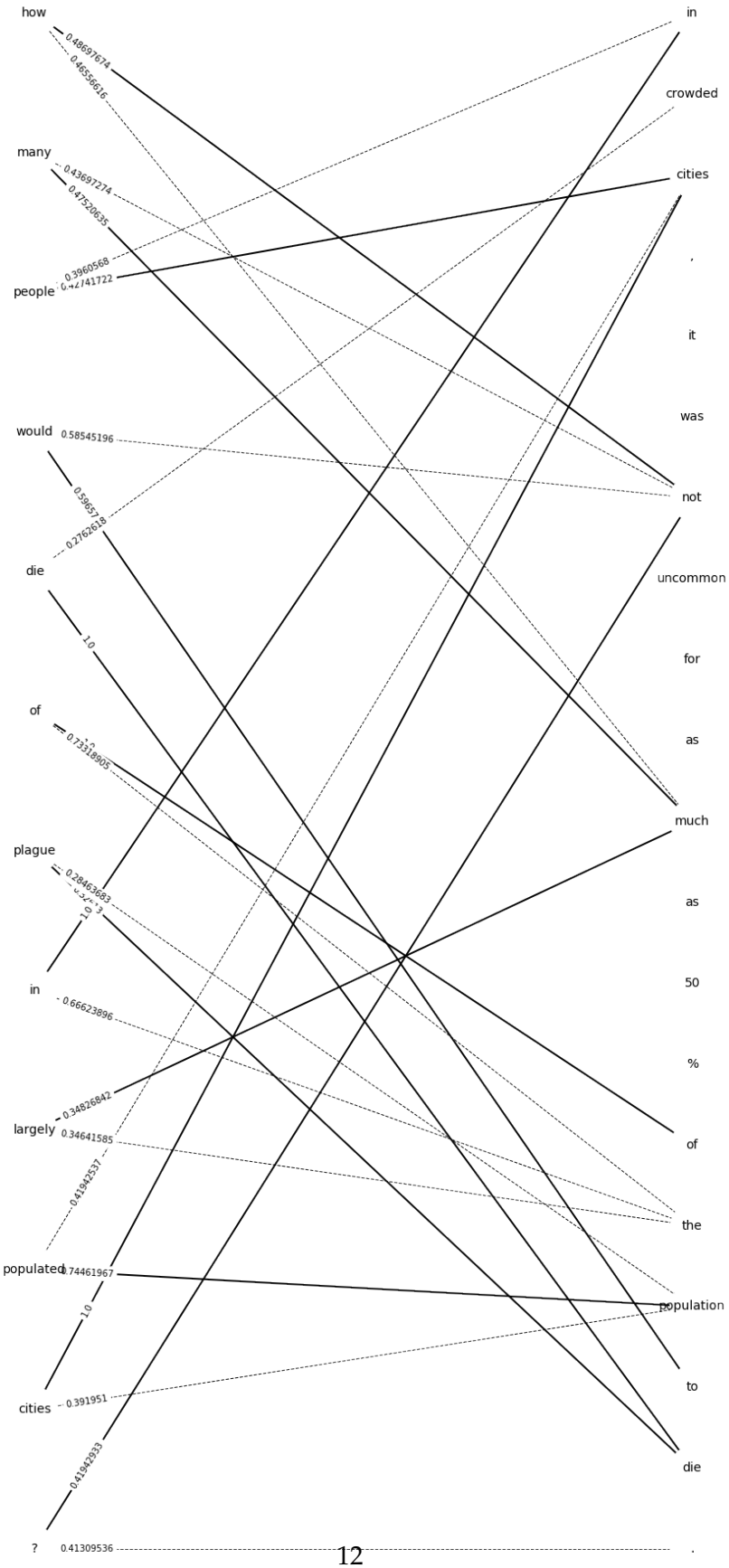
```
In [19]: Alignement_graph_bipartite(" What work of Luther's is seen as a clear religious teachin
```


what
work
of
luther
's
is
seen
as
a
clear
religious
teaching
?
the
small
catechism
has
earned
a
reputation
as
a
model
of
clear
religious
teaching
.

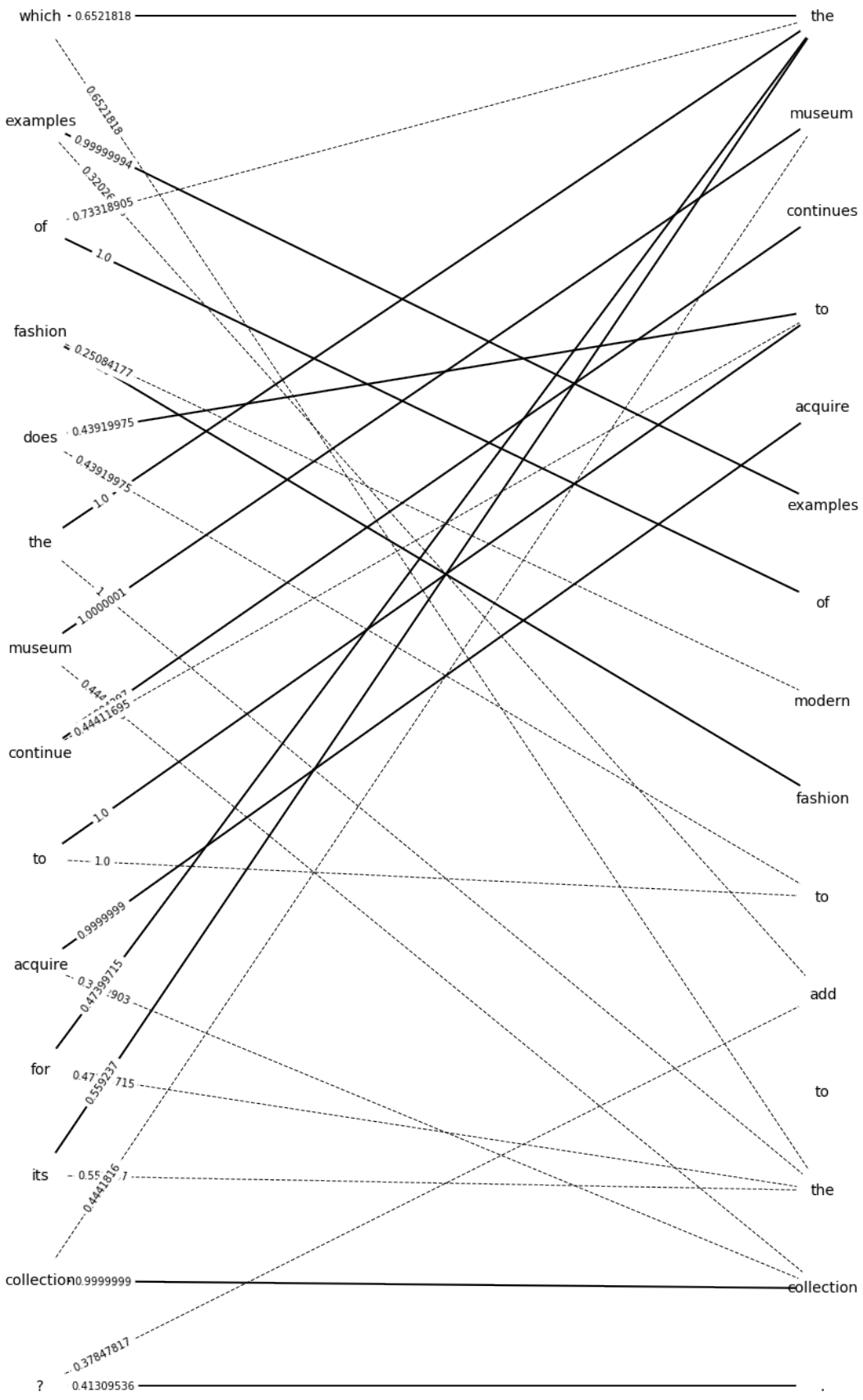
```
In [20]: Alignement_graph_bipartite(" Which is the largest city not connected to an interstate h
```


which
is
the
largest
city
not
connected
to
an
interstate
highway
?
fresno
is
the
largest
u.s.
city
not
directly
linked
to
an
interstate
highway
.
0.6521818
0.53450567
0.58236724
1.0
0.37570366
0.36471927
1.0
0.99999994
0.5149973
0.3525316
0.53539217
1.0
0.46540684
1.0
0.64630264
0.64630264
0.41942933
0.41309536

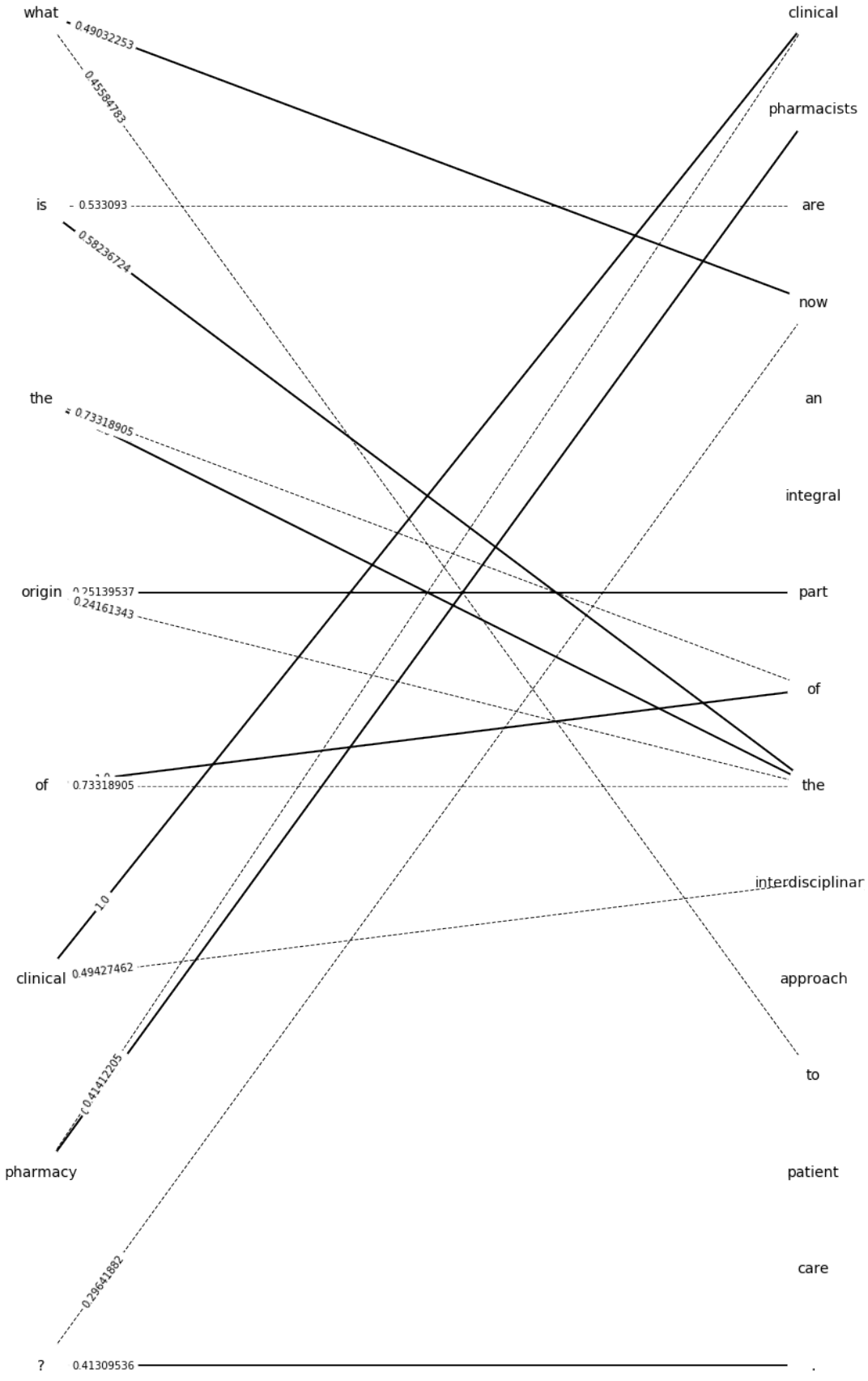
```
In [21]: Alignement_graph_bipartite(" How many people would die of plague in largely populated c
```


how
many
people
would
die
of
plague
in
largely
populated
cities
?
in
crowded
cities
,
it
was
not
uncommon
for
as
much
as
50
%
of
the
population
to
die
.
0.48697674
0.46556616
0.43697274
0.47520635
0.39605568
0.42741722
0.58545196
0.59657
0.27626218
1.0
0.73318905
0.28463683
1.0
0.66623896
0.34826842
0.34641585
0.41942537
0.74461967
1.0
0.391951
0.41942933
0.41309536

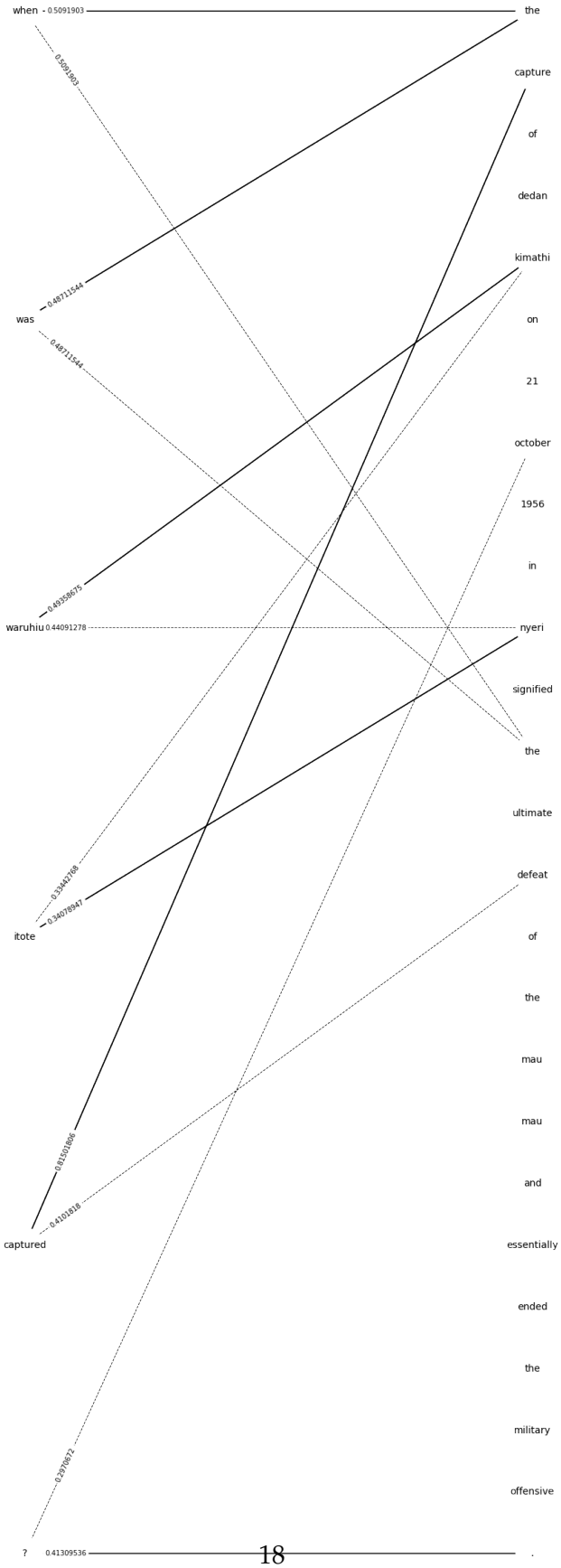
```
In [22]: Alignement_graph_bipartite(" Which examples of fashion does the museum continue to acqu
```


which
examples
of
fashion
does
the
museum
continue
to
acquire
for
its
collection
?
the
museum
continues
to
acquire
examples
of
modern
fashion
to
add
to
the
collection
.
0.6521818
0.99999994
0.73318905
1.0
0.25084177
0.43919975
1.0
1.0000001
1.0
1.0
0.9999999
0.47399715
0.55923706
0.9999999
0.37847817
0.41309536

```
In [23]: Alignement_graph_bipartite(" What is the origin of clinical pharmacy? "," Clinical phar
```


what
is
the
origin
of
clinical
pharmacy
?
clinical
pharmacists
are
now
an
integral
part
of
the
interdisciplinar
approach
to
patient
care
.
0.49032253
0.45584783
0.533093
0.58236724
0.73318905
0.25139537
0.24161343
0.73318905
1.0
0.49427462
0.41412205
0.29641882
0.41309536

```
In [24]: Alignement_graph_bipartite(" When was Waruhiu Itote captured? "," The capture of Dedan
```


when
was
waruhiu
itote
captured
?
the
capture
of
dedan
kimathi
on
21
october
1956
in
nyeri
signified
the
ultimate
defeat
of
the
mau
mau
and
essentially
ended
the
military
offensive
.
0.5091903
0.5091903
0.48711544
0.48711544
0.49358675
0.44091278
0.33441768
0.34078947
0.81501806
0.4101818
0.2970872
0.41309536