

Exercise 6 - Vector space model

First name: Brian

Last name: Schweigler

Matriculation number: 16-102-071

Q1: Represent each play by a vector with only the tf component. You can apply some preprocessing before generating this vector representation.

General imports and setup:

In [37]:

```
%load_ext autoreload
%autoreload 2
%matplotlib inline

import matplotlib.pyplot as plt
import pandas as pd
import re
import numpy as np
import lxml.etree
import os
from scipy import stats
import nltk
import nltk.tokenize
import collections

np.random.seed(6) # for reproducibility

PUNCT_RE = re.compile(r'^\w\s]+$')

def is_punct(string):
    """Check if STRING is a punctuation marker or a sequence of
    punctuation markers.
    """
    return PUNCT_RE.match(string) is not None

def preprocess_text(text, language='French', lowercase=True):
    if lowercase:
        text = text.lower()
    if language == 'French':
        text = re.sub("-", " ", text)
        text = re.sub("l'", "le ", text)
        text = re.sub("d'", "de ", text)
        text = re.sub("c'", "ce ", text)
        text = re.sub("j'", "je ", text)
        text = re.sub("m'", "me ", text)
        text = re.sub("qu'", "que ", text)
        text = re.sub("'", " ' ", text)
        text = re.sub("quelqu'", "quelque ", text)
        text = re.sub("aujourd'hui", "aujourd'hui", text)
        text = re.sub(r'\d+', '', text)
        text = re.sub(r'^\w\s]', '', text)
    tokens = nltk.tokenize.word_tokenize(text, language=language)
    tokens = [token for token in tokens if not is_punct(token)]
    return tokens

subgenres = ('Comédie', 'Tragédie', 'Tragi-comédie')
plays, titles, genres = [], [], []
authors, years = [], []

for fn in os.listdir('theatre-classique'):
    # Only include XML files
```

```

if not fn.name.endswith('.xml'):
    continue
tree = lxml.etree.parse(fn.path)
genre = tree.find('//genre')
title = tree.find('//title')
author = tree.find('//author')
year = tree.find('//date')
if genre is not None and genre.text in subgenres:
    lines = []
    for line in tree.xpath('//l//p'):
        lines.append(' '.join(line.itertext()))
    text = '\n'.join(lines)
    plays.append(text)
    genres.append(genre.text)
    titles.append(title.text)
    authors.append(author.text)
    if year is not None:
        years.append(year.text)

print(len(plays), len(genres), len(titles), len(authors), len(years))

```

The autoreload extension is already loaded. To reload it, use:
 %reload_ext autoreload
 498 498 498 498 208

In [38]:

```

#nltk.download('punkt')
plays_tok = [preprocess_text(play, 'French') for play in plays]

def extract_vocabulary(tokenized_corpus, min_count=1, max_count=float('inf')):
    vocabulary = collections.Counter()
    for document in tokenized_corpus:
        vocabulary.update(document)
    vocabulary = {word for word, count in vocabulary.items()
                  if count >= min_count and count <= max_count}
    return sorted(vocabulary)

vocabulary = extract_vocabulary(plays_tok, min_count=2)
len(vocabulary)

```

Out[38]: 37824

In [39]:

```

# Representation in a doc x term matrix for the French plays
def corpus2dtm(tokenized_corpus, vocabulary):
    "Transform a tokenized corpus into a document-term matrix"
    document_term_matrix = []
    for document in tokenized_corpus:
        document_counts = collections.Counter(document)
        row = [document_counts[word] for word in vocabulary]
        document_term_matrix.append(row)
    return document_term_matrix

# building the doc/term matrix in a few seconds for the French example
document_term_matrix = np.array(corpus2dtm(plays_tok, vocabulary))
print(f"document-term matrix with "
      f"|D| = {document_term_matrix.shape[0]} documents and "
      f"|V| = {document_term_matrix.shape[1]} words.")

```

document-term matrix with |D| = 498 documents and |V| = 37824 words.

Now we can calculate the tf component, which is the occurrence frequency of a term in the whole corpus. For this, we stem the words.

In [40]:

```

import operator
from nltk import word_tokenize
from nltk.stem import SnowballStemmer
fr = SnowballStemmer('french')

```

```

stemmed_vocabulary = [fr.stem(word) for word in vocabulary]

counts = dict()

for word in stemmed_vocabulary:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict = counts
for key, value in counts.items():
    tf_dict[key] = value/len(stemmed_vocabulary)

tf_dict

```

```

Out[40]: {'a': 2.643824027072758e-05,
'aa': 2.643824027072758e-05,
'aah': 2.643824027072758e-05,
'aaron': 2.643824027072758e-05,
'ab': 2.643824027072758e-05,
'abaiss': 0.00042301184433164127,
'abaissement': 5.287648054145516e-05,
'abaissent': 2.643824027072758e-05,
'abandon': 0.0006609560067681895,
'abandonnent': 2.643824027072758e-05,
'abandonnon': 2.643824027072758e-05,
'abat': 5.287648054145516e-05,
'abatt': 0.0001321912013536379,
'abattent': 2.643824027072758e-05,
'abattr': 2.643824027072758e-05,
'abattu': 7.931472081218273e-05,
'abattus': 2.643824027072758e-05,
'abbess': 2.643824027072758e-05,
'abbé': 5.287648054145516e-05,
'abces': 2.643824027072758e-05,
'abdened': 2.643824027072758e-05,
'abder': 2.643824027072758e-05,
'abdiqu': 7.931472081218273e-05,
'abdit': 2.643824027072758e-05,
'abdérit': 2.643824027072758e-05,
'abec': 2.643824027072758e-05,
'abeil': 5.287648054145516e-05,
'abencerrag': 2.643824027072758e-05,
'abendax': 2.643824027072758e-05,
'abez': 2.643824027072758e-05,
'abhorrr': 0.00018506768189509307,
'abhorrent': 2.643824027072758e-05,
'abincerrag': 2.643824027072758e-05,
'abject': 7.931472081218273e-05,
'abjur': 0.0001321912013536379,
'ablativo': 2.643824027072758e-05,
'abner': 2.643824027072758e-05,
'abois': 2.643824027072758e-05,
'abol': 0.00015862944162436547,
'abomin': 7.931472081218273e-05,
'abond': 0.00018506768189509307,
'abon': 2.643824027072758e-05,
'abord': 0.00034369712351945856,
'abordag': 2.643824027072758e-05,
'abordent': 2.643824027072758e-05,
'abordon': 2.643824027072758e-05,
'abouch': 5.287648054145516e-05,
'aboumékares': 2.643824027072758e-05,
'about': 0.00018506768189509307,
'aboi': 0.00010575296108291032,
'abra': 2.643824027072758e-05,

```

'abraham': 2.643824027072758e-05,
'abre': 2.643824027072758e-05,
'abreuv': 0.0001321912013536379,
'abreuvoir': 2.643824027072758e-05,
'abri': 2.643824027072758e-05,
'abreg': 0.00023794416243654823,
'abrègent': 2.643824027072758e-05,
'abrégeon': 2.643824027072758e-05,
'absalon': 2.643824027072758e-05,
'absenc': 5.287648054145516e-05,
'absent': 0.00010575296108291032,
'absit': 2.643824027072758e-05,
'absolu': 0.0001321912013536379,
'absolus': 2.643824027072758e-05,
'absorb': 2.643824027072758e-05,
'absoudr': 5.287648054145516e-05,
'absous': 2.643824027072758e-05,
'absout': 2.643824027072758e-05,
'absten': 5.287648054145516e-05,
'abstenu': 2.643824027072758e-05,
'abstien': 2.643824027072758e-05,
'abstinent': 2.643824027072758e-05,
'abstract': 2.643824027072758e-05,
'abstrait': 2.643824027072758e-05,
'absurd': 0.00010575296108291032,
'abus': 0.0006345177664974619,
'abusent': 2.643824027072758e-05,
'abuson': 2.643824027072758e-05,
'abym': 2.643824027072758e-05,
'abîm': 0.00018506768189509307,
'ac': 2.643824027072758e-05,
'acab': 2.643824027072758e-05,
'académicien': 2.643824027072758e-05,
'academ': 0.0001321912013536379,
'acant': 2.643824027072758e-05,
'acanth': 2.643824027072758e-05,
'acariâtr': 2.643824027072758e-05,
'acast': 2.643824027072758e-05,
'acbitophel': 2.643824027072758e-05,
'accabl': 0.00044945008460236884,
'accablent': 2.643824027072758e-05,
'accablon': 2.643824027072758e-05,
'accapareur': 2.643824027072758e-05,
'accent': 5.287648054145516e-05,
'accept': 0.00047588832487309646,
'accepton': 2.643824027072758e-05,
'accessibl': 5.287648054145516e-05,
'accessoir': 5.287648054145516e-05,
'accident': 5.287648054145516e-05,
'accin': 2.643824027072758e-05,
'acclam': 2.643824027072758e-05,
'accoint': 2.643824027072758e-05,
'accolad': 2.643824027072758e-05,
'accommod': 0.000502326565143824,
'accommodent': 2.643824027072758e-05,
'accommodon': 2.643824027072758e-05,
'accomod': 2.643824027072758e-05,
'accompagn': 0.000502326565143824,
'accompagnent': 2.643824027072758e-05,
'accompl': 0.00044945008460236884,
'accord': 0.0006609560067681895,
'accordent': 2.643824027072758e-05,
'accordon': 2.643824027072758e-05,
'accort': 5.287648054145516e-05,
'accortis': 2.643824027072758e-05,
'accost': 7.931472081218273e-05,
'accouch': 0.00018506768189509307,
'accour': 0.0001321912013536379,
'accourc': 7.931472081218273e-05,

'accourent': 2.643824027072758e-05,
'accourt': 2.643824027072758e-05,
'accouru': 5.287648054145516e-05,
'accourus': 2.643824027072758e-05,
'accourut': 2.643824027072758e-05,
'accout': 5.287648054145516e-05,
'accoutr': 2.643824027072758e-05,
'accoutum': 0.00029082064297800337,
'accoutu': 2.643824027072758e-05,
'accoutumon': 2.643824027072758e-05,
'accroch': 0.00010575296108291032,
'accrochent': 2.643824027072758e-05,
'accroir': 2.643824027072758e-05,
'accrois': 2.643824027072758e-05,
'accroiss': 5.287648054145516e-05,
'accroissement': 2.643824027072758e-05,
'accroissent': 2.643824027072758e-05,
'accroît': 2.643824027072758e-05,
'accroît': 2.643824027072758e-05,
'accroître': 7.931472081218273e-05,
'accroître': 2.643824027072758e-05,
'accru': 5.287648054145516e-05,
'accrurent': 2.643824027072758e-05,
'accrus': 2.643824027072758e-05,
'accrut': 2.643824027072758e-05,
'accrédit': 2.643824027072758e-05,
'accred': 2.643824027072758e-05,
'accrût': 2.643824027072758e-05,
'accueil': 0.00021150592216582064,
'accueillent': 2.643824027072758e-05,
'accumul': 7.931472081218273e-05,
'accus': 0.0007667089678510998,
'accusent': 2.643824027072758e-05,
'accuson': 2.643824027072758e-05,
'acced': 2.643824027072758e-05,
'acces': 2.643824027072758e-05,
'achab': 2.643824027072758e-05,
'acharn': 0.00015862944162436547,
'achas': 2.643824027072758e-05,
'achat': 7.931472081218273e-05,
'achemin': 5.287648054145516e-05,
'acheminon': 2.643824027072758e-05,
'achet': 0.00047588832487309646,
'acheteur': 5.287648054145516e-05,
'acheton': 2.643824027072758e-05,
'achev': 0.0006345177664974619,
'achevon': 2.643824027072758e-05,
'achil': 2.643824027072758e-05,
'achill': 5.287648054145516e-05,
'achillea': 2.643824027072758e-05,
'achior': 2.643824027072758e-05,
'achis': 2.643824027072758e-05,
'achitophel': 2.643824027072758e-05,
'achmat': 2.643824027072758e-05,
'achmet': 2.643824027072758e-05,
'achomat': 2.643824027072758e-05,
'achopp': 2.643824027072758e-05,
'achètent': 2.643824027072758e-05,
'achèvent': 2.643824027072758e-05,
'achéron': 2.643824027072758e-05,
'acid': 2.643824027072758e-05,
'aci': 2.643824027072758e-05,
'acmat': 2.643824027072758e-05,
'acomat': 2.643824027072758e-05,
'acon': 2.643824027072758e-05,
'acoquin': 2.643824027072758e-05,
'acord': 2.643824027072758e-05,
'acout': 7.931472081218273e-05,
'acquerr': 5.287648054145516e-05,

'acquier': 5.287648054145516e-05,
'acquiert': 2.643824027072758e-05,
'acquiesc': 2.643824027072758e-05,
'acquirent': 2.643824027072758e-05,
'acquis': 7.931472081218273e-05,
'acquisit': 5.287648054145516e-05,
'acquit': 5.287648054145516e-05,
'acquitt': 0.00037013536379018613,
'acquittent': 2.643824027072758e-05,
'acquitton': 2.643824027072758e-05,
'acqui': 2.643824027072758e-05,
'acquer': 0.00010575296108291032,
'acquéreur': 2.643824027072758e-05,
'acquéron': 2.643824027072758e-05,
'acquêt': 5.287648054145516e-05,
'acris': 2.643824027072758e-05,
'acte': 5.287648054145516e-05,
'acteon': 2.643824027072758e-05,
'acteur': 5.287648054145516e-05,
'actif': 5.287648054145516e-05,
'action': 5.287648054145516e-05,
'actium': 2.643824027072758e-05,
'activ': 5.287648054145516e-05,
'actor': 2.643824027072758e-05,
'actric': 5.287648054145516e-05,
'actuel': 7.931472081218273e-05,
'actéon': 2.643824027072758e-05,
'acer': 2.643824027072758e-05,
'ad': 2.643824027072758e-05,
'adag': 2.643824027072758e-05,
'adam': 5.287648054145516e-05,
'adamir': 2.643824027072758e-05,
'adapt': 2.643824027072758e-05,
'addison': 2.643824027072758e-05,
'addit': 5.287648054145516e-05,
'adest': 2.643824027072758e-05,
'adhuc': 2.643824027072758e-05,
'adhérent': 2.643824027072758e-05,
'adieu': 2.643824027072758e-05,
'adieux': 2.643824027072758e-05,
'adin': 2.643824027072758e-05,
'adio': 2.643824027072758e-05,
'adject': 2.643824027072758e-05,
'adjoint': 2.643824027072758e-05,
'adjug': 7.931472081218273e-05,
'admet': 0.00010575296108291032,
'admettent': 2.643824027072758e-05,
'admettr': 5.287648054145516e-05,
'administr': 2.643824027072758e-05,
'admir': 0.0006609560067681895,
'admirent': 2.643824027072758e-05,
'admiron': 2.643824027072758e-05,
'admis': 5.287648054145516e-05,
'adolescent': 5.287648054145516e-05,
'adonc': 2.643824027072758e-05,
'adoni': 2.643824027072758e-05,
'adon': 0.00010575296108291032,
'adopt': 0.00031725888324873094,
'ador': 0.0007138324873096446,
'adorec': 2.643824027072758e-05,
'adorent': 2.643824027072758e-05,
'adoro': 2.643824027072758e-05,
'adoron': 2.643824027072758e-05,
'adouc': 0.00044945008460236884,
'adrast': 2.643824027072758e-05,
'adress': 0.00037013536379018613,
'adressent': 2.643824027072758e-05,
'adresson': 2.643824027072758e-05,
'adrian': 2.643824027072758e-05,

'adrien': 5.287648054145516e-05,
'adroit': 0.0001321912013536379,
'adul': 5.287648054145516e-05,
'adult': 2.643824027072758e-05,
'adulter': 5.287648054145516e-05,
'adven': 2.643824027072758e-05,
'advenu': 2.643824027072758e-05,
'advenus': 2.643824027072758e-05,
'adversair': 5.287648054145516e-05,
'advers': 7.931472081218273e-05,
'advien': 2.643824027072758e-05,
'advient': 2.643824027072758e-05,
'advocatus': 2.643824027072758e-05,
'adélaïd': 2.643824027072758e-05,
'aeacid': 2.643824027072758e-05,
'aegid': 2.643824027072758e-05,
'aegin': 2.643824027072758e-05,
'aegl': 2.643824027072758e-05,
'aenarus': 2.643824027072758e-05,
'aequor': 2.643824027072758e-05,
'aerop': 2.643824027072758e-05,
'afec': 2.643824027072758e-05,
'affabl': 5.287648054145516e-05,
'affad': 2.643824027072758e-05,
'affaibl': 0.00034369712351945856,
'affair': 7.931472081218273e-05,
'affaiss': 2.643824027072758e-05,
'affam': 0.0001321912013536379,
'affect': 0.00042301184433164127,
'affectent': 2.643824027072758e-05,
'affection': 5.287648054145516e-05,
'affecton': 2.643824027072758e-05,
'afferme': 0.00031725888324873094,
'affich': 0.00018506768189509307,
'affichent': 2.643824027072758e-05,
'affid': 2.643824027072758e-05,
'affi': 2.643824027072758e-05,
'affil': 2.643824027072758e-05,
'affin': 7.931472081218273e-05,
'affiquet': 2.643824027072758e-05,
'affirm': 0.00010575296108291032,
'afflict': 5.287648054145516e-05,
'afflig': 0.00044945008460236884,
'affligent': 2.643824027072758e-05,
'affligeon': 2.643824027072758e-05,
'affluenc': 2.643824027072758e-05,
'affol': 7.931472081218273e-05,
'affranch': 0.00031725888324873094,
'affreux': 7.931472081218273e-05,
'affront': 0.0002643824027072758,
'affronteur': 2.643824027072758e-05,
'affubl': 2.643824027072758e-05,
'affût': 2.643824027072758e-05,
'afin': 2.643824027072758e-05,
'afoir': 2.643824027072758e-05,
'afre': 2.643824027072758e-05,
'africain': 0.00010575296108291032,
'afriqu': 2.643824027072758e-05,
'aga': 2.643824027072758e-05,
'agac': 7.931472081218273e-05,
'agag': 2.643824027072758e-05,
'agamemnon': 2.643824027072758e-05,
'agant': 2.643824027072758e-05,
'agat': 2.643824027072758e-05,
'agath': 2.643824027072758e-05,
'agathocl': 2.643824027072758e-05,
'agathon': 2.643824027072758e-05,
'age': 2.643824027072758e-05,
'agenc': 2.643824027072758e-05,

'agend': 2.643824027072758e-05,
'agent': 7.931472081218273e-05,
'ager': 2.643824027072758e-05,
'agesilan': 2.643824027072758e-05,
'aggrav': 5.287648054145516e-05,
'agi': 2.643824027072758e-05,
'agil': 5.287648054145516e-05,
'agioteur': 2.643824027072758e-05,
'agir': 0.0001321912013536379,
'agiron': 2.643824027072758e-05,
'agiront': 2.643824027072758e-05,
'agis': 2.643824027072758e-05,
'agiss': 0.00015862944162436547,
'agissent': 2.643824027072758e-05,
'agisson': 2.643824027072758e-05,
'agit': 0.0003965736040609137,
'agitatus': 2.643824027072758e-05,
'agitent': 2.643824027072758e-05,
'agla': 2.643824027072758e-05,
'agnant': 2.643824027072758e-05,
'agneau': 5.287648054145516e-05,
'agnelet': 2.643824027072758e-05,
'agnonid': 2.643824027072758e-05,
'agnes': 2.643824027072758e-05,
'agné': 2.643824027072758e-05,
'agon': 2.643824027072758e-05,
'agonis': 7.931472081218273e-05,
'agraf': 2.643824027072758e-05,
'agrand': 0.0001321912013536379,
'agresseur': 2.643824027072758e-05,
'agriabl': 2.643824027072758e-05,
'agricultur': 2.643824027072758e-05,
'agri': 2.643824027072758e-05,
'agrip': 2.643824027072758e-05,
'agripin': 2.643824027072758e-05,
'agripp': 5.287648054145516e-05,
'agrippin': 2.643824027072758e-05,
'agré': 0.00021150592216582064,
'agréabl': 7.931472081218273e-05,
'agrément': 5.287648054145516e-05,
'aguerr': 0.00010575296108291032,
'aguet': 2.643824027072758e-05,
'agel': 2.643824027072758e-05,
'agénor': 2.643824027072758e-05,
'agésil': 2.643824027072758e-05,
'ah': 2.643824027072758e-05,
'ahi': 2.643824027072758e-05,
'ahihu': 2.643824027072758e-05,
'ahur': 2.643824027072758e-05,
'ahy': 2.643824027072758e-05,
'ai': 2.643824027072758e-05,
'aid': 0.0005287648054145516,
'aident': 2.643824027072758e-05,
'aidion': 2.643824027072758e-05,
'aidon': 2.643824027072758e-05,
'aie': 5.287648054145516e-05,
'aient': 2.643824027072758e-05,
'aigl': 5.287648054145516e-05,
'aigr': 0.00042301184433164127,
'aigrefin': 5.287648054145516e-05,
'aigret': 2.643824027072758e-05,
'aigreur': 5.287648054145516e-05,
'aigu': 5.287648054145516e-05,
'aigueberr': 2.643824027072758e-05,
'aiguill': 5.287648054145516e-05,
'aiguillet': 2.643824027072758e-05,
'aiguillon': 0.00010575296108291032,
'aiguis': 5.287648054145516e-05,
'aigui': 2.643824027072758e-05,

'aigus': 2.643824027072758e-05,
'aih': 2.643824027072758e-05,
'ail': 0.00010575296108291032,
'aill': 5.287648054145516e-05,
'aillent': 2.643824027072758e-05,
'ailleur': 2.643824027072758e-05,
'aim': 0.0008724619289340102,
'aimabl': 7.931472081218273e-05,
'aiment': 2.643824027072758e-05,
'aimion': 2.643824027072758e-05,
'aimiont': 2.643824027072758e-05,
'aimon': 5.287648054145516e-05,
'aimont': 2.643824027072758e-05,
'ain': 0.0001321912013536379,
'ains': 2.643824027072758e-05,
'air': 5.287648054145516e-05,
'airain': 2.643824027072758e-05,
'ais': 0.00021150592216582064,
'aisanc': 2.643824027072758e-05,
'aisi': 2.643824027072758e-05,
'ait': 2.643824027072758e-05,
'aix': 2.643824027072758e-05,
'ajax': 2.643824027072758e-05,
'ajourn': 5.287648054145516e-05,
'ajout': 0.000502326565143824,
'ajoutent': 2.643824027072758e-05,
'ajouton': 2.643824027072758e-05,
'ajust': 0.0002643824027072758,
'ajuston': 2.643824027072758e-05,
'al': 2.643824027072758e-05,
'alain': 2.643824027072758e-05,
'alambic': 2.643824027072758e-05,
'alambiqu': 2.643824027072758e-05,
'alam': 2.643824027072758e-05,
'alaric': 2.643824027072758e-05,
'alarm': 0.00029082064297800337,
'alarment': 2.643824027072758e-05,
'alarmon': 2.643824027072758e-05,
'alban': 2.643824027072758e-05,
'albert': 2.643824027072758e-05,
'albin': 5.287648054145516e-05,
'albion': 2.643824027072758e-05,
'albis': 2.643824027072758e-05,
'albâtr': 2.643824027072758e-05,
'alcalc': 2.643824027072758e-05,
'alcandr': 2.643824027072758e-05,
'alcest': 2.643824027072758e-05,
'alcet': 2.643824027072758e-05,
'alciat': 2.643824027072758e-05,
'alcibiad': 2.643824027072758e-05,
'alcid': 2.643824027072758e-05,
'alcidon': 2.643824027072758e-05,
'alcidor': 2.643824027072758e-05,
'alcim': 2.643824027072758e-05,
'alcimed': 2.643824027072758e-05,
'alcimédon': 2.643824027072758e-05,
'alcin': 2.643824027072758e-05,
'alcinous': 2.643824027072758e-05,
'alcion': 7.931472081218273e-05,
'alcip': 2.643824027072758e-05,
'alcipp': 2.643824027072758e-05,
'alcir': 2.643824027072758e-05,
'alcmen': 2.643824027072758e-05,
'alcméon': 2.643824027072758e-05,
'alcomir': 2.643824027072758e-05,
'alcoran': 2.643824027072758e-05,
'alcyon': 5.287648054145516e-05,
'alcôv': 2.643824027072758e-05,
'aldamon': 2.643824027072758e-05,

'ale': 2.643824027072758e-05,
'alecton': 2.643824027072758e-05,
'alenc': 2.643824027072758e-05,
'alent': 5.287648054145516e-05,
'alentour': 2.643824027072758e-05,
'alençon': 2.643824027072758e-05,
'alerin': 2.643824027072758e-05,
'alert': 5.287648054145516e-05,
'alexandr': 7.931472081218273e-05,
'alexandrin': 5.287648054145516e-05,
'alex': 5.287648054145516e-05,
'alezan': 2.643824027072758e-05,
'alfonc': 2.643824027072758e-05,
'alfort': 2.643824027072758e-05,
'alger': 2.643824027072758e-05,
'alguasil': 2.643824027072758e-05,
'algebr': 2.643824027072758e-05,
'ali': 5.287648054145516e-05,
'alii': 2.643824027072758e-05,
'aliment': 5.287648054145516e-05,
'aliquando': 2.643824027072758e-05,
'alison': 2.643824027072758e-05,
'alix': 2.643824027072758e-05,
'alien': 5.287648054145516e-05,
'alla': 2.643824027072758e-05,
'allai': 2.643824027072758e-05,
'allaient': 2.643824027072758e-05,
'allais': 2.643824027072758e-05,
'allait': 0.00010575296108291032,
'allant': 2.643824027072758e-05,
'allarm': 5.287648054145516e-05,
'allas': 2.643824027072758e-05,
'allass': 5.287648054145516e-05,
'alle': 5.287648054145516e-05,
'allegretto': 2.643824027072758e-05,
'allegrezz': 2.643824027072758e-05,
'allemagn': 2.643824027072758e-05,
'allemand': 7.931472081218273e-05,
'aller': 2.643824027072758e-05,
'allez': 2.643824027072758e-05,
'alliag': 2.643824027072758e-05,
'alli': 0.00021150592216582064,
'allianc': 5.287648054145516e-05,
'allic': 2.643824027072758e-05,
'allion': 2.643824027072758e-05,
'allobrog': 2.643824027072758e-05,
'alloit': 2.643824027072758e-05,
'allomic': 2.643824027072758e-05,
'allong': 0.00018506768189509307,
'allongent': 2.643824027072758e-05,
'allon': 2.643824027072758e-05,
'allont': 2.643824027072758e-05,
'allum': 0.00042301184433164127,
'allument': 2.643824027072758e-05,
'allumet': 5.287648054145516e-05,
'allumon': 2.643824027072758e-05,
'allur': 5.287648054145516e-05,
'allus': 5.287648054145516e-05,
'allâm': 2.643824027072758e-05,
'allât': 2.643824027072758e-05,
'allech': 2.643824027072758e-05,
'alleg': 0.0001321912013536379,
'allegn': 2.643824027072758e-05,
'allègu': 2.643824027072758e-05,
'allèrent': 2.643824027072758e-05,
'allé': 0.00010575296108291032,
'allége': 2.643824027072758e-05,
'allégor': 2.643824027072758e-05,
'allégress': 2.643824027072758e-05,

'allégu': 0.00018506768189509307,
'allélu': 2.643824027072758e-05,
'almanach': 5.287648054145516e-05,
'almanzor': 2.643824027072758e-05,
'almaviv': 2.643824027072758e-05,
'almedor': 2.643824027072758e-05,
'almen': 2.643824027072758e-05,
'almodobar': 2.643824027072758e-05,
'aloi': 2.643824027072758e-05,
'alons': 2.643824027072758e-05,
'alonzo': 2.643824027072758e-05,
'alor': 2.643824027072758e-05,
'alouet': 2.643824027072758e-05,
'aloyau': 2.643824027072758e-05,
'alpe': 2.643824027072758e-05,
'alphabet': 5.287648054145516e-05,
'alphons': 2.643824027072758e-05,
'alphé': 2.643824027072758e-05,
'alphénor': 2.643824027072758e-05,
'alquif': 2.643824027072758e-05,
'altdorff': 2.643824027072758e-05,
'alterc': 5.287648054145516e-05,
'altern': 5.287648054145516e-05,
'altess': 2.643824027072758e-05,
'alti': 2.643824027072758e-05,
'altier': 7.931472081218273e-05,
'altis': 2.643824027072758e-05,
'altomir': 2.643824027072758e-05,
'altro': 2.643824027072758e-05,
'altrui': 2.643824027072758e-05,
'alter': 0.00029082064297800337,
'altèrent': 2.643824027072758e-05,
'alvar': 5.287648054145516e-05,
'alvarad': 2.643824027072758e-05,
'alvarados': 2.643824027072758e-05,
'alvares': 2.643824027072758e-05,
'alve': 2.643824027072758e-05,
'alzaid': 2.643824027072758e-05,
'alzir': 2.643824027072758e-05,
'alzond': 2.643824027072758e-05,
'amabil': 2.643824027072758e-05,
'amad': 2.643824027072758e-05,
'amadou': 2.643824027072758e-05,
'amaigr': 2.643824027072758e-05,
'amalec': 2.643824027072758e-05,
'amalécit': 2.643824027072758e-05,
'aman': 5.287648054145516e-05,
'amand': 2.643824027072758e-05,
'amant': 0.00010575296108291032,
'amar': 2.643824027072758e-05,
'amarill': 5.287648054145516e-05,
'amart': 2.643824027072758e-05,
'amas': 0.00010575296108291032,
'amass': 0.0001321912013536379,
'amassent': 2.643824027072758e-05,
'amateur': 5.287648054145516e-05,
'amatont': 2.643824027072758e-05,
'amatrox': 2.643824027072758e-05,
'amazon': 5.287648054145516e-05,
'ambag': 2.643824027072758e-05,
'ambassad': 5.287648054145516e-05,
'ambassadeur': 5.287648054145516e-05,
'ambigu': 5.287648054145516e-05,
'ambigus': 2.643824027072758e-05,
'ambiguït': 2.643824027072758e-05,
'ambiti': 7.931472081218273e-05,
'ambit': 5.287648054145516e-05,
'ambition': 2.643824027072758e-05,
'ambre': 2.643824027072758e-05,

'ambrois': 5.287648054145516e-05,
'ambros': 2.643824027072758e-05,
'ambul': 5.287648054145516e-05,
'ambulator': 2.643824027072758e-05,
'ame': 5.287648054145516e-05,
'amelin': 2.643824027072758e-05,
'amen': 0.0005816412859560067,
'amend': 0.00010575296108291032,
'amenon': 2.643824027072758e-05,
'amer': 0.00015862944162436547,
'amerin': 2.643824027072758e-05,
'amertum': 2.643824027072758e-05,
'amestr': 2.643824027072758e-05,
'amet': 2.643824027072758e-05,
'ameubl': 2.643824027072758e-05,
'ameut': 5.287648054145516e-05,
'amez': 2.643824027072758e-05,
'ami': 7.931472081218273e-05,
'amiabl': 5.287648054145516e-05,
'amico': 2.643824027072758e-05,
'amidor': 2.643824027072758e-05,
'amien': 2.643824027072758e-05,
'amilcar': 2.643824027072758e-05,
'aminus': 2.643824027072758e-05,
'amint': 5.287648054145516e-05,
'aminth': 2.643824027072758e-05,
'amiqui': 2.643824027072758e-05,
'amiral': 7.931472081218273e-05,
'amis': 2.643824027072758e-05,
'amiti': 5.287648054145516e-05,
'ammon': 2.643824027072758e-05,
'amnon': 2.643824027072758e-05,
'amo': 2.643824027072758e-05,
'amoindr': 5.287648054145516e-05,
'amoll': 0.00023794416243654823,
'amor': 7.931472081218273e-05,
'amorc': 7.931472081218273e-05,
'amore': 2.643824027072758e-05,
'amores': 2.643824027072758e-05,
'amort': 7.931472081218273e-05,
'amour': 0.00015862944162436547,
'amouret': 5.287648054145516e-05,
'amphib': 2.643824027072758e-05,
'amphict': 2.643824027072758e-05,
'amphigourist': 2.643824027072758e-05,
'amphion': 2.643824027072758e-05,
'amphitryon': 5.287648054145516e-05,
'amphrys': 2.643824027072758e-05,
'amphytrion': 2.643824027072758e-05,
'ample': 5.287648054145516e-05,
'amplement': 2.643824027072758e-05,
'amplif': 2.643824027072758e-05,
'ampoul': 2.643824027072758e-05,
'amsterdam': 2.643824027072758e-05,
'amurat': 2.643824027072758e-05,
'amus': 0.0006345177664974619,
'amusent': 2.643824027072758e-05,
'amuson': 2.643824027072758e-05,
'amynt': 2.643824027072758e-05,
'amèment': 2.643824027072758e-05,
'amé': 5.287648054145516e-05,
'amed': 2.643824027072758e-05,
'amel': 2.643824027072758e-05,
'amélior': 2.643824027072758e-05,
'aménaïd': 2.643824027072758e-05,
'aménoph': 2.643824027072758e-05,
'américain': 5.287648054145516e-05,
'amérin': 2.643824027072758e-05,
'an': 5.287648054145516e-05,

'anachron': 2.643824027072758e-05,
'anacréon': 2.643824027072758e-05,
'anagramm': 2.643824027072758e-05,
'analog': 2.643824027072758e-05,
'analys': 7.931472081218273e-05,
'analyson': 2.643824027072758e-05,
'anarch': 2.643824027072758e-05,
'anathem': 2.643824027072758e-05,
'anatom': 2.643824027072758e-05,
'anatomis': 2.643824027072758e-05,
'anaxagor': 2.643824027072758e-05,
'anaxandr': 2.643824027072758e-05,
'anaxarqu': 2.643824027072758e-05,
'anaximandr': 2.643824027072758e-05,
'anaximen': 2.643824027072758e-05,
'anche': 2.643824027072758e-05,
'anchis': 2.643824027072758e-05,
'anchois': 2.643824027072758e-05,
'ancien': 0.0001321912013536379,
'anciennet': 2.643824027072758e-05,
'anci': 2.643824027072758e-05,
'ancor': 2.643824027072758e-05,
'ancre': 2.643824027072758e-05,
'ancré': 5.287648054145516e-05,
'ancêtr': 5.287648054145516e-05,
'andalous': 2.643824027072758e-05,
'andar': 2.643824027072758e-05,
'andre': 2.643824027072758e-05,
'andrien': 2.643824027072758e-05,
'andrigen': 2.643824027072758e-05,
'andrinopl': 2.643824027072758e-05,
'andrion': 2.643824027072758e-05,
'androclid': 2.643824027072758e-05,
'andromach': 2.643824027072758e-05,
'andromaqu': 2.643824027072758e-05,
'andromed': 2.643824027072758e-05,
'andronic': 2.643824027072758e-05,
'andres': 2.643824027072758e-05,
'andré': 7.931472081218273e-05,
'anecdot': 7.931472081218273e-05,
'anen': 2.643824027072758e-05,
'ange': 5.287648054145516e-05,
'angelo': 2.643824027072758e-05,
'anglais': 7.931472081218273e-05,
'angle': 5.287648054145516e-05,
'angleterr': 2.643824027072758e-05,
'angloman': 2.643824027072758e-05,
'angoiss': 5.287648054145516e-05,
'angol': 2.643824027072758e-05,
'angoulêm': 2.643824027072758e-05,
'anguill': 5.287648054145516e-05,
'angust': 2.643824027072758e-05,
'angel': 5.287648054145516e-05,
'anicroch': 2.643824027072758e-05,
'anim': 0.00034369712351945856,
'animal': 7.931472081218273e-05,
'animent': 2.643824027072758e-05,
'animon': 2.643824027072758e-05,
'animos': 5.287648054145516e-05,
'anitus': 2.643824027072758e-05,
'annal': 2.643824027072758e-05,
'anne': 2.643824027072758e-05,
'anneau': 5.287648054145516e-05,
'annibal': 2.643824027072758e-05,
'annihil': 2.643824027072758e-05,
'annis': 2.643824027072758e-05,
'annonc': 0.00042301184433164127,
'annoncent': 2.643824027072758e-05,
'annonçon': 2.643824027072758e-05,

'annul': 5.287648054145516e-05,
'anné': 5.287648054145516e-05,
'anobl': 0.00010575296108291032,
'anodin': 5.287648054145516e-05,
'anonym': 5.287648054145516e-05,
'anselm': 2.643824027072758e-05,
'ant': 2.643824027072758e-05,
'antagon': 2.643824027072758e-05,
'antarct': 2.643824027072758e-05,
'ante': 2.643824027072758e-05,
'anthenor': 2.643824027072758e-05,
'anthoin': 2.643824027072758e-05,
'anthony': 2.643824027072758e-05,
'anthropophag': 2.643824027072758e-05,
'anthym': 2.643824027072758e-05,
'anthé': 2.643824027072758e-05,
'anti': 2.643824027072758e-05,
'antichambr': 2.643824027072758e-05,
'anticip': 0.00010575296108291032,
'antidot': 2.643824027072758e-05,
'antigon': 2.643824027072758e-05,
'antigonus': 2.643824027072758e-05,
'antinoü': 2.643824027072758e-05,
'antioch': 2.643824027072758e-05,
'antiochus': 2.643824027072758e-05,
'antiop': 2.643824027072758e-05,
'antipat': 2.643824027072758e-05,
'antipath': 5.287648054145516e-05,
'antipati': 2.643824027072758e-05,
'antipatr': 2.643824027072758e-05,
'antipod': 5.287648054145516e-05,
'antipéristas': 2.643824027072758e-05,
'antiquair': 2.643824027072758e-05,
'antiqu': 0.00010575296108291032,
'antistroph': 2.643824027072758e-05,
'antithes': 5.287648054145516e-05,
'antoin': 2.643824027072758e-05,
'antoinet': 2.643824027072758e-05,
'antoniât': 2.643824027072758e-05,
'antonin': 2.643824027072758e-05,
'antonio': 2.643824027072758e-05,
'antre': 5.287648054145516e-05,
'anté': 2.643824027072758e-05,
'anténor': 2.643824027072758e-05,
'antérieur': 2.643824027072758e-05,
'anx': 2.643824027072758e-05,
'anéant': 0.00015862944162436547,
'aouf': 2.643824027072758e-05,
'août': 2.643824027072758e-05,
'apais': 0.00034369712351945856,
'apaisent': 2.643824027072758e-05,
'apaison': 2.643824027072758e-05,
'apam': 2.643824027072758e-05,
'apanag': 2.643824027072758e-05,
'apart': 2.643824027072758e-05,
'aparçois': 2.643824027072758e-05,
'aparçu': 2.643824027072758e-05,
'apel': 2.643824027072758e-05,
'apennin': 2.643824027072758e-05,
'apen': 2.643824027072758e-05,
'apeps': 2.643824027072758e-05,
'apercev': 7.931472081218273e-05,
'apercevoir': 2.643824027072758e-05,
'apercevon': 2.643824027072758e-05,
'apercevr': 0.00010575296108291032,
'aperçois': 2.643824027072758e-05,
'aperçoit': 2.643824027072758e-05,
'aperço': 2.643824027072758e-05,
'aperçoivent': 2.643824027072758e-05,

'aperçu': 5.287648054145516e-05,
'aperçus': 2.643824027072758e-05,
'aperçut': 2.643824027072758e-05,
'aperçût': 2.643824027072758e-05,
'aplan': 0.00018506768189509307,
'apocalyps': 2.643824027072758e-05,
'apocryph': 2.643824027072758e-05,
'apollon': 5.287648054145516e-05,
'apollonius': 2.643824027072758e-05,
'apolog': 7.931472081218273e-05,
'apologu': 2.643824027072758e-05,
'apophtegm': 2.643824027072758e-05,
'apoplex': 5.287648054145516e-05,
'apostat': 5.287648054145516e-05,
'apostill': 2.643824027072758e-05,
'apostroph': 5.287648054145516e-05,
'apost': 7.931472081218273e-05,
'apothicair': 5.287648054145516e-05,
'apothéos': 2.643824027072758e-05,
'appais': 2.643824027072758e-05,
'appanag': 2.643824027072758e-05,
'appar': 2.643824027072758e-05,
'apparaît': 2.643824027072758e-05,
'apparaître': 2.643824027072758e-05,
'appareil': 5.287648054145516e-05,
'apparent': 0.00021150592216582064,
'apparit': 5.287648054145516e-05,
'appartemen': 2.643824027072758e-05,
'appart': 7.931472081218273e-05,
'apparten': 0.00018506768189509307,
'appartenu': 2.643824027072758e-05,
'apparti': 2.643824027072758e-05,
'appartiendr': 2.643824027072758e-05,
'appartien': 5.287648054145516e-05,
'appartiennent': 2.643824027072758e-05,
'appartient': 2.643824027072758e-05,
'apparu': 5.287648054145516e-05,
'apparut': 2.643824027072758e-05,
'appas': 2.643824027072758e-05,
'appauvr': 7.931472081218273e-05,
'appel': 0.000819585448392555,
'appelet': 2.643824027072758e-05,
'appellent': 2.643824027072758e-05,
'appelon': 2.643824027072758e-05,
'appelont': 2.643824027072758e-05,
'appendr': 2.643824027072758e-05,
'appert': 2.643824027072758e-05,
'appesant': 0.00015862944162436547,
'appian': 2.643824027072758e-05,
'appius': 2.643824027072758e-05,
'applaud': 0.00044945008460236884,
'applic': 5.287648054145516e-05,
'appliqu': 0.0002643824027072758,
'appliquent': 2.643824027072758e-05,
'appliquon': 2.643824027072758e-05,
'appoint': 7.931472081218273e-05,
'apport': 0.0005552030456852792,
'apportent': 2.643824027072758e-05,
'apporton': 2.643824027072758e-05,
'appos': 2.643824027072758e-05,
'appren': 0.00018506768189509307,
'apprend': 5.287648054145516e-05,
'apprendr': 0.00023794416243654823,
'apprendron': 2.643824027072758e-05,
'apprendront': 2.643824027072758e-05,
'apprennent': 2.643824027072758e-05,
'apprenon': 2.643824027072758e-05,
'apprent': 5.287648054145516e-05,
'apprentissag': 2.643824027072758e-05,

'appriand': 2.643824027072758e-05,
'apprirent': 2.643824027072758e-05,
'appris': 7.931472081218273e-05,
'apprit': 2.643824027072758e-05,
'apprivois': 0.00010575296108291032,
'approb': 7.931472081218273e-05,
'approch': 0.00044945008460236884,
'approchent': 2.643824027072758e-05,
'approchon': 2.643824027072758e-05,
'approfond': 0.00015862944162436547,
'appropr': 2.643824027072758e-05,
'appropri': 5.287648054145516e-05,
'approuv': 0.0006080795262267344,
'approuvent': 2.643824027072758e-05,
'approuvon': 2.643824027072758e-05,
'apprec': 2.643824027072758e-05,
'appréci': 5.287648054145516e-05,
'appréhend': 0.00023794416243654823,
'appréhendent': 2.643824027072758e-05,
'appréhendon': 2.643824027072758e-05,
'appréhens': 5.287648054145516e-05,
'apprêt': 0.00031725888324873094,
'apprêtent': 2.643824027072758e-05,
'apprêton': 2.643824027072758e-05,
'apprît': 5.287648054145516e-05,
'appui': 0.00037013536379018613,
'appuient': 2.643824027072758e-05,
'appuis': 2.643824027072758e-05,
'appuy': 2.643824027072758e-05,
'appuyon': 2.643824027072758e-05,
'appât': 5.287648054145516e-05,
'appet': 7.931472081218273e-05,
'appétit': 2.643824027072758e-05,
'apre': 2.643824027072758e-05,
'apries': 2.643824027072758e-05,
'apres': 2.643824027072758e-05,
'après': 2.643824027072758e-05,
'apôtr': 5.287648054145516e-05,
'aquilien': 2.643824027072758e-05,
'aquilin': 2.643824027072758e-05,
'aquillin': 2.643824027072758e-05,
'aquilon': 5.287648054145516e-05,
'aquitain': 2.643824027072758e-05,
'ar': 5.287648054145516e-05,
'arab': 0.00010575296108291032,
'arabel': 2.643824027072758e-05,
'aragon': 2.643824027072758e-05,
'araign': 2.643824027072758e-05,
'aram': 2.643824027072758e-05,
'aramint': 2.643824027072758e-05,
'aramont': 2.643824027072758e-05,
'aras': 2.643824027072758e-05,
'arasp': 2.643824027072758e-05,
'arax': 5.287648054145516e-05,
'araxid': 2.643824027072758e-05,
'araxes': 2.643824027072758e-05,
'arbac': 2.643824027072758e-05,
'arbact': 2.643824027072758e-05,
'arbaces': 2.643824027072758e-05,
'arbalet': 2.643824027072758e-05,
'arbant': 2.643824027072758e-05,
'arbas': 2.643824027072758e-05,
'arbat': 2.643824027072758e-05,
'arbazan': 2.643824027072758e-05,
'arbitrair': 5.287648054145516e-05,
'arbitr': 5.287648054145516e-05,
'arbois': 2.643824027072758e-05,
'arbor': 0.00010575296108291032,
'arbre': 5.287648054145516e-05,


```
'arbrisseau': 5.287648054145516e-05,
'arc': 5.287648054145516e-05,
'arcagamb': 2.643824027072758e-05,
'arcas': 2.643824027072758e-05,
'arce': 2.643824027072758e-05,
'archang': 2.643824027072758e-05,
'arche': 2.643824027072758e-05,
'archel': 2.643824027072758e-05,
'archelaus': 2.643824027072758e-05,
'archer': 5.287648054145516e-05,
'archerot': 2.643824027072758e-05,
'archet': 2.643824027072758e-05,
'archevêqu': 2.643824027072758e-05,
'archi': 2.643824027072758e-05,
'archidemid': 2.643824027072758e-05,
'archipel': 2.643824027072758e-05,
'architalass': 2.643824027072758e-05,
'architect': 2.643824027072758e-05,
'architectur': 2.643824027072758e-05,
'architrav': 2.643824027072758e-05,
'archiv': 2.643824027072758e-05,
'archont': 2.643824027072758e-05,
'archélaü': 2.643824027072758e-05,
'arctodem': 2.643824027072758e-05,
'arde': 2.643824027072758e-05,
'ardemment': 2.643824027072758e-05,
'arden': 5.287648054145516e-05,
'ardent': 0.00010575296108291032,
'ardeuil': 2.643824027072758e-05,
'ardeur': 5.287648054145516e-05,
'ardez': 2.643824027072758e-05,
'ardois': 2.643824027072758e-05,
'ardor': 2.643824027072758e-05,
'ardé': 2.643824027072758e-05,
'arethus': 2.643824027072758e-05,
'argan': 2.643824027072758e-05,
'argant': 5.287648054145516e-05,
'argatiphontid': 2.643824027072758e-05,
'argen': 7.931472081218273e-05,
'argenson': 2.643824027072758e-05,
'argent': 5.287648054145516e-05,
'argentac': 2.643824027072758e-05,
'argental': 2.643824027072758e-05,
'argentin': 2.643824027072758e-05,
'argid': 2.643824027072758e-05,
'argi': 2.643824027072758e-05,
'argien': 5.287648054145516e-05,
'argil': 2.643824027072758e-05,
'argin': 2.643824027072758e-05,
'argir': 2.643824027072758e-05,
'argos': 2.643824027072758e-05,
'argoulet': 2.643824027072758e-05,
...}
```

Q2: For each genre, it is possible to generate a “profile”, in the form of a single vector representing the entire set of plays corresponding to this genre. Build such a profile for each of the three genres (Comedy, Tragedy and Tragicomedy).

In [41]:

```
plays_com, titles_com, genres_com = [], [], []
authors_com, years_com = [], []

plays_tra, titles_tra, genres_tra = [], [], []
authors_tra, years_tra = [], []

plays_tra_com, titles_tra_com, genres_tra_com = [], [], []
authors_tra_com, years_tra_com = [], []
for fn in os.scandir('theatre-classique'):
    # Only include XML files
```

```

if not fn.name.endswith('.xml'):
    continue
tree = lxml.etree.parse(fn.path)
genre = tree.find('//genre')
title = tree.find('//title')
author = tree.find('//author')
year = tree.find('//date')
if genre is not None and genre.text == 'Comédie':
    lines = []
    for line in tree.xpath('//l//p'):
        lines.append(' '.join(line.itertext()))
    text = '\n'.join(lines)
    plays_com.append(text)
    genres_com.append(genre.text)
    titles_com.append(title.text)
    authors_com.append(author.text)
    if year is not None:
        years_com.append(year.text)
elif genre is not None and genre.text == 'Tragédie':
    lines = []
    for line in tree.xpath('//l//p'):
        lines.append(' '.join(line.itertext()))
    text = '\n'.join(lines)
    plays_tra.append(text)
    genres_tra.append(genre.text)
    titles_tra.append(title.text)
    authors_tra.append(author.text)
    if year is not None:
        years_tra.append(year.text)
elif genre is not None and genre.text == 'Tragi-comédie':
    lines = []
    for line in tree.xpath('//l//p'):
        lines.append(' '.join(line.itertext()))
    text = '\n'.join(lines)
    plays_tra_com.append(text)
    genres_tra_com.append(genre.text)
    titles_tra_com.append(title.text)
    authors_tra_com.append(author.text)
    if year is not None:
        years_tra_com.append(year.text)

plays_com_tok = [preprocess_text(play, 'French') for play in plays_com]
plays_tra_tok = [preprocess_text(play, 'French') for play in plays_tra]
plays_tra_com_tok = [preprocess_text(play, 'French') for play in plays_tra_com]
vocabulary_com = extract_vocabulary(plays_com_tok, min_count=2)
vocabulary_tra = extract_vocabulary(plays_tra_tok, min_count=2)
vocabulary_tra_com = extract_vocabulary(plays_tra_com_tok, min_count=2)

```

In [42]:

```

stemmed_vocabulary_com = [fr.stem(word) for word in vocabulary_com]
counts = dict()
for word in stemmed_vocabulary_com:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_com = counts
for key, value in counts.items():
    tf_dict_com[key] = value / len(stemmed_vocabulary_com)
tf_dict_com

```

Out[42]:

```

{'a': 3.389141191622043e-05,
 'aah': 3.389141191622043e-05,
 'ab': 3.389141191622043e-05,
 'abaiss': 6.778282383244086e-05,
 'abaissement': 3.389141191622043e-05,
 'abandon': 0.0006100454144919677,

```

'abandonnent': 3.389141191622043e-05,
'abandonnon': 3.389141191622043e-05,
'abat': 6.778282383244086e-05,
'abatt': 3.389141191622043e-05,
'abattr': 3.389141191622043e-05,
'abattu': 6.778282383244086e-05,
'abattus': 3.389141191622043e-05,
'abbess': 3.389141191622043e-05,
'abbé': 6.778282383244086e-05,
'abces': 3.389141191622043e-05,
'abder': 3.389141191622043e-05,
'abdérit': 3.389141191622043e-05,
'abec': 3.389141191622043e-05,
'abeil': 3.389141191622043e-05,
'abez': 3.389141191622043e-05,
'abhorr': 6.778282383244086e-05,
'abject': 3.389141191622043e-05,
'abjur': 0.00013556564766488172,
'ablativo': 3.389141191622043e-05,
'abois': 3.389141191622043e-05,
'abol': 6.778282383244086e-05,
'abomin': 0.00010167423574866129,
'abond': 0.00016945705958110215,
'abon': 3.389141191622043e-05,
'abord': 0.00027113129532976344,
'abordag': 3.389141191622043e-05,
'abordent': 3.389141191622043e-05,
'abordon': 3.389141191622043e-05,
'abouch': 6.778282383244086e-05,
'aboumékares': 3.389141191622043e-05,
'about': 0.000237239883413543,
'aboi': 0.00010167423574866129,
'abraham': 3.389141191622043e-05,
'abre': 3.389141191622043e-05,
'abreuv': 3.389141191622043e-05,
'abri': 3.389141191622043e-05,
'abreg': 0.00020334847149732258,
'abrégeon': 3.389141191622043e-05,
'absenc': 6.778282383244086e-05,
'absent': 0.00013556564766488172,
'absit': 3.389141191622043e-05,
'absolu': 0.00013556564766488172,
'absolus': 3.389141191622043e-05,
'absorb': 3.389141191622043e-05,
'absoudr': 3.389141191622043e-05,
'absous': 3.389141191622043e-05,
'absout': 3.389141191622043e-05,
'absten': 6.778282383244086e-05,
'abstien': 3.389141191622043e-05,
'abstinent': 3.389141191622043e-05,
'abstract': 3.389141191622043e-05,
'abstrait': 3.389141191622043e-05,
'absurd': 0.00013556564766488172,
'abus': 0.0006100454144919677,
'abusent': 3.389141191622043e-05,
'abuson': 3.389141191622043e-05,
'abîm': 0.00020334847149732258,
'ac': 3.389141191622043e-05,
'acab': 3.389141191622043e-05,
'academ': 0.00016945705958110215,
'acant': 3.389141191622043e-05,
'acanth': 3.389141191622043e-05,
'acariâtr': 3.389141191622043e-05,
'acast': 3.389141191622043e-05,
'accabl': 0.00044058835491086557,
'accablent': 3.389141191622043e-05,
'accent': 6.778282383244086e-05,
'accept': 0.0005422625906595269,
'accepton': 3.389141191622043e-05,

'accessibl': 3.389141191622043e-05,
'accessoir': 6.778282383244086e-05,
'accident': 6.778282383244086e-05,
'accin': 3.389141191622043e-05,
'accoint': 3.389141191622043e-05,
'accolad': 3.389141191622043e-05,
'accommod': 0.0006439368264081882,
'accommodent': 3.389141191622043e-05,
'accommodon': 3.389141191622043e-05,
'accomod': 3.389141191622043e-05,
'accompagn': 0.0003728055310784247,
'accompagnent': 3.389141191622043e-05,
'accompl': 0.0003728055310784247,
'accord': 0.0006100454144919677,
'accordent': 3.389141191622043e-05,
'accordon': 3.389141191622043e-05,
'accort': 3.389141191622043e-05,
'accost': 6.778282383244086e-05,
'accouch': 0.00020334847149732258,
'accour': 0.00013556564766488172,
'accourent': 3.389141191622043e-05,
'accourt': 3.389141191622043e-05,
'accouru': 6.778282383244086e-05,
'accout': 6.778282383244086e-05,
'accoutr': 3.389141191622043e-05,
'accoutum': 0.0003728055310784247,
'accoutu': 3.389141191622043e-05,
'accroch': 0.00010167423574866129,
'accrochent': 3.389141191622043e-05,
'accroir': 3.389141191622043e-05,
'accrois': 3.389141191622043e-05,
'accroiss': 3.389141191622043e-05,
'accroît': 3.389141191622043e-05,
'accroître': 3.389141191622043e-05,
'accru': 3.389141191622043e-05,
'accrut': 3.389141191622043e-05,
'accrédit': 3.389141191622043e-05,
'accueil': 0.00016945705958110215,
'accueillent': 3.389141191622043e-05,
'accumul': 6.778282383244086e-05,
'accus': 0.0006778282383244086,
'accusent': 3.389141191622043e-05,
'accuson': 3.389141191622043e-05,
'acced': 3.389141191622043e-05,
'acces': 3.389141191622043e-05,
'acharn': 0.00016945705958110215,
'achat': 3.389141191622043e-05,
'achemin': 6.778282383244086e-05,
'achet': 0.0005422625906595269,
'acheteur': 6.778282383244086e-05,
'acheton': 3.389141191622043e-05,
'achev': 0.0007456110621568494,
'achevon': 3.389141191622043e-05,
'achil': 3.389141191622043e-05,
'achill': 6.778282383244086e-05,
'achopp': 3.389141191622043e-05,
'achètent': 3.389141191622043e-05,
'achéron': 3.389141191622043e-05,
'acid': 3.389141191622043e-05,
'aci': 3.389141191622043e-05,
'acoquin': 3.389141191622043e-05,
'acord': 3.389141191622043e-05,
'acout': 0.00010167423574866129,
'acquerr': 3.389141191622043e-05,
'acquier': 3.389141191622043e-05,
'acquiert': 3.389141191622043e-05,
'acquis': 6.778282383244086e-05,
'acquisit': 6.778282383244086e-05,
'acquit': 3.389141191622043e-05,

'acquitt': 0.0003728055310784247,
'acquittent': 3.389141191622043e-05,
'acquer': 0.00010167423574866129,
'acquéreur': 3.389141191622043e-05,
'acquêt': 6.778282383244086e-05,
'acte': 6.778282383244086e-05,
'acteur': 6.778282383244086e-05,
'actif': 6.778282383244086e-05,
'action': 6.778282383244086e-05,
'activ': 6.778282383244086e-05,
'actric': 6.778282383244086e-05,
'actuel': 0.00010167423574866129,
'actéon': 3.389141191622043e-05,
'ad': 3.389141191622043e-05,
'adag': 3.389141191622043e-05,
'adam': 6.778282383244086e-05,
'adamir': 3.389141191622043e-05,
'adapt': 3.389141191622043e-05,
'addit': 6.778282383244086e-05,
'adest': 3.389141191622043e-05,
'adieu': 3.389141191622043e-05,
'adieux': 3.389141191622043e-05,
'adin': 3.389141191622043e-05,
'adio': 3.389141191622043e-05,
'adject': 3.389141191622043e-05,
'adjoint': 3.389141191622043e-05,
'adjug': 6.778282383244086e-05,
'admet': 3.389141191622043e-05,
'admettr': 3.389141191622043e-05,
'administr': 3.389141191622043e-05,
'admir': 0.0006439368264081882,
'admiron': 3.389141191622043e-05,
'admis': 6.778282383244086e-05,
'adolescent': 6.778282383244086e-05,
'adonc': 3.389141191622043e-05,
'adon': 0.00013556564766488172,
'adopt': 0.00020334847149732258,
'ador': 0.000711719650240629,
'adorec': 3.389141191622043e-05,
'adorent': 3.389141191622043e-05,
'adoro': 3.389141191622043e-05,
'adouc': 0.00040669694299464516,
'adress': 0.00044058835491086557,
'adressent': 3.389141191622043e-05,
'adresson': 3.389141191622043e-05,
'adrian': 3.389141191622043e-05,
'adrien': 3.389141191622043e-05,
'adroit': 0.00013556564766488172,
'adul': 6.778282383244086e-05,
'adulter': 3.389141191622043e-05,
'adven': 3.389141191622043e-05,
'advenu': 3.389141191622043e-05,
'adversair': 6.778282383244086e-05,
'advers': 0.00010167423574866129,
'advien': 3.389141191622043e-05,
'advient': 3.389141191622043e-05,
'advocatus': 3.389141191622043e-05,
'adélaïd': 3.389141191622043e-05,
'aegin': 3.389141191622043e-05,
'afec': 3.389141191622043e-05,
'affabl': 6.778282383244086e-05,
'affad': 3.389141191622043e-05,
'affaibl': 0.00027113129532976344,
'affair': 0.00010167423574866129,
'affam': 0.00013556564766488172,
'affect': 0.00040669694299464516,
'affectent': 3.389141191622043e-05,
'affection': 6.778282383244086e-05,
'afferm': 0.00020334847149732258,

'affich': 0.000237239883413543,
'affichent': 3.389141191622043e-05,
'affi': 3.389141191622043e-05,
'affil': 3.389141191622043e-05,
'affin': 3.389141191622043e-05,
'affiquet': 3.389141191622043e-05,
'affirm': 0.00010167423574866129,
'afflict': 6.778282383244086e-05,
'afflig': 0.000474479766827086,
'affligent': 3.389141191622043e-05,
'affluenc': 3.389141191622043e-05,
'affol': 3.389141191622043e-05,
'affranch': 0.00016945705958110215,
'affreux': 0.00010167423574866129,
'affront': 0.00020334847149732258,
'affronteur': 3.389141191622043e-05,
'affubl': 3.389141191622043e-05,
'affût': 3.389141191622043e-05,
'afin': 3.389141191622043e-05,
'afoir': 3.389141191622043e-05,
'afre': 3.389141191622043e-05,
'africain': 6.778282383244086e-05,
'afriqu': 3.389141191622043e-05,
'aga': 3.389141191622043e-05,
'agac': 0.00010167423574866129,
'agamemnon': 3.389141191622043e-05,
'agat': 3.389141191622043e-05,
'agath': 3.389141191622043e-05,
'age': 3.389141191622043e-05,
'agenc': 3.389141191622043e-05,
'agend': 3.389141191622043e-05,
'agent': 0.00010167423574866129,
'ager': 3.389141191622043e-05,
'aggrav': 3.389141191622043e-05,
'agi': 3.389141191622043e-05,
'agil': 6.778282383244086e-05,
'agioteur': 3.389141191622043e-05,
'agir': 0.00016945705958110215,
'agiron': 3.389141191622043e-05,
'agis': 3.389141191622043e-05,
'agiss': 0.00016945705958110215,
'agissent': 3.389141191622043e-05,
'agisson': 3.389141191622043e-05,
'agit': 0.0003728055310784247,
'agitent': 3.389141191622043e-05,
'agnant': 3.389141191622043e-05,
'agneau': 3.389141191622043e-05,
'agnelet': 3.389141191622043e-05,
'agnes': 3.389141191622043e-05,
'agné': 3.389141191622043e-05,
'agon': 3.389141191622043e-05,
'agonis': 0.00010167423574866129,
'agraf': 3.389141191622043e-05,
'agrand': 6.778282383244086e-05,
'agresseur': 3.389141191622043e-05,
'agriabl': 3.389141191622043e-05,
'agricultur': 3.389141191622043e-05,
'agri': 3.389141191622043e-05,
'agripp': 3.389141191622043e-05,
'agrippin': 3.389141191622043e-05,
'agré': 0.00020334847149732258,
'agréabl': 0.00010167423574866129,
'agrément': 6.778282383244086e-05,
'aguerr': 0.00013556564766488172,
'aguet': 3.389141191622043e-05,
'agel': 3.389141191622043e-05,
'agénor': 3.389141191622043e-05,
'ah': 3.389141191622043e-05,
'ahi': 3.389141191622043e-05,

'ahihu': 3.389141191622043e-05,
'ahur': 3.389141191622043e-05,
'ahy': 3.389141191622043e-05,
'ai': 3.389141191622043e-05,
'aid': 0.0005761540025757473,
'aident': 3.389141191622043e-05,
'aidon': 3.389141191622043e-05,
'aie': 6.778282383244086e-05,
'aient': 3.389141191622043e-05,
'aigl': 6.778282383244086e-05,
'aigr': 0.0003389141191622043,
'aigrefin': 6.778282383244086e-05,
'aigret': 3.389141191622043e-05,
'aigreur': 6.778282383244086e-05,
'aigu': 3.389141191622043e-05,
'aigueberr': 3.389141191622043e-05,
'aiguill': 6.778282383244086e-05,
'aiguillet': 3.389141191622043e-05,
'aiguillon': 6.778282383244086e-05,
'aiguis': 6.778282383244086e-05,
'aigui': 3.389141191622043e-05,
'aigus': 3.389141191622043e-05,
'aih': 3.389141191622043e-05,
'ail': 0.00013556564766488172,
'aill': 6.778282383244086e-05,
'aillent': 3.389141191622043e-05,
'ailleur': 3.389141191622043e-05,
'aim': 0.001016742357486613,
'aimabl': 0.00010167423574866129,
'aiment': 3.389141191622043e-05,
'aimion': 3.389141191622043e-05,
'aimiont': 3.389141191622043e-05,
'aimon': 6.778282383244086e-05,
'aimont': 3.389141191622043e-05,
'ain': 0.00010167423574866129,
'ains': 3.389141191622043e-05,
'air': 6.778282383244086e-05,
'airain': 3.389141191622043e-05,
'ais': 0.00027113129532976344,
'aisanc': 3.389141191622043e-05,
'aisi': 3.389141191622043e-05,
'ait': 3.389141191622043e-05,
'aix': 3.389141191622043e-05,
'ajax': 3.389141191622043e-05,
'ajourn': 6.778282383244086e-05,
'ajout': 0.00044058835491086557,
'ajoutent': 3.389141191622043e-05,
'ajouton': 3.389141191622043e-05,
'ajust': 0.0003389141191622043,
'ajuston': 3.389141191622043e-05,
'al': 3.389141191622043e-05,
'alain': 3.389141191622043e-05,
'alambic': 3.389141191622043e-05,
'alambiqu': 3.389141191622043e-05,
'alarm': 0.000237239883413543,
'alarment': 3.389141191622043e-05,
'alban': 3.389141191622043e-05,
'albert': 3.389141191622043e-05,
'albin': 3.389141191622043e-05,
'albâtr': 3.389141191622043e-05,
'alcalc': 3.389141191622043e-05,
'alcandr': 3.389141191622043e-05,
'alcest': 3.389141191622043e-05,
'alciat': 3.389141191622043e-05,
'alcibiad': 3.389141191622043e-05,
'alcid': 3.389141191622043e-05,
'alcidon': 3.389141191622043e-05,
'alcidor': 3.389141191622043e-05,
'alcion': 3.389141191622043e-05,

'alcip': 3.389141191622043e-05,
'alcipp': 3.389141191622043e-05,
'alcmen': 3.389141191622043e-05,
'alcoran': 3.389141191622043e-05,
'alcyon': 3.389141191622043e-05,
'alcôv': 3.389141191622043e-05,
'ale': 3.389141191622043e-05,
'alecton': 3.389141191622043e-05,
'alenc': 3.389141191622043e-05,
'alentour': 3.389141191622043e-05,
'alert': 6.778282383244086e-05,
'alexandr': 6.778282383244086e-05,
'alexandrin': 3.389141191622043e-05,
'alex': 3.389141191622043e-05,
'alezan': 3.389141191622043e-05,
'alfonc': 3.389141191622043e-05,
'alfort': 3.389141191622043e-05,
'alger': 3.389141191622043e-05,
'alguasil': 3.389141191622043e-05,
'algebr': 3.389141191622043e-05,
'ali': 3.389141191622043e-05,
'aliment': 6.778282383244086e-05,
'alison': 3.389141191622043e-05,
'alix': 3.389141191622043e-05,
'alien': 6.778282383244086e-05,
'alla': 3.389141191622043e-05,
'allai': 3.389141191622043e-05,
'allaient': 3.389141191622043e-05,
'allais': 3.389141191622043e-05,
'allait': 3.389141191622043e-05,
'allant': 3.389141191622043e-05,
'allas': 3.389141191622043e-05,
'allass': 6.778282383244086e-05,
'alle': 6.778282383244086e-05,
'allegretto': 3.389141191622043e-05,
'allegrezz': 3.389141191622043e-05,
'allemagn': 3.389141191622043e-05,
'allemand': 0.00010167423574866129,
'aller': 3.389141191622043e-05,
'allez': 3.389141191622043e-05,
'allianc': 6.778282383244086e-05,
'alli': 0.000237239883413543,
'allic': 3.389141191622043e-05,
'allion': 3.389141191622043e-05,
'alloit': 3.389141191622043e-05,
'allong': 0.00020334847149732258,
'allongent': 3.389141191622043e-05,
'allon': 3.389141191622043e-05,
'allont': 3.389141191622043e-05,
'allum': 0.000237239883413543,
'allument': 3.389141191622043e-05,
'allumet': 3.389141191622043e-05,
'allur': 6.778282383244086e-05,
'allus': 6.778282383244086e-05,
'allâm': 3.389141191622043e-05,
'allât': 3.389141191622043e-05,
'alleg': 0.00016945705958110215,
'allegr': 3.389141191622043e-05,
'allègu': 3.389141191622043e-05,
'allé': 0.00013556564766488172,
'allége': 3.389141191622043e-05,
'allégor': 3.389141191622043e-05,
'allégress': 3.389141191622043e-05,
'allégu': 0.00013556564766488172,
'allélu': 3.389141191622043e-05,
'almanach': 6.778282383244086e-05,
'almanzor': 3.389141191622043e-05,
'almaviv': 3.389141191622043e-05,
'almedor': 3.389141191622043e-05,

'almen': 3.389141191622043e-05,
'almodobar': 3.389141191622043e-05,
'aloi': 3.389141191622043e-05,
'alonzo': 3.389141191622043e-05,
'alor': 3.389141191622043e-05,
'alouet': 3.389141191622043e-05,
'aloyau': 3.389141191622043e-05,
'alphabet': 6.778282383244086e-05,
'alphons': 3.389141191622043e-05,
'alterc': 6.778282383244086e-05,
'altern': 6.778282383244086e-05,
'altess': 3.389141191622043e-05,
'alti': 3.389141191622043e-05,
'altier': 6.778282383244086e-05,
'altro': 3.389141191622043e-05,
'alter': 0.000237239883413543,
'alvar': 6.778282383244086e-05,
'alvarad': 3.389141191622043e-05,
'alvarados': 3.389141191622043e-05,
'alzir': 3.389141191622043e-05,
'amabil': 3.389141191622043e-05,
'amad': 3.389141191622043e-05,
'amadou': 3.389141191622043e-05,
'amaign': 3.389141191622043e-05,
'amand': 3.389141191622043e-05,
'aman': 3.389141191622043e-05,
'amant': 0.00013556564766488172,
'amar': 3.389141191622043e-05,
'amarill': 3.389141191622043e-05,
'amart': 3.389141191622043e-05,
'amas': 3.389141191622043e-05,
'amass': 0.00013556564766488172,
'amateur': 6.778282383244086e-05,
'amatont': 3.389141191622043e-05,
'amatrox': 3.389141191622043e-05,
'amazon': 3.389141191622043e-05,
'ambag': 3.389141191622043e-05,
'ambassad': 6.778282383244086e-05,
'ambassadeur': 6.778282383244086e-05,
'ambigu': 6.778282383244086e-05,
'ambigus': 3.389141191622043e-05,
'ambiti': 0.00010167423574866129,
'ambit': 3.389141191622043e-05,
'ambition': 3.389141191622043e-05,
'ambre': 3.389141191622043e-05,
'ambrois': 6.778282383244086e-05,
'ambros': 3.389141191622043e-05,
'ambul': 6.778282383244086e-05,
'ambulator': 3.389141191622043e-05,
'ame': 6.778282383244086e-05,
'amelin': 3.389141191622043e-05,
'amen': 0.0006439368264081882,
'amend': 0.00013556564766488172,
'amenon': 3.389141191622043e-05,
'amer': 0.00020334847149732258,
'amertum': 3.389141191622043e-05,
'ameubl': 3.389141191622043e-05,
'ameut': 3.389141191622043e-05,
'amez': 3.389141191622043e-05,
'ami': 0.00010167423574866129,
'amiabl': 6.778282383244086e-05,
'amico': 3.389141191622043e-05,
'amidor': 3.389141191622043e-05,
'amien': 3.389141191622043e-05,
'amilcar': 3.389141191622043e-05,
'amint': 6.778282383244086e-05,
'amiqui': 3.389141191622043e-05,
'amiral': 0.00010167423574866129,
'amis': 3.389141191622043e-05,

'amiti': 6.778282383244086e-05,
'amo': 3.389141191622043e-05,
'amoindr': 3.389141191622043e-05,
'amoll': 0.00010167423574866129,
'amor': 0.00010167423574866129,
'amorc': 0.00010167423574866129,
'amoros': 3.389141191622043e-05,
'amort': 3.389141191622043e-05,
'amour': 0.00020334847149732258,
'amouret': 6.778282383244086e-05,
'amphib': 3.389141191622043e-05,
'amphigourist': 3.389141191622043e-05,
'amphion': 3.389141191622043e-05,
'amphitryon': 6.778282383244086e-05,
'amphytrion': 3.389141191622043e-05,
'ample': 6.778282383244086e-05,
'amplement': 3.389141191622043e-05,
'amplif': 3.389141191622043e-05,
'ampoul': 3.389141191622043e-05,
'amsterdam': 3.389141191622043e-05,
'amus': 0.0007795024740730699,
'amusent': 3.389141191622043e-05,
'amuson': 3.389141191622043e-05,
'amènent': 3.389141191622043e-05,
'amé': 6.778282383244086e-05,
'amed': 3.389141191622043e-05,
'amel': 3.389141191622043e-05,
'amélior': 3.389141191622043e-05,
'an': 6.778282383244086e-05,
'anacréon': 3.389141191622043e-05,
'anagramm': 3.389141191622043e-05,
'analys': 0.00010167423574866129,
'anarch': 3.389141191622043e-05,
'anathem': 3.389141191622043e-05,
'anatom': 3.389141191622043e-05,
'anatomis': 3.389141191622043e-05,
'anaxagor': 3.389141191622043e-05,
'anaxarqu': 3.389141191622043e-05,
'anaximandr': 3.389141191622043e-05,
'anaximen': 3.389141191622043e-05,
'anchis': 3.389141191622043e-05,
'anchois': 3.389141191622043e-05,
'ancien': 0.00013556564766488172,
'anciennet': 3.389141191622043e-05,
'anci': 3.389141191622043e-05,
'ancor': 3.389141191622043e-05,
'ancre': 3.389141191622043e-05,
'ancré': 6.778282383244086e-05,
'ancêtr': 6.778282383244086e-05,
'andalous': 3.389141191622043e-05,
'andar': 3.389141191622043e-05,
'andrien': 3.389141191622043e-05,
'andromed': 3.389141191622043e-05,
'andres': 3.389141191622043e-05,
'andré': 6.778282383244086e-05,
'anecdote': 0.00010167423574866129,
'anen': 3.389141191622043e-05,
'ange': 6.778282383244086e-05,
'angelo': 3.389141191622043e-05,
'anglais': 0.00010167423574866129,
'angle': 6.778282383244086e-05,
'angleterr': 3.389141191622043e-05,
'angoiss': 6.778282383244086e-05,
'angol': 3.389141191622043e-05,
'angoulêm': 3.389141191622043e-05,
'anguill': 6.778282383244086e-05,
'angust': 3.389141191622043e-05,
'angel': 3.389141191622043e-05,
'anicroch': 3.389141191622043e-05,

'anim': 0.0003728055310784247,
'animal': 0.00010167423574866129,
'animos': 3.389141191622043e-05,
'annal': 3.389141191622043e-05,
'anne': 3.389141191622043e-05,
'anneau': 6.778282383244086e-05,
'annibal': 3.389141191622043e-05,
'annis': 3.389141191622043e-05,
'annonc': 0.0003389141191622043,
'annoncent': 3.389141191622043e-05,
'annonçon': 3.389141191622043e-05,
'annul': 6.778282383244086e-05,
'anné': 6.778282383244086e-05,
'anobl': 0.00010167423574866129,
'anodin': 6.778282383244086e-05,
'anonym': 6.778282383244086e-05,
'anselm': 3.389141191622043e-05,
'ant': 3.389141191622043e-05,
'antagon': 3.389141191622043e-05,
'antarct': 3.389141191622043e-05,
'anthony': 3.389141191622043e-05,
'anti': 3.389141191622043e-05,
'antichambr': 3.389141191622043e-05,
'anticip': 6.778282383244086e-05,
'antidot': 3.389141191622043e-05,
'antiochus': 3.389141191622043e-05,
'antipath': 6.778282383244086e-05,
'antipati': 3.389141191622043e-05,
'antipod': 6.778282383244086e-05,
'antipéristas': 3.389141191622043e-05,
'antiquair': 3.389141191622043e-05,
'antiqu': 0.00013556564766488172,
'antistroph': 3.389141191622043e-05,
'antithes': 6.778282383244086e-05,
'antoin': 3.389141191622043e-05,
'antonin': 3.389141191622043e-05,
'antonio': 3.389141191622043e-05,
'antre': 6.778282383244086e-05,
'antérieur': 3.389141191622043e-05,
'anx': 3.389141191622043e-05,
'anéant': 0.00020334847149732258,
'aouf': 3.389141191622043e-05,
'août': 3.389141191622043e-05,
'apais': 0.00027113129532976344,
'apanag': 3.389141191622043e-05,
'aparçois': 3.389141191622043e-05,
'aparçu': 3.389141191622043e-05,
'apel': 3.389141191622043e-05,
'apen': 3.389141191622043e-05,
'apeps': 3.389141191622043e-05,
'apercev': 0.00010167423574866129,
'apercevoir': 3.389141191622043e-05,
'apercevon': 3.389141191622043e-05,
'apercevr': 0.00013556564766488172,
'aperçois': 3.389141191622043e-05,
'aperçoit': 3.389141191622043e-05,
'aperço': 3.389141191622043e-05,
'aperçoivent': 3.389141191622043e-05,
'aperçu': 6.778282383244086e-05,
'aperçus': 3.389141191622043e-05,
'aperçut': 3.389141191622043e-05,
'aperçût': 3.389141191622043e-05,
'aplan': 3.389141191622043e-05,
'apocalyps': 3.389141191622043e-05,
'apocryph': 3.389141191622043e-05,
'apollon': 3.389141191622043e-05,
'apolog': 0.00010167423574866129,
'apologu': 3.389141191622043e-05,
'apophtegm': 3.389141191622043e-05,

'apoplex': 6.778282383244086e-05,
'apostat': 3.389141191622043e-05,
'apostill': 3.389141191622043e-05,
'apostroph': 6.778282383244086e-05,
'apost': 0.00010167423574866129,
'apothicaire': 6.778282383244086e-05,
'apothéos': 3.389141191622043e-05,
'appais': 3.389141191622043e-05,
'appar': 3.389141191622043e-05,
'appareil': 3.389141191622043e-05,
'apparent': 0.00027113129532976344,
'apparit': 3.389141191622043e-05,
'appartement': 3.389141191622043e-05,
'appart': 0.00010167423574866129,
'apparten': 0.000237239883413543,
'appartenu': 3.389141191622043e-05,
'apparti': 3.389141191622043e-05,
'appartiendr': 3.389141191622043e-05,
'appartien': 6.778282383244086e-05,
'appartiennent': 3.389141191622043e-05,
'appartient': 3.389141191622043e-05,
'apparu': 3.389141191622043e-05,
'appas': 3.389141191622043e-05,
'appauvr': 6.778282383244086e-05,
'appel': 0.0008811767098217311,
'appelet': 3.389141191622043e-05,
'appellent': 3.389141191622043e-05,
'appelon': 3.389141191622043e-05,
'appelont': 3.389141191622043e-05,
'appert': 3.389141191622043e-05,
'appesant': 3.389141191622043e-05,
'applaud': 0.00044058835491086557,
'applic': 6.778282383244086e-05,
'appliqu': 0.000237239883413543,
'appliquon': 3.389141191622043e-05,
'appoint': 0.00010167423574866129,
'apport': 0.0006439368264081882,
'apportent': 3.389141191622043e-05,
'apporton': 3.389141191622043e-05,
'appos': 3.389141191622043e-05,
'appren': 0.00020334847149732258,
'apprend': 6.778282383244086e-05,
'apprendr': 0.000237239883413543,
'apprendron': 3.389141191622043e-05,
'apprendront': 3.389141191622043e-05,
'apprennent': 3.389141191622043e-05,
'apprenon': 3.389141191622043e-05,
'apprent': 6.778282383244086e-05,
'apprentissag': 3.389141191622043e-05,
'appriand': 3.389141191622043e-05,
'appris': 0.00010167423574866129,
'apprit': 3.389141191622043e-05,
'apprivois': 0.00010167423574866129,
'approb': 6.778282383244086e-05,
'approch': 0.0003728055310784247,
'approchent': 3.389141191622043e-05,
'approchon': 3.389141191622043e-05,
'approfond': 0.00020334847149732258,
'appropr': 3.389141191622043e-05,
'appropri': 6.778282383244086e-05,
'approuv': 0.0005761540025757473,
'approuvent': 3.389141191622043e-05,
'apprec': 3.389141191622043e-05,
'appréci': 6.778282383244086e-05,
'appréhend': 0.00020334847149732258,
'appréhendent': 3.389141191622043e-05,
'appréhens': 6.778282383244086e-05,
'apprêt': 0.00030502270724598384,
'apprêtent': 3.389141191622043e-05,

'appréton': 3.389141191622043e-05,
'apprît': 6.778282383244086e-05,
'appui': 0.0003389141191622043,
'appuient': 3.389141191622043e-05,
'appuis': 3.389141191622043e-05,
'appât': 6.778282383244086e-05,
'appet': 0.00010167423574866129,
'appétit': 3.389141191622043e-05,
'apres': 3.389141191622043e-05,
'apré': 3.389141191622043e-05,
'apôtr': 6.778282383244086e-05,
'aquilin': 3.389141191622043e-05,
'aquilon': 3.389141191622043e-05,
'ar': 6.778282383244086e-05,
'arab': 0.00013556564766488172,
'arabel': 3.389141191622043e-05,
'aramint': 3.389141191622043e-05,
'aramont': 3.389141191622043e-05,
'arasp': 3.389141191622043e-05,
'arbact': 3.389141191622043e-05,
'arbalet': 3.389141191622043e-05,
'arbitr': 6.778282383244086e-05,
'arbois': 3.389141191622043e-05,
'arbor': 6.778282383244086e-05,
'arbre': 6.778282383244086e-05,
'arbrisseau': 3.389141191622043e-05,
'arc': 6.778282383244086e-05,
'arce': 3.389141191622043e-05,
'archer': 6.778282383244086e-05,
'archerot': 3.389141191622043e-05,
'archet': 3.389141191622043e-05,
'archi': 3.389141191622043e-05,
'archidemid': 3.389141191622043e-05,
'architect': 3.389141191622043e-05,
'architectur': 3.389141191622043e-05,
'architrav': 3.389141191622043e-05,
'archiv': 3.389141191622043e-05,
'archélaü': 3.389141191622043e-05,
'ardemment': 3.389141191622043e-05,
'ardent': 0.00013556564766488172,
'ardeuil': 3.389141191622043e-05,
'ardeur': 6.778282383244086e-05,
'ardez': 3.389141191622043e-05,
'ardor': 3.389141191622043e-05,
'ardé': 3.389141191622043e-05,
'argan': 3.389141191622043e-05,
'argant': 6.778282383244086e-05,
'argatiphontid': 3.389141191622043e-05,
'argenson': 3.389141191622043e-05,
'argent': 6.778282383244086e-05,
'argentac': 3.389141191622043e-05,
'argentin': 3.389141191622043e-05,
'argos': 3.389141191622043e-05,
'argoulet': 3.389141191622043e-05,
'argument': 0.00010167423574866129,
'argus': 3.389141191622043e-05,
'arid': 3.389141191622043e-05,
'ariet': 6.778282383244086e-05,
'aristarqu': 3.389141191622043e-05,
'arist': 3.389141191622043e-05,
'aristem': 6.778282383244086e-05,
'aristid': 3.389141191622043e-05,
'aristion': 3.389141191622043e-05,
'aristipp': 3.389141191622043e-05,
'aristocrat': 6.778282383244086e-05,
'ariston': 3.389141191622043e-05,
'aristophan': 3.389141191622043e-05,
'aristot': 3.389141191622043e-05,
'arithmet': 3.389141191622043e-05,

'arjan': 3.389141191622043e-05,
'arlequin': 6.778282383244086e-05,
'arlequinec': 3.389141191622043e-05,
'arlequinoc': 3.389141191622043e-05,
'arlequinou': 3.389141191622043e-05,
'arma': 3.389141191622043e-05,
'armait': 3.389141191622043e-05,
'armand': 6.778282383244086e-05,
'armant': 3.389141191622043e-05,
'arme': 6.778282383244086e-05,
'armenonvill': 3.389141191622043e-05,
'armer': 3.389141191622043e-05,
'armet': 3.389141191622043e-05,
'armez': 3.389141191622043e-05,
'armid': 6.778282383244086e-05,
'armoir': 0.00010167423574866129,
'armon': 3.389141191622043e-05,
'armur': 3.389141191622043e-05,
'armé': 0.00013556564766488172,
'armen': 3.389141191622043e-05,
'arménien': 3.389141191622043e-05,
'arnolph': 3.389141191622043e-05,
'aronc': 3.389141191622043e-05,
'arpen': 3.389141191622043e-05,
'arpent': 6.778282383244086e-05,
'arquebus': 3.389141191622043e-05,
'arrach': 0.000474479766827086,
'arrachent': 3.389141191622043e-05,
'arrang': 0.000474479766827086,
'arrangent': 3.389141191622043e-05,
'arrangeon': 3.389141191622043e-05,
'arras': 3.389141191622043e-05,
'arrhe': 3.389141191622043e-05,
'arriv': 0.0007456110621568494,
'arrivent': 3.389141191622043e-05,
'arrivon': 3.389141191622043e-05,
'arrier': 3.389141191622043e-05,
'arrog': 0.00020334847149732258,
'arrond': 3.389141191622043e-05,
'arros': 0.00013556564766488172,
'arrérage': 3.389141191622043e-05,
'arrêt': 0.0006778282383244086,
'arrêtent': 3.389141191622043e-05,
'arrêtic': 3.389141191622043e-05,
'arrêton': 3.389141191622043e-05,
'arsenal': 3.389141191622043e-05,
'arsenic': 3.389141191622043e-05,
'arsino': 3.389141191622043e-05,
'art': 6.778282383244086e-05,
'artabaz': 3.389141191622043e-05,
'arthénic': 3.389141191622043e-05,
'articl': 6.778282383244086e-05,
'articul': 6.778282383244086e-05,
'artific': 6.778282383244086e-05,
'artifici': 3.389141191622043e-05,
'artiller': 3.389141191622043e-05,
'artiphil': 3.389141191622043e-05,
'artisan': 6.778282383244086e-05,
'artist': 0.00010167423574866129,
'artois': 3.389141191622043e-05,
'artus': 3.389141191622043e-05,
'arter': 6.778282383244086e-05,
'artémis': 3.389141191622043e-05,
'artémius': 3.389141191622043e-05,
'arténic': 3.389141191622043e-05,
'arçon': 3.389141191622043e-05,
'aren': 3.389141191622043e-05,
'aréopag': 3.389141191622043e-05,
'as': 3.389141191622043e-05,

'ascagn': 3.389141191622043e-05,
'ascend': 3.389141191622043e-05,
'asi': 3.389141191622043e-05,
'asil': 6.778282383244086e-05,
'asmod': 3.389141191622043e-05,
'aspas': 3.389141191622043e-05,
'aspect': 6.778282383244086e-05,
'aspic': 3.389141191622043e-05,
'aspir': 0.000237239883413543,
'aspirent': 3.389141191622043e-05,
'assai': 3.389141191622043e-05,
'assaill': 6.778282383244086e-05,
'assaison': 0.00016945705958110215,
'assaisonnent': 3.389141191622043e-05,
'assassin': 0.00030502270724598384,
'assassinat': 3.389141191622043e-05,
'assaut': 6.778282383244086e-05,
'assembl': 0.0003728055310784247,
'assemblag': 3.389141191622043e-05,
'assemblent': 3.389141191622043e-05,
'asseoir': 3.389141191622043e-05,
'assert': 3.389141191622043e-05,
'asserv': 0.00016945705958110215,
'assesseur': 3.389141191622043e-05,
'assei': 3.389141191622043e-05,
'asseyon': 3.389141191622043e-05,
'assez': 3.389141191622043e-05,
'assidu': 0.00013556564766488172,
'assidus': 3.389141191622043e-05,
'assidû': 3.389141191622043e-05,
'assied': 6.778282383244086e-05,
'assiet': 6.778282383244086e-05,
'assign': 0.00013556564766488172,
'assimil': 3.389141191622043e-05,
'assis': 6.778282383244086e-05,
'assist': 0.000237239883413543,
'assieg': 0.00013556564766488172,
'assiègent': 3.389141191622043e-05,
'associ': 0.00016945705958110215,
'assoc': 0.00010167423574866129,
'assomm': 0.0003728055310784247,
'assomment': 3.389141191622043e-05,
'assommeur': 3.389141191622043e-05,
'assort': 0.00016945705958110215,
'assott': 3.389141191622043e-05,
'assot': 3.389141191622043e-05,
'assoup': 0.00020334847149732258,
'assouv': 0.00010167423574866129,
'assujet': 0.00016945705958110215,
'assur': 0.0006778282383244086,
'assurent': 3.389141191622043e-05,
'assuron': 3.389141191622043e-05,
'assen': 3.389141191622043e-05,
'asteur': 3.389141191622043e-05,
'asthmat': 3.389141191622043e-05,
'asthme': 3.389141191622043e-05,
'astorgu': 3.389141191622043e-05,
'astre': 6.778282383244086e-05,
'astrolab': 3.389141191622043e-05,
'astrolog': 3.389141191622043e-05,
'astrologu': 6.778282383244086e-05,
'astre': 3.389141191622043e-05,
'atal': 3.389141191622043e-05,
'ateli': 3.389141191622043e-05,
'athen': 6.778282383244086e-05,
'athénien': 3.389141191622043e-05,
'atlant': 3.389141191622043e-05,
'atlas': 3.389141191622043e-05,
'atom': 6.778282383244086e-05,

'atour': 3.389141191622043e-05,
'atque': 3.389141191622043e-05,
'atrabilair': 3.389141191622043e-05,
'atroc': 0.00010167423574866129,
'atropos': 3.389141191622043e-05,
'attach': 0.0005083711787433065,
'attachent': 3.389141191622043e-05,
'attaqu': 0.0003728055310784247,
'attaquent': 3.389141191622043e-05,
'attaquon': 3.389141191622043e-05,
'atteindr': 3.389141191622043e-05,
'atteint': 0.00013556564766488172,
'attelag': 3.389141191622043e-05,
'attel': 0.00010167423574866129,
'atten': 3.389141191622043e-05,
'attend': 0.0003728055310784247,
'attendent': 3.389141191622043e-05,
'attendois': 3.389141191622043e-05,
'attendon': 3.389141191622043e-05,
'attendent': 3.389141191622043e-05,
'attendr': 0.0006100454144919677,
'attendron': 3.389141191622043e-05,
'attendu': 6.778282383244086e-05,
'attentat': 6.778282383244086e-05,
'attent': 0.0003728055310784247,
'attest': 6.778282383244086e-05,
'attif': 3.389141191622043e-05,
'attiqu': 6.778282383244086e-05,
'attir': 0.00044058835491086557,
'attirail': 3.389141191622043e-05,
'attirent': 3.389141191622043e-05,
'attitud': 6.778282383244086e-05,
'attract': 6.778282383244086e-05,
'attrait': 6.778282383244086e-05,
'attrap': 0.00044058835491086557,
'attrapent': 3.389141191622043e-05,
'attrapon': 3.389141191622043e-05,
'attrai': 3.389141191622043e-05,
'attribu': 0.00020334847149732258,
'attribuent': 3.389141191622043e-05,
'attribut': 6.778282383244086e-05,
'attrist': 0.00013556564766488172,
'aty': 3.389141191622043e-05,
'au': 3.389141191622043e-05,
'aubad': 3.389141191622043e-05,
'aubain': 3.389141191622043e-05,
'aub': 3.389141191622043e-05,
'auberg': 6.778282383244086e-05,
'aubert': 3.389141191622043e-05,
'aubignac': 3.389141191622043e-05,
'aubin': 3.389141191622043e-05,
'aubouyn': 3.389141191622043e-05,
'aucun': 0.00016945705958110215,
'audac': 3.389141191622043e-05,
'audaci': 6.778282383244086e-05,
'audienc': 6.778282383244086e-05,
'audit': 3.389141191622043e-05,
'auditeur': 6.778282383244086e-05,
'auditoir': 3.389141191622043e-05,
'auditric': 3.389141191622043e-05,
'augment': 0.00040669694299464516,
'augmentent': 3.389141191622043e-05,
'augur': 0.00013556564766488172,
'august': 6.778282383244086e-05,
'augustin': 6.778282383244086e-05,
'ai': 3.389141191622043e-05,
'aujordy': 3.389141191622043e-05,
'aujourn': 3.389141191622043e-05,
'aujourdhui': 3.389141191622043e-05,


```

'aulin': 6.778282383244086e-05,
'aumôn': 6.778282383244086e-05,
'aumôni': 3.389141191622043e-05,
'aun': 6.778282383244086e-05,
'auparav': 3.389141191622043e-05,
'aupres': 3.389141191622043e-05,
'aupr': 3.389141191622043e-05,
'auquel': 3.389141191622043e-05,
'aur': 0.00030502270724598384,
'auret': 3.389141191622043e-05,
'auri': 3.389141191622043e-05,
'auriculaire': 3.389141191622043e-05,
'aurion': 3.389141191622043e-05,
'auroient': 3.389141191622043e-05,
'aurois': 3.389141191622043e-05,
'auroit': 3.389141191622043e-05,
'auron': 3.389141191622043e-05,
'auront': 3.389141191622043e-05,
'auror': 3.389141191622043e-05,
'aurel': 3.389141191622043e-05,
'aurélio': 3.389141191622043e-05,
'ausbourg': 3.389141191622043e-05,
'auspic': 3.389141191622043e-05,
'auss': 3.389141191622043e-05,
'aussitôt': 3.389141191622043e-05,
'auster': 0.00013556564766488172,
'aut': 0.00013556564766488172,
'autel': 6.778282383244086e-05,
...}

```

In [43]:

```

stemmed_vocabulary_tra = [fr.stem(word) for word in vocabulary_tra]
counts = dict()
for word in stemmed_vocabulary_tra:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_tra = counts
for key, value in counts.items():
    tf_dict_tra[key] = value / len(stemmed_vocabulary_tra)
tf_dict_tra

```

Out[43]:

```

{'a': 5.0020008003201284e-05,
'aaron': 5.0020008003201284e-05,
'ab': 5.0020008003201284e-05,
'abaiss': 0.0007002801120448179,
'abaissement': 0.00010004001600640257,
'abaissent': 5.0020008003201284e-05,
'abandon': 0.0010504201680672268,
'abandonnent': 5.0020008003201284e-05,
'abandonnon': 5.0020008003201284e-05,
'abat': 5.0020008003201284e-05,
'abatt': 0.00015006002400960383,
'abattr': 5.0020008003201284e-05,
'abattu': 0.00010004001600640257,
'abattus': 5.0020008003201284e-05,
'abbé': 5.0020008003201284e-05,
'abdened': 5.0020008003201284e-05,
'abdiqu': 0.00015006002400960383,
'abdit': 5.0020008003201284e-05,
'abencerrag': 5.0020008003201284e-05,
'abendax': 5.0020008003201284e-05,
'abhorrr': 0.00035014005602240897,
'abincerrag': 5.0020008003201284e-05,
'abject': 0.00010004001600640257,
'abjur': 0.0002501000400160064,
'abner': 5.0020008003201284e-05,

```

'abois': 5.0020008003201284e-05,
'abol': 0.00020008003201280514,
'abomin': 0.00010004001600640257,
'abond': 0.00020008003201280514,
'abord': 0.0004501800720288115,
'abra': 5.0020008003201284e-05,
'abraham': 5.0020008003201284e-05,
'abreuv': 0.0002501000400160064,
'abri': 5.0020008003201284e-05,
'abreg': 0.0002501000400160064,
'abrègent': 5.0020008003201284e-05,
'absalon': 5.0020008003201284e-05,
'absenc': 5.0020008003201284e-05,
'absent': 0.00010004001600640257,
'absolu': 0.00020008003201280514,
'absolus': 5.0020008003201284e-05,
'absorb': 5.0020008003201284e-05,
'absoudr': 0.00010004001600640257,
'absous': 5.0020008003201284e-05,
'absout': 5.0020008003201284e-05,
'abstinent': 5.0020008003201284e-05,
'absurd': 0.00020008003201280514,
'abus': 0.000900360144057623,
'abusent': 5.0020008003201284e-05,
'abuson': 5.0020008003201284e-05,
'abym': 5.0020008003201284e-05,
'abîm': 0.00030012004801920766,
'ac': 5.0020008003201284e-05,
'académicien': 5.0020008003201284e-05,
'academ': 5.0020008003201284e-05,
'acast': 5.0020008003201284e-05,
'acbitophel': 5.0020008003201284e-05,
'accabl': 0.0008003201280512205,
'accablent': 5.0020008003201284e-05,
'accablon': 5.0020008003201284e-05,
'accent': 0.00010004001600640257,
'accept': 0.0007002801120448179,
'accepton': 5.0020008003201284e-05,
'accessibl': 5.0020008003201284e-05,
'accessoir': 5.0020008003201284e-05,
'accident': 0.00010004001600640257,
'acclam': 5.0020008003201284e-05,
'accommod': 0.00015006002400960383,
'accompagn': 0.0007002801120448179,
'accompagnent': 5.0020008003201284e-05,
'accompl': 0.0007002801120448179,
'accord': 0.0008003201280512205,
'accordent': 5.0020008003201284e-05,
'accordon': 5.0020008003201284e-05,
'accouch': 5.0020008003201284e-05,
'accourc': 0.00015006002400960383,
'accourent': 5.0020008003201284e-05,
'accour': 0.00015006002400960383,
'accourt': 5.0020008003201284e-05,
'accouru': 0.00010004001600640257,
'accourus': 5.0020008003201284e-05,
'accourut': 5.0020008003201284e-05,
'accoutum': 0.00035014005602240897,
'accroir': 5.0020008003201284e-05,
'accroissement': 5.0020008003201284e-05,
'accroissent': 5.0020008003201284e-05,
'accroît': 5.0020008003201284e-05,
'accroît': 5.0020008003201284e-05,
'accroître': 5.0020008003201284e-05,
'accru': 0.00010004001600640257,
'accrus': 5.0020008003201284e-05,
'accrut': 5.0020008003201284e-05,
'accueil': 0.00030012004801920766,
'accumul': 0.00010004001600640257,

'accus': 0.0011504601840736295,
'accusent': 5.0020008003201284e-05,
'accuson': 5.0020008003201284e-05,
'acces': 5.0020008003201284e-05,
'achab': 5.0020008003201284e-05,
'acharn': 0.00020008003201280514,
'achas': 5.0020008003201284e-05,
'achat': 5.0020008003201284e-05,
'achemin': 5.0020008003201284e-05,
'achet': 0.0004501800720288115,
'achev': 0.0007503001200480192,
'achevon': 5.0020008003201284e-05,
'achill': 5.0020008003201284e-05,
'achillea': 5.0020008003201284e-05,
'achior': 5.0020008003201284e-05,
'achis': 5.0020008003201284e-05,
'achitophel': 5.0020008003201284e-05,
'achmat': 5.0020008003201284e-05,
'achmet': 5.0020008003201284e-05,
'achomat': 5.0020008003201284e-05,
'achètent': 5.0020008003201284e-05,
'achèvent': 5.0020008003201284e-05,
'achéron': 5.0020008003201284e-05,
'aci': 5.0020008003201284e-05,
'acmat': 5.0020008003201284e-05,
'acomat': 5.0020008003201284e-05,
'acquier': 0.00010004001600640257,
'acquiert': 5.0020008003201284e-05,
'acquirent': 5.0020008003201284e-05,
'acquis': 0.00015006002400960383,
'acquit': 5.0020008003201284e-05,
'acquitt': 0.0002501000400160064,
'acquitton': 5.0020008003201284e-05,
'acqui': 5.0020008003201284e-05,
'acquer': 0.00020008003201280514,
'acte': 0.00010004001600640257,
'acteur': 0.00010004001600640257,
'actif': 5.0020008003201284e-05,
'action': 0.00010004001600640257,
'actium': 5.0020008003201284e-05,
'activ': 0.00010004001600640257,
'actor': 5.0020008003201284e-05,
'actric': 0.00010004001600640257,
'actuel': 5.0020008003201284e-05,
'ad': 5.0020008003201284e-05,
'adam': 5.0020008003201284e-05,
'addison': 5.0020008003201284e-05,
'adhuc': 5.0020008003201284e-05,
'adieu': 5.0020008003201284e-05,
'adieux': 5.0020008003201284e-05,
'admet': 0.00010004001600640257,
'admettr': 5.0020008003201284e-05,
'admir': 0.0009503801520608243,
'admirent': 5.0020008003201284e-05,
'admiron': 5.0020008003201284e-05,
'admis': 5.0020008003201284e-05,
'adoni': 5.0020008003201284e-05,
'adopt': 0.0005002000800320128,
'ador': 0.0010004001600640256,
'adorent': 5.0020008003201284e-05,
'adoron': 5.0020008003201284e-05,
'adouc': 0.0006502601040416166,
'adrast': 5.0020008003201284e-05,
'adress': 0.0005502200880352141,
'adressent': 5.0020008003201284e-05,
'adrian': 5.0020008003201284e-05,
'adrien': 0.00010004001600640257,
'adroit': 0.0002501000400160064,
'adul': 0.00010004001600640257,

'adulter': 0.00010004001600640257,
'adven': 5.0020008003201284e-05,
'advenu': 5.0020008003201284e-05,
'advenus': 5.0020008003201284e-05,
'adversair': 0.00010004001600640257,
'advers': 0.00015006002400960383,
'advient': 5.0020008003201284e-05,
'adélaïd': 5.0020008003201284e-05,
'aeacid': 5.0020008003201284e-05,
'aegl': 5.0020008003201284e-05,
'aenarus': 5.0020008003201284e-05,
'aequor': 5.0020008003201284e-05,
'aerop': 5.0020008003201284e-05,
'affabl': 5.0020008003201284e-05,
'affaibl': 0.00040016006402561027,
'affair': 0.00010004001600640257,
'affaiss': 5.0020008003201284e-05,
'affam': 0.00015006002400960383,
'affect': 0.0007002801120448179,
'affectent': 5.0020008003201284e-05,
'affection': 5.0020008003201284e-05,
'afferme': 0.0005002000800320128,
'affich': 5.0020008003201284e-05,
'affid': 5.0020008003201284e-05,
'affin': 5.0020008003201284e-05,
'affirm': 0.00010004001600640257,
'afflict': 0.00010004001600640257,
'afflig': 0.0005002000800320128,
'affligent': 5.0020008003201284e-05,
'affligeon': 5.0020008003201284e-05,
'affranch': 0.00040016006402561027,
'affreux': 0.00015006002400960383,
'affront': 0.00030012004801920766,
'afin': 5.0020008003201284e-05,
'africain': 0.00015006002400960383,
'afriqu': 5.0020008003201284e-05,
'aga': 5.0020008003201284e-05,
'agag': 5.0020008003201284e-05,
'agamemnon': 5.0020008003201284e-05,
'agathocl': 5.0020008003201284e-05,
'agathon': 5.0020008003201284e-05,
'agent': 0.00010004001600640257,
'aggrav': 0.00010004001600640257,
'agi': 5.0020008003201284e-05,
'agil': 5.0020008003201284e-05,
'agir': 0.00020008003201280514,
'agiront': 5.0020008003201284e-05,
'agis': 5.0020008003201284e-05,
'agiss': 0.00020008003201280514,
'agissent': 5.0020008003201284e-05,
'agisson': 5.0020008003201284e-05,
'agit': 0.0006002400960384153,
'agitatus': 5.0020008003201284e-05,
'agitent': 5.0020008003201284e-05,
'agla': 5.0020008003201284e-05,
'agneau': 0.00010004001600640257,
'agnonid': 5.0020008003201284e-05,
'agnes': 5.0020008003201284e-05,
'agrand': 0.00015006002400960383,
'agrip': 5.0020008003201284e-05,
'agripin': 5.0020008003201284e-05,
'agripp': 0.00010004001600640257,
'agrippin': 5.0020008003201284e-05,
'agréabl': 0.00015006002400960383,
'agré': 0.00010004001600640257,
'agrément': 0.00010004001600640257,
'aguerr': 5.0020008003201284e-05,
'agénor': 5.0020008003201284e-05,
'ah': 5.0020008003201284e-05,

'ai': 5.0020008003201284e-05,
'aid': 0.0005502200880352141,
'aidon': 5.0020008003201284e-05,
'aie': 5.0020008003201284e-05,
'aient': 5.0020008003201284e-05,
'aigl': 0.00010004001600640257,
'aigr': 0.0006502601040416166,
'aigreur': 5.0020008003201284e-05,
'aiguillon': 5.0020008003201284e-05,
'aiguiss': 5.0020008003201284e-05,
'ail': 0.00010004001600640257,
'aill': 5.0020008003201284e-05,
'aillent': 5.0020008003201284e-05,
'ailleur': 5.0020008003201284e-05,
'aim': 0.001250500200080032,
'aimabl': 0.00010004001600640257,
'aiment': 5.0020008003201284e-05,
'aimion': 5.0020008003201284e-05,
'aimon': 5.0020008003201284e-05,
'ain': 0.00010004001600640257,
'ains': 5.0020008003201284e-05,
'air': 0.00010004001600640257,
'airain': 5.0020008003201284e-05,
'ais': 0.00030012004801920766,
'aisanc': 5.0020008003201284e-05,
'ait': 5.0020008003201284e-05,
'ajax': 5.0020008003201284e-05,
'ajout': 0.0006002400960384153,
'ajouton': 5.0020008003201284e-05,
'ajust': 0.00010004001600640257,
'al': 5.0020008003201284e-05,
'alam': 5.0020008003201284e-05,
'alaric': 5.0020008003201284e-05,
'alarm': 0.0004501800720288115,
'alarment': 5.0020008003201284e-05,
'albert': 5.0020008003201284e-05,
'albin': 0.00010004001600640257,
'albion': 5.0020008003201284e-05,
'albis': 5.0020008003201284e-05,
'alcest': 5.0020008003201284e-05,
'alcet': 5.0020008003201284e-05,
'alcibiad': 5.0020008003201284e-05,
'alcid': 5.0020008003201284e-05,
'alcim': 5.0020008003201284e-05,
'alcimed': 5.0020008003201284e-05,
'alcimédon': 5.0020008003201284e-05,
'alcinous': 5.0020008003201284e-05,
'alcion': 5.0020008003201284e-05,
'alcir': 5.0020008003201284e-05,
'alcmen': 5.0020008003201284e-05,
'alcméon': 5.0020008003201284e-05,
'alcomir': 5.0020008003201284e-05,
'alcoran': 5.0020008003201284e-05,
'aldamon': 5.0020008003201284e-05,
'alecton': 5.0020008003201284e-05,
'alenc': 5.0020008003201284e-05,
'alent': 5.0020008003201284e-05,
'alentour': 5.0020008003201284e-05,
'alençon': 5.0020008003201284e-05,
'alexandr': 0.00010004001600640257,
'alexandrin': 5.0020008003201284e-05,
'alex': 0.00010004001600640257,
'ali': 0.00010004001600640257,
'alii': 5.0020008003201284e-05,
'aliment': 0.00010004001600640257,
'alix': 5.0020008003201284e-05,
'alla': 5.0020008003201284e-05,
'allai': 5.0020008003201284e-05,
'allaient': 5.0020008003201284e-05,

'allais': 5.0020008003201284e-05,
'allait': 0.00015006002400960383,
'allant': 5.0020008003201284e-05,
'allarm': 0.00010004001600640257,
'allemagn': 5.0020008003201284e-05,
'allemand': 0.00010004001600640257,
'aller': 5.0020008003201284e-05,
'allez': 5.0020008003201284e-05,
'alli': 0.00040016006402561027,
'allianc': 5.0020008003201284e-05,
'allion': 5.0020008003201284e-05,
'allobrog': 5.0020008003201284e-05,
'allong': 5.0020008003201284e-05,
'allon': 5.0020008003201284e-05,
'allum': 0.0006002400960384153,
'allument': 5.0020008003201284e-05,
'allât': 5.0020008003201284e-05,
'allech': 5.0020008003201284e-05,
'alleg': 0.00020008003201280514,
'allé': 0.00015006002400960383,
'allége': 5.0020008003201284e-05,
'allégress': 5.0020008003201284e-05,
'allégu': 0.0002501000400160064,
'alor': 5.0020008003201284e-05,
'alpe': 5.0020008003201284e-05,
'alphons': 5.0020008003201284e-05,
'alphé': 5.0020008003201284e-05,
'altdorff': 5.0020008003201284e-05,
'altess': 5.0020008003201284e-05,
'alti': 5.0020008003201284e-05,
'altier': 0.00015006002400960383,
'altis': 5.0020008003201284e-05,
'alter': 0.00035014005602240897,
'altèrent': 5.0020008003201284e-05,
'alvares': 5.0020008003201284e-05,
'alzaid': 5.0020008003201284e-05,
'alzir': 5.0020008003201284e-05,
'alzond': 5.0020008003201284e-05,
'amad': 5.0020008003201284e-05,
'amalec': 5.0020008003201284e-05,
'amalécit': 5.0020008003201284e-05,
'aman': 0.00010004001600640257,
'amant': 0.00015006002400960383,
'amas': 0.00020008003201280514,
'amass': 0.00020008003201280514,
'amateur': 0.00010004001600640257,
'amazon': 0.00010004001600640257,
'ambassad': 5.0020008003201284e-05,
'ambassadeur': 0.00010004001600640257,
'ambiguït': 5.0020008003201284e-05,
'ambiti': 0.00010004001600640257,
'ambit': 5.0020008003201284e-05,
'ambre': 5.0020008003201284e-05,
'ambrois': 5.0020008003201284e-05,
'ame': 5.0020008003201284e-05,
'amen': 0.0006002400960384153,
'amend': 5.0020008003201284e-05,
'amer': 0.00030012004801920766,
'amertum': 5.0020008003201284e-05,
'amestr': 5.0020008003201284e-05,
'ami': 0.00015006002400960383,
'amilcar': 5.0020008003201284e-05,
'aminus': 5.0020008003201284e-05,
'amint': 0.00010004001600640257,
'amis': 5.0020008003201284e-05,
'amiti': 0.00010004001600640257,
'ammon': 5.0020008003201284e-05,
'amnon': 5.0020008003201284e-05,
'amoindr': 5.0020008003201284e-05,

'amoll': 0.00030012004801920766,
'amorc': 0.00010004001600640257,
'amor': 5.0020008003201284e-05,
'amort': 0.00010004001600640257,
'amour': 0.0002501000400160064,
'amphict': 5.0020008003201284e-05,
'amphrys': 5.0020008003201284e-05,
'ample': 0.00010004001600640257,
'amplement': 5.0020008003201284e-05,
'amsterdam': 5.0020008003201284e-05,
'amurat': 5.0020008003201284e-05,
'amus': 0.00035014005602240897,
'amusent': 5.0020008003201284e-05,
'amynt': 5.0020008003201284e-05,
'amènent': 5.0020008003201284e-05,
'amé': 0.00010004001600640257,
'amel': 5.0020008003201284e-05,
'aménaïd': 5.0020008003201284e-05,
'aménoph': 5.0020008003201284e-05,
'américain': 5.0020008003201284e-05,
'an': 0.00010004001600640257,
'analys': 0.00010004001600640257,
'anathem': 5.0020008003201284e-05,
'anaxandr': 5.0020008003201284e-05,
'ancien': 0.0002501000400160064,
'ancree': 5.0020008003201284e-05,
'ancêtr': 5.0020008003201284e-05,
'andrinopl': 5.0020008003201284e-05,
'androclid': 5.0020008003201284e-05,
'andromach': 5.0020008003201284e-05,
'andromaqu': 5.0020008003201284e-05,
'andromed': 5.0020008003201284e-05,
'andronic': 5.0020008003201284e-05,
'andré': 5.0020008003201284e-05,
'anecdote': 0.00010004001600640257,
'ange': 0.00010004001600640257,
'anglais': 0.00010004001600640257,
'angle': 5.0020008003201284e-05,
'angleterr': 5.0020008003201284e-05,
'angloman': 5.0020008003201284e-05,
'anim': 0.0005502200880352141,
'animal': 0.00010004001600640257,
'animent': 5.0020008003201284e-05,
'animos': 0.00010004001600640257,
'anitus': 5.0020008003201284e-05,
'annal': 5.0020008003201284e-05,
'anneau': 5.0020008003201284e-05,
'annibal': 5.0020008003201284e-05,
'annonc': 0.0005502200880352141,
'annoncent': 5.0020008003201284e-05,
'annonçon': 5.0020008003201284e-05,
'année': 0.00010004001600640257,
'anobl': 5.0020008003201284e-05,
'anonym': 5.0020008003201284e-05,
'ante': 5.0020008003201284e-05,
'anthenor': 5.0020008003201284e-05,
'anthoin': 5.0020008003201284e-05,
'anthropophag': 5.0020008003201284e-05,
'anthym': 5.0020008003201284e-05,
'anticip': 5.0020008003201284e-05,
'antidot': 5.0020008003201284e-05,
'antigon': 5.0020008003201284e-05,
'antigonus': 5.0020008003201284e-05,
'antinoü': 5.0020008003201284e-05,
'antiochus': 5.0020008003201284e-05,
'antiop': 5.0020008003201284e-05,
'antipat': 5.0020008003201284e-05,
'antipatr': 5.0020008003201284e-05,
'antiqu': 0.00015006002400960383,

'antoin': 5.0020008003201284e-05,
'antoinet': 5.0020008003201284e-05,
'antoniât': 5.0020008003201284e-05,
'antre': 0.00010004001600640257,
'anténor': 5.0020008003201284e-05,
'anéant': 0.0002501000400160064,
'août': 5.0020008003201284e-05,
'apais': 0.00040016006402561027,
'apaisent': 5.0020008003201284e-05,
'apaison': 5.0020008003201284e-05,
'apam': 5.0020008003201284e-05,
'apanag': 5.0020008003201284e-05,
'apennin': 5.0020008003201284e-05,
'apercev': 0.00010004001600640257,
'apercevoir': 5.0020008003201284e-05,
'aperçois': 5.0020008003201284e-05,
'aperçoit': 5.0020008003201284e-05,
'aperçoivent': 5.0020008003201284e-05,
'aperçu': 0.00010004001600640257,
'aperçus': 5.0020008003201284e-05,
'aperçut': 5.0020008003201284e-05,
'aperçût': 5.0020008003201284e-05,
'aplan': 0.00020008003201280514,
'apollon': 5.0020008003201284e-05,
'apollonius': 5.0020008003201284e-05,
'apolog': 5.0020008003201284e-05,
'apoplex': 5.0020008003201284e-05,
'apostat': 5.0020008003201284e-05,
'apost': 5.0020008003201284e-05,
'appaïs': 5.0020008003201284e-05,
'appanag': 5.0020008003201284e-05,
'apparaît': 5.0020008003201284e-05,
'apparaître': 5.0020008003201284e-05,
'appareil': 0.00010004001600640257,
'apparent': 0.00030012004801920766,
'apparit': 0.00010004001600640257,
'appart': 5.0020008003201284e-05,
'apparten': 0.00015006002400960383,
'appartiendr': 5.0020008003201284e-05,
'appartien': 0.00010004001600640257,
'appartiennent': 5.0020008003201284e-05,
'appartient': 5.0020008003201284e-05,
'apparu': 5.0020008003201284e-05,
'appas': 5.0020008003201284e-05,
'appel': 0.0008003201280512205,
'appellent': 5.0020008003201284e-05,
'appelon': 5.0020008003201284e-05,
'appesant': 0.00030012004801920766,
'appius': 5.0020008003201284e-05,
'applaud': 0.0006502601040416166,
'applic': 5.0020008003201284e-05,
'appliqu': 0.00030012004801920766,
'appliquent': 5.0020008003201284e-05,
'apport': 0.0007503001200480192,
'apportent': 5.0020008003201284e-05,
'appren': 0.00030012004801920766,
'apprend': 0.00010004001600640257,
'apprendr': 0.00040016006402561027,
'apprendron': 5.0020008003201284e-05,
'apprendront': 5.0020008003201284e-05,
'apprennent': 5.0020008003201284e-05,
'apprenon': 5.0020008003201284e-05,
'apprent': 5.0020008003201284e-05,
'apprentissag': 5.0020008003201284e-05,
'apprirent': 5.0020008003201284e-05,
'appris': 0.00010004001600640257,
'apprit': 5.0020008003201284e-05,
'apprivois': 0.00010004001600640257,
'approb': 0.00015006002400960383,

'approch': 0.0006502601040416166,
'approchent': 5.0020008003201284e-05,
'approchon': 5.0020008003201284e-05,
'approfond': 0.00010004001600640257,
'approuv': 0.0008003201280512205,
'approuvon': 5.0020008003201284e-05,
'appréci': 5.0020008003201284e-05,
'appréhend': 0.00035014005602240897,
'appréhendon': 5.0020008003201284e-05,
'appréhens': 0.00010004001600640257,
'apprêt': 0.00040016006402561027,
'apprêtent': 5.0020008003201284e-05,
'apprît': 0.00010004001600640257,
'appui': 0.0006502601040416166,
'appuient': 5.0020008003201284e-05,
'appuis': 5.0020008003201284e-05,
'appuy': 5.0020008003201284e-05,
'appât': 0.00010004001600640257,
'appet': 5.0020008003201284e-05,
'apries': 5.0020008003201284e-05,
'apres': 5.0020008003201284e-05,
'apré': 5.0020008003201284e-05,
'apôtr': 5.0020008003201284e-05,
'aquilien': 5.0020008003201284e-05,
'aquillin': 5.0020008003201284e-05,
'aquilon': 5.0020008003201284e-05,
'aquitain': 5.0020008003201284e-05,
'arab': 0.00015006002400960383,
'aram': 5.0020008003201284e-05,
'aramint': 5.0020008003201284e-05,
'aras': 5.0020008003201284e-05,
'arasp': 5.0020008003201284e-05,
'araxid': 5.0020008003201284e-05,
'arax': 5.0020008003201284e-05,
'arbac': 5.0020008003201284e-05,
'arbaces': 5.0020008003201284e-05,
'arbant': 5.0020008003201284e-05,
'arbas': 5.0020008003201284e-05,
'arbat': 5.0020008003201284e-05,
'arbazan': 5.0020008003201284e-05,
'arbitrair': 0.00010004001600640257,
'arbitr': 0.00010004001600640257,
'arbor': 5.0020008003201284e-05,
'arbre': 0.00010004001600640257,
'arbrisseau': 5.0020008003201284e-05,
'arc': 0.00010004001600640257,
'arcagamb': 5.0020008003201284e-05,
'arcas': 5.0020008003201284e-05,
'archang': 5.0020008003201284e-05,
'arche': 5.0020008003201284e-05,
'archipel': 5.0020008003201284e-05,
'architectur': 5.0020008003201284e-05,
'archiv': 5.0020008003201284e-05,
'archont': 5.0020008003201284e-05,
'archélaü': 5.0020008003201284e-05,
'arde': 5.0020008003201284e-05,
'ardemment': 5.0020008003201284e-05,
'ardent': 0.00020008003201280514,
'ardeur': 0.00010004001600640257,
'arethus': 5.0020008003201284e-05,
'argent': 5.0020008003201284e-05,
'argental': 5.0020008003201284e-05,
'argid': 5.0020008003201284e-05,
'argi': 5.0020008003201284e-05,
'argien': 0.00010004001600640257,
'argil': 5.0020008003201284e-05,
'argin': 5.0020008003201284e-05,
'argir': 5.0020008003201284e-05,
'argos': 5.0020008003201284e-05,

'argument': 0.00010004001600640257,
'argus': 5.0020008003201284e-05,
'arian': 0.00010004001600640257,
'ariarath': 5.0020008003201284e-05,
'aric': 5.0020008003201284e-05,
'arid': 0.00010004001600640257,
'ari': 5.0020008003201284e-05,
'ariet': 5.0020008003201284e-05,
'ariov': 5.0020008003201284e-05,
'arib': 5.0020008003201284e-05,
'aristid': 5.0020008003201284e-05,
'aristobul': 5.0020008003201284e-05,
'aristocrat': 5.0020008003201284e-05,
'ariston': 5.0020008003201284e-05,
'aristot': 5.0020008003201284e-05,
'arma': 5.0020008003201284e-05,
'armai': 5.0020008003201284e-05,
'armaient': 5.0020008003201284e-05,
'armait': 5.0020008003201284e-05,
'armant': 5.0020008003201284e-05,
'arme': 0.00010004001600640257,
'armenonvill': 5.0020008003201284e-05,
'arment': 5.0020008003201284e-05,
'armer': 0.00020008003201280514,
'armeront': 5.0020008003201284e-05,
'armet': 5.0020008003201284e-05,
'armez': 5.0020008003201284e-05,
'armid': 5.0020008003201284e-05,
'armi': 5.0020008003201284e-05,
'arminius': 5.0020008003201284e-05,
'armon': 5.0020008003201284e-05,
'armur': 5.0020008003201284e-05,
'armât': 5.0020008003201284e-05,
'armèrent': 5.0020008003201284e-05,
'armé': 0.00020008003201280514,
'armen': 5.0020008003201284e-05,
'arnauld': 5.0020008003201284e-05,
'aron': 0.00010004001600640257,
'arondel': 5.0020008003201284e-05,
'arpinum': 5.0020008003201284e-05,
'arrach': 0.0010504201680672268,
'arrachent': 5.0020008003201284e-05,
'arrachon': 5.0020008003201284e-05,
'arrang': 0.00010004001600640257,
'arriv': 0.0006502601040416166,
'arrivent': 5.0020008003201284e-05,
'arrivon': 5.0020008003201284e-05,
'arrier': 5.0020008003201284e-05,
'arrog': 0.0002501000400160064,
'arros': 0.0004501800720288115,
'arrosent': 5.0020008003201284e-05,
'arrêt': 0.0010004001600640256,
'arrêtent': 5.0020008003201284e-05,
'arrêton': 5.0020008003201284e-05,
'arsac': 5.0020008003201284e-05,
'arsam': 5.0020008003201284e-05,
'arsenal': 5.0020008003201284e-05,
'arsino': 5.0020008003201284e-05,
'arsen': 5.0020008003201284e-05,
'art': 0.00010004001600640257,
'artaban': 5.0020008003201284e-05,
'artabaz': 5.0020008003201284e-05,
'artamen': 5.0020008003201284e-05,
'artanasd': 5.0020008003201284e-05,
'artaxat': 5.0020008003201284e-05,
'artaxerc': 5.0020008003201284e-05,
'artaxerx': 5.0020008003201284e-05,
'artazir': 5.0020008003201284e-05,
'arthémis': 5.0020008003201284e-05,

'arthenic': 5.0020008003201284e-05,
'articl': 0.00010004001600640257,
'artific': 0.00010004001600640257,
'artifici': 0.00010004001600640257,
'artisan': 0.00010004001600640257,
'artist': 5.0020008003201284e-05,
'arter': 0.00010004001600640257,
'artémis': 5.0020008003201284e-05,
'arzac': 5.0020008003201284e-05,
'arzam': 5.0020008003201284e-05,
'arzan': 5.0020008003201284e-05,
'arzémon': 5.0020008003201284e-05,
'aren': 5.0020008003201284e-05,
'aréthus': 5.0020008003201284e-05,
'as': 5.0020008003201284e-05,
'asaph': 5.0020008003201284e-05,
'ascalaph': 5.0020008003201284e-05,
'ascalon': 5.0020008003201284e-05,
'ascend': 5.0020008003201284e-05,
'ascendimus': 5.0020008003201284e-05,
'asdrubal': 5.0020008003201284e-05,
'asiat': 5.0020008003201284e-05,
'asi': 5.0020008003201284e-05,
'asil': 0.00010004001600640257,
'asmon': 5.0020008003201284e-05,
'asmonéen': 0.00010004001600640257,
'aspar': 5.0020008003201284e-05,
'aspas': 5.0020008003201284e-05,
'aspect': 5.0020008003201284e-05,
'asper': 5.0020008003201284e-05,
'aspic': 5.0020008003201284e-05,
'aspir': 0.00030012004801920766,
'aspirent': 5.0020008003201284e-05,
'aspiiron': 5.0020008003201284e-05,
'assaill': 0.00035014005602240897,
'assaillent': 5.0020008003201284e-05,
'assaison': 5.0020008003201284e-05,
'assassin': 0.00040016006402561027,
'assassinat': 0.00010004001600640257,
'assaut': 0.00010004001600640257,
'assembl': 0.0005502200880352141,
'assemblag': 5.0020008003201284e-05,
'assemblent': 5.0020008003201284e-05,
'asseoir': 5.0020008003201284e-05,
'asser': 5.0020008003201284e-05,
'asserv': 0.00040016006402561027,
'assez': 5.0020008003201284e-05,
'assidu': 0.00015006002400960383,
'assidus': 5.0020008003201284e-05,
'assied': 5.0020008003201284e-05,
'assiet': 5.0020008003201284e-05,
'assign': 0.00010004001600640257,
'assis': 0.00010004001600640257,
'assist': 0.00040016006402561027,
'assieg': 0.00030012004801920766,
'assiègent': 5.0020008003201284e-05,
'associ': 0.00015006002400960383,
'assoc': 0.00015006002400960383,
'assomm': 5.0020008003201284e-05,
'assort': 0.0002501000400160064,
'assoup': 0.00020008003201280514,
'assouv': 0.00030012004801920766,
'assujet': 0.00030012004801920766,
'assur': 0.0010004001600640256,
'assurent': 5.0020008003201284e-05,
'assuron': 5.0020008003201284e-05,
'assuérus': 5.0020008003201284e-05,
'assyri': 5.0020008003201284e-05,
'assyrien': 5.0020008003201284e-05,

'assel': 5.0020008003201284e-05,
'ast': 5.0020008003201284e-05,
'astarb': 5.0020008003201284e-05,
'astre': 0.00010004001600640257,
'astreint': 5.0020008003201284e-05,
'astré': 5.0020008003201284e-05,
'astyanax': 5.0020008003201284e-05,
'aster': 5.0020008003201284e-05,
'atalaric': 5.0020008003201284e-05,
'atalid': 5.0020008003201284e-05,
'atamir': 5.0020008003201284e-05,
'ataulph': 5.0020008003201284e-05,
'athal': 5.0020008003201284e-05,
'athen': 5.0020008003201284e-05,
'athé': 0.00010004001600640257,
'athénaï': 5.0020008003201284e-05,
'athénien': 0.00010004001600640257,
'atlas': 5.0020008003201284e-05,
'atom': 5.0020008003201284e-05,
'atrid': 0.00010004001600640257,
'atroc': 0.00020008003201280514,
'atré': 5.0020008003201284e-05,
'attach': 0.000900360144057623,
'attachent': 5.0020008003201284e-05,
'attal': 5.0020008003201284e-05,
'attaqu': 0.0008003201280512205,
'attaquent': 5.0020008003201284e-05,
'attaquon': 5.0020008003201284e-05,
'atteignent': 5.0020008003201284e-05,
'atteign': 5.0020008003201284e-05,
'atteindr': 5.0020008003201284e-05,
'attein': 5.0020008003201284e-05,
'atteint': 0.00020008003201280514,
'attend': 0.0005502200880352141,
'attendent': 5.0020008003201284e-05,
'attendois': 5.0020008003201284e-05,
'attendon': 5.0020008003201284e-05,
'attendr': 0.0008503401360544217,
'attendron': 5.0020008003201284e-05,
'attendront': 5.0020008003201284e-05,
'attendu': 0.00010004001600640257,
'attendus': 5.0020008003201284e-05,
'attent': 0.0005002000800320128,
'attentat': 0.00010004001600640257,
'atterr': 5.0020008003201284e-05,
'attest': 0.00040016006402561027,
'attestent': 5.0020008003201284e-05,
'attil': 0.00010004001600640257,
'attiqu': 5.0020008003201284e-05,
'attir': 0.0007002801120448179,
'attirent': 5.0020008003201284e-05,
'attis': 0.00010004001600640257,
'attitud': 5.0020008003201284e-05,
'attouch': 5.0020008003201284e-05,
'attrait': 0.00010004001600640257,
'attribu': 0.00020008003201280514,
'attribut': 5.0020008003201284e-05,
'attrist': 0.00010004001600640257,
'aty': 5.0020008003201284e-05,
'au': 5.0020008003201284e-05,
'aub': 5.0020008003201284e-05,
'aucun': 0.00020008003201280514,
'audac': 5.0020008003201284e-05,
'audaci': 0.00010004001600640257,
'audienc': 5.0020008003201284e-05,
'audit': 5.0020008003201284e-05,
'auecqu': 5.0020008003201284e-05,
'augment': 0.0005502200880352141,
'augmentent': 5.0020008003201284e-05,

'augmenton': 5.0020008003201284e-05,
'augur': 0.00015006002400960383,
'august': 0.00010004001600640257,
'augustin': 5.0020008003201284e-05,
'aujourd': 5.0020008003201284e-05,
'aujourd'hui': 5.0020008003201284e-05,
'aulerqu': 5.0020008003201284e-05,
'aulid': 5.0020008003201284e-05,
'aulu': 5.0020008003201284e-05,
'aumal': 5.0020008003201284e-05,
'auparav': 5.0020008003201284e-05,
'aupres': 5.0020008003201284e-05,
'aupr': 5.0020008003201284e-05,
'auquel': 5.0020008003201284e-05,
'aur': 0.00040016006402561027,
'aurilcar': 5.0020008003201284e-05,
'aurion': 5.0020008003201284e-05,
'aurois': 5.0020008003201284e-05,
'auroit': 5.0020008003201284e-05,
'auron': 5.0020008003201284e-05,
'auront': 5.0020008003201284e-05,
'auron': 5.0020008003201284e-05,
'aurel': 5.0020008003201284e-05,
'aurélien': 5.0020008003201284e-05,
'auspic': 0.00010004001600640257,
'auss': 5.0020008003201284e-05,
'aussitôt': 5.0020008003201284e-05,
'auster': 0.00015006002400960383,
'aut': 5.0020008003201284e-05,
'autel': 0.00010004001600640257,
'auteur': 0.00010004001600640257,
'authent': 0.00010004001600640257,
'autoris': 0.0004501800720288115,
'autorisent': 5.0020008003201284e-05,
'autor': 0.00010004001600640257,
'autour': 5.0020008003201284e-05,
'autr': 0.00015006002400960383,
'autrefois': 5.0020008003201284e-05,
'autrich': 5.0020008003201284e-05,
'autrui': 5.0020008003201284e-05,
'aux': 5.0020008003201284e-05,
'auxquel': 0.00010004001600640257,
'avaient': 5.0020008003201284e-05,
'avais': 5.0020008003201284e-05,
'avait': 5.0020008003201284e-05,
'aval': 5.0020008003201284e-05,
'avanc': 0.0006502601040416166,
'avancent': 5.0020008003201284e-05,
'avant': 5.0020008003201284e-05,
'avantag': 0.00030012004801920766,
'avançon': 5.0020008003201284e-05,
'avar': 0.00010004001600640257,
'avaric': 5.0020008003201284e-05,
'avec': 5.0020008003201284e-05,
'avecqu': 0.00010004001600640257,
'aven': 0.00010004001600640257,
'aventin': 5.0020008003201284e-05,
'aventur': 0.00010004001600640257,
'avenu': 5.0020008003201284e-05,
'avers': 5.0020008003201284e-05,
'avert': 0.0005002000800320128,
'aveu': 5.0020008003201284e-05,
'aveugl': 0.0005502200880352141,
'aveux': 5.0020008003201284e-05,
'avev': 5.0020008003201284e-05,
'avez': 5.0020008003201284e-05,
'avid': 0.00020008003201280514,
'avi': 5.0020008003201284e-05,
'avil': 0.0004501800720288115,

'avion': 5.0020008003201284e-05,
'avis': 0.00015006002400960383,
'avocat': 0.00010004001600640257,
'avoient': 5.0020008003201284e-05,
'avoir': 5.0020008003201284e-05,
'avoit': 5.0020008003201284e-05,
'avon': 5.0020008003201284e-05,
'avort': 0.00015006002400960383,
'avou': 0.0006002400960384153,
'avouent': 5.0020008003201284e-05,
'avour': 5.0020008003201284e-05,
'avril': 5.0020008003201284e-05,
'aver': 0.00015006002400960383,
'axian': 5.0020008003201284e-05,
'ay': 5.0020008003201284e-05,
'ayant': 0.00010004001600640257,
'ayez': 5.0020008003201284e-05,
'ayon': 5.0020008003201284e-05,
'azari': 5.0020008003201284e-05,
'azaël': 5.0020008003201284e-05,
'azur': 0.00015006002400960383,
'azem': 5.0020008003201284e-05,
'azémir': 5.0020008003201284e-05,
'azémon': 5.0020008003201284e-05,
'aïness': 5.0020008003201284e-05,
'aîn': 0.00010004001600640257,
'aïeul': 0.00010004001600640257,
'aïeux': 5.0020008003201284e-05,
'b': 5.0020008003201284e-05,
'baal': 5.0020008003201284e-05,
'babylon': 5.0020008003201284e-05,
'bacazar': 5.0020008003201284e-05,
'bacchus': 5.0020008003201284e-05,
'bactr': 5.0020008003201284e-05,
'bagatel': 5.0020008003201284e-05,
'bagu': 5.0020008003201284e-05,
'baign': 0.00035014005602240897,
'baignent': 5.0020008003201284e-05,
'baill': 0.00010004001600640257,
'baillif': 5.0020008003201284e-05,
'bain': 5.0020008003201284e-05,
'bais': 0.0002501000400160064,
'baisent': 5.0020008003201284e-05,
'baiser': 5.0020008003201284e-05,
'baiss': 0.0004501800720288115,
'baissent': 5.0020008003201284e-05,
'bajazet': 5.0020008003201284e-05,
'bal': 5.0020008003201284e-05,
'balanc': 0.0007002801120448179,
'balancent': 5.0020008003201284e-05,
'balançon': 5.0020008003201284e-05,
'bambin': 5.0020008003201284e-05,
'bandeau': 5.0020008003201284e-05,
'band': 5.0020008003201284e-05,
'bann': 0.0005502200880352141,
'banqui': 5.0020008003201284e-05,
'baptist': 5.0020008003201284e-05,
'baptêm': 5.0020008003201284e-05,
'barbar': 0.00020008003201280514,
'barb': 5.0020008003201284e-05,
'barbin': 5.0020008003201284e-05,
'barc': 5.0020008003201284e-05,
'baron': 0.00010004001600640257,
'barqu': 5.0020008003201284e-05,
'barrab': 5.0020008003201284e-05,
'barreau': 5.0020008003201284e-05,
'barri': 0.00010004001600640257,
'barrer': 5.0020008003201284e-05,
'barsin': 5.0020008003201284e-05,

'barses': 5.0020008003201284e-05,
'barthélem': 5.0020008003201284e-05,
'bas': 0.00010004001600640257,
'bass': 0.0002501000400160064,
'bassess': 0.00010004001600640257,
'bassin': 5.0020008003201284e-05,
'bat': 5.0020008003201284e-05,
'bataill': 0.00010004001600640257,
'bataillon': 0.00010004001600640257,
'batav': 5.0020008003201284e-05,
'batt': 0.00010004001600640257,
'batteux': 5.0020008003201284e-05,
'battr': 5.0020008003201284e-05,
'battu': 5.0020008003201284e-05,
'battus': 5.0020008003201284e-05,
'baudr': 5.0020008003201284e-05,
'baum': 5.0020008003201284e-05,
'bav': 5.0020008003201284e-05,
'beau': 0.00010004001600640257,
'beaucoup': 5.0020008003201284e-05,
'beauregard': 5.0020008003201284e-05,
'beaut': 0.00010004001600640257,
'beauval': 5.0020008003201284e-05,
'bel': 0.00015006002400960383,
'belliqu': 0.00010004001600640257,
'bellon': 0.00010004001600640257,
'belloy': 5.0020008003201284e-05,
'belmont': 5.0020008003201284e-05,
'ben': 0.0008003201280512205,
'bengal': 5.0020008003201284e-05,
'benjamin': 5.0020008003201284e-05,
'bensérad': 5.0020008003201284e-05,
'berceau': 0.00010004001600640257,
'berg': 5.0020008003201284e-05,
'bergerac': 5.0020008003201284e-05,
'berger': 5.0020008003201284e-05,
'bernard': 5.0020008003201284e-05,
'bernardht': 5.0020008003201284e-05,
'bertios': 5.0020008003201284e-05,
'besoin': 0.00010004001600640257,
'besongn': 5.0020008003201284e-05,
'bi': 5.0020008003201284e-05,
'bibl': 5.0020008003201284e-05,
'bibliothequ': 5.0020008003201284e-05,
'bien': 0.00010004001600640257,
'bienfais': 0.00020008003201280514,
'bienf': 5.0020008003201284e-05,
'bienfaiteur': 0.00010004001600640257,
'bienfaitric': 5.0020008003201284e-05,
'bienfait': 5.0020008003201284e-05,
'bienheur': 0.00010004001600640257,
'bienséanc': 0.00010004001600640257,
'bientôt': 5.0020008003201284e-05,
'bienveil': 5.0020008003201284e-05,
'bil': 5.0020008003201284e-05,
'billain': 5.0020008003201284e-05,
'billet': 0.00010004001600640257,
'bin': 5.0020008003201284e-05,
'bithyn': 5.0020008003201284e-05,
'bithynien': 5.0020008003201284e-05,
'bitum': 5.0020008003201284e-05,
'bizarr': 0.00010004001600640257,
'bizarrer': 5.0020008003201284e-05,
'bier': 5.0020008003201284e-05,
'blanc': 0.00010004001600640257,
'blanch': 0.00020008003201280514,
'blasphem': 0.00035014005602240897,
'bless': 0.0006502601040416166,
'blessent': 5.0020008003201284e-05,

```

'blessur': 0.00010004001600640257,
'blois': 5.0020008003201284e-05,
'blond': 0.00010004001600640257,
'blâmabl': 5.0020008003201284e-05,
'blâm': 0.0006502601040416166,
'blâment': 5.0020008003201284e-05,
'blé': 5.0020008003201284e-05,
'blêm': 0.00010004001600640257,
'bocag': 0.00010004001600640257,
...}

```

In [44]:

```

stemmed_vocabulary_tra_com = [fr.stem(word) for word in vocabulary_tra_com]
counts = dict()
for word in stemmed_vocabulary_tra_com:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_tra_com = counts
for key, value in counts.items():
    tf_dict_tra_com[key] = value / len(stemmed_vocabulary_tra_com)
tf_dict_tra_com

```

Out[44]:

```

{'a': 9.220839096357768e-05,
'abaiss': 0.0005532503457814661,
'abaissement': 9.220839096357768e-05,
'abandon': 0.0009220839096357768,
'abandonnent': 9.220839096357768e-05,
'abat': 9.220839096357768e-05,
'abattr': 9.220839096357768e-05,
'abattu': 0.00018441678192715537,
'abattus': 9.220839096357768e-05,
'abhorr': 9.220839096357768e-05,
'abject': 0.00018441678192715537,
'abois': 9.220839096357768e-05,
'abomin': 9.220839096357768e-05,
'abond': 0.00027662517289073305,
'abord': 0.00036883356385431073,
'abordon': 9.220839096357768e-05,
'abri': 9.220839096357768e-05,
'abreg': 0.00036883356385431073,
'absenc': 9.220839096357768e-05,
'absent': 0.00018441678192715537,
'absolu': 0.00027662517289073305,
'absolus': 9.220839096357768e-05,
'absous': 9.220839096357768e-05,
'absout': 9.220839096357768e-05,
'abus': 0.0010142923005993546,
'abusent': 9.220839096357768e-05,
'abîm': 0.00027662517289073305,
'accabl': 0.00036883356385431073,
'accablent': 9.220839096357768e-05,
'accapareur': 9.220839096357768e-05,
'accent': 9.220839096357768e-05,
'accept': 0.0005532503457814661,
'accident': 0.00018441678192715537,
'accommod': 0.00018441678192715537,
'accompagn': 0.0004610419548178884,
'accompagnent': 9.220839096357768e-05,
'accompl': 0.0006454587367450438,
'accord': 0.00119870908252651,
'accordent': 9.220839096357768e-05,
'accost': 9.220839096357768e-05,
'accouch': 9.220839096357768e-05,
'accourut': 9.220839096357768e-05,
'accoutr': 9.220839096357768e-05,
'accoutum': 0.00036883356385431073,

```


'accroir': 9.220839096357768e-05,
'accrois': 9.220839096357768e-05,
'accroissent': 9.220839096357768e-05,
'accroît': 9.220839096357768e-05,
'accroître': 9.220839096357768e-05,
'accrurent': 9.220839096357768e-05,
'accrut': 9.220839096357768e-05,
'accueil': 9.220839096357768e-05,
'accus': 0.0012909174734900876,
'accusent': 9.220839096357768e-05,
'accuson': 9.220839096357768e-05,
'acces': 9.220839096357768e-05,
'acharn': 9.220839096357768e-05,
'achat': 9.220839096357768e-05,
'achemin': 9.220839096357768e-05,
'achet': 0.00027662517289073305,
'achev': 0.0007376671277086215,
'achevon': 9.220839096357768e-05,
'achill': 9.220839096357768e-05,
'achèvent': 9.220839096357768e-05,
'aci': 9.220839096357768e-05,
'acquier': 9.220839096357768e-05,
'acquiert': 9.220839096357768e-05,
'acquis': 0.00018441678192715537,
'acquisit': 9.220839096357768e-05,
'acquit': 9.220839096357768e-05,
'acquitt': 0.00036883356385431073,
'acquitton': 9.220839096357768e-05,
'acquer': 9.220839096357768e-05,
'acte': 0.00018441678192715537,
'acteon': 9.220839096357768e-05,
'acteur': 9.220839096357768e-05,
'action': 0.00018441678192715537,
'actuel': 9.220839096357768e-05,
'adieu': 9.220839096357768e-05,
'adieux': 9.220839096357768e-05,
'admet': 0.00018441678192715537,
'admettr': 9.220839096357768e-05,
'admir': 0.0007376671277086215,
'admiron': 9.220839096357768e-05,
'adon': 9.220839096357768e-05,
'ador': 0.0013831258644536654,
'adorent': 9.220839096357768e-05,
'adoron': 9.220839096357768e-05,
'adouc': 0.0006454587367450438,
'adrast': 9.220839096357768e-05,
'adress': 0.0005532503457814661,
'adressent': 9.220839096357768e-05,
'adroit': 0.00027662517289073305,
'adulter': 0.00018441678192715537,
'advenu': 9.220839096357768e-05,
'advenus': 9.220839096357768e-05,
'adversair': 0.00018441678192715537,
'advers': 0.00027662517289073305,
'advien': 9.220839096357768e-05,
'advient': 9.220839096357768e-05,
'aegid': 9.220839096357768e-05,
'affabl': 9.220839096357768e-05,
'affaibl': 0.00036883356385431073,
'affair': 0.00018441678192715537,
'affam': 0.00018441678192715537,
'affect': 0.00036883356385431073,
'afferme': 0.00027662517289073305,
'afflict': 0.00018441678192715537,
'afflig': 0.0005532503457814661,
'affligent': 9.220839096357768e-05,
'affranch': 0.0004610419548178884,
'affreux': 0.00018441678192715537,
'affront': 0.00036883356385431073,

'afin': 9.220839096357768e-05,
'africain': 9.220839096357768e-05,
'afriqu': 9.220839096357768e-05,
'agent': 9.220839096357768e-05,
'agent': 0.00018441678192715537,
'agesilan': 9.220839096357768e-05,
'aggrav': 0.00018441678192715537,
'agi': 9.220839096357768e-05,
'agir': 9.220839096357768e-05,
'agis': 9.220839096357768e-05,
'agiss': 0.00036883356385431073,
'agit': 0.0005532503457814661,
'agitent': 9.220839096357768e-05,
'agrand': 0.00027662517289073305,
'agréabl': 0.00027662517289073305,
'agré': 0.00018441678192715537,
'agrément': 0.00018441678192715537,
'ah': 9.220839096357768e-05,
'ahi': 9.220839096357768e-05,
'ai': 9.220839096357768e-05,
'aid': 0.0006454587367450438,
'aident': 9.220839096357768e-05,
'aie': 0.00018441678192715537,
'aient': 9.220839096357768e-05,
'aigl': 9.220839096357768e-05,
'aigr': 0.0004610419548178884,
'aigreur': 9.220839096357768e-05,
'aigus': 9.220839096357768e-05,
'ail': 0.00018441678192715537,
'aill': 9.220839096357768e-05,
'ailleur': 9.220839096357768e-05,
'aim': 0.0021207929921622866,
'aimabl': 0.00018441678192715537,
'aient': 9.220839096357768e-05,
'aimon': 9.220839096357768e-05,
'ains': 9.220839096357768e-05,
'air': 0.00018441678192715537,
'airain': 9.220839096357768e-05,
'ais': 0.0005532503457814661,
'ait': 9.220839096357768e-05,
'ajout': 0.0006454587367450438,
'ajust': 0.00018441678192715537,
'alarm': 0.0005532503457814661,
'albion': 9.220839096357768e-05,
'albâtr': 9.220839096357768e-05,
'alcid': 9.220839096357768e-05,
'alcidor': 9.220839096357768e-05,
'alcimédon': 9.220839096357768e-05,
'alcin': 9.220839096357768e-05,
'alcion': 9.220839096357768e-05,
'alcir': 9.220839096357768e-05,
'alent': 9.220839096357768e-05,
'alentour': 9.220839096357768e-05,
'alerin': 9.220839096357768e-05,
'alexandr': 9.220839096357768e-05,
'aliment': 9.220839096357768e-05,
'alla': 9.220839096357768e-05,
'allaient': 9.220839096357768e-05,
'allais': 9.220839096357768e-05,
'allait': 9.220839096357768e-05,
'allant': 9.220839096357768e-05,
'aller': 9.220839096357768e-05,
'allez': 9.220839096357768e-05,
'allianc': 9.220839096357768e-05,
'alli': 0.0004610419548178884,
'allion': 9.220839096357768e-05,
'allomic': 9.220839096357768e-05,
'allong': 9.220839096357768e-05,
'allon': 9.220839096357768e-05,

'allum': 0.0006454587367450438,
'allument': 9.220839096357768e-05,
'allât': 9.220839096357768e-05,
'allech': 9.220839096357768e-05,
'alleg': 0.00027662517289073305,
'allé': 0.00018441678192715537,
'allége': 9.220839096357768e-05,
'allégor': 9.220839096357768e-05,
'allégress': 9.220839096357768e-05,
'allégu': 0.00018441678192715537,
'alons': 9.220839096357768e-05,
'alor': 9.220839096357768e-05,
'alphons': 9.220839096357768e-05,
'alphénor': 9.220839096357768e-05,
'alquif': 9.220839096357768e-05,
'altess': 9.220839096357768e-05,
'alti': 9.220839096357768e-05,
'altier': 9.220839096357768e-05,
'altomir': 9.220839096357768e-05,
'alter': 0.00036883356385431073,
'alve': 9.220839096357768e-05,
'amad': 9.220839096357768e-05,
'amant': 0.00027662517289073305,
'amarill': 9.220839096357768e-05,
'amas': 0.00018441678192715537,
'amazon': 9.220839096357768e-05,
'ambassadeur': 0.00018441678192715537,
'ambiti': 0.00018441678192715537,
'ambit': 9.220839096357768e-05,
'amen': 0.0004610419548178884,
'amend': 0.00018441678192715537,
'amer': 0.00027662517289073305,
'amerin': 9.220839096357768e-05,
'amertum': 9.220839096357768e-05,
'ami': 0.00018441678192715537,
'aminth': 9.220839096357768e-05,
'amiral': 9.220839096357768e-05,
'amis': 9.220839096357768e-05,
'amiti': 0.00018441678192715537,
'amoindr': 9.220839096357768e-05,
'amoll': 0.00027662517289073305,
'amorc': 0.00018441678192715537,
'amort': 9.220839096357768e-05,
'amour': 0.0004610419548178884,
'ample': 9.220839096357768e-05,
'amplement': 9.220839096357768e-05,
'amus': 0.00036883356385431073,
'amé': 0.00018441678192715537,
'amel': 9.220839096357768e-05,
'amérin': 9.220839096357768e-05,
'an': 0.00018441678192715537,
'anaxandr': 9.220839096357768e-05,
'ancien': 0.00036883356385431073,
'ancêtr': 9.220839096357768e-05,
'andalous': 9.220839096357768e-05,
'andre': 9.220839096357768e-05,
'andrigen': 9.220839096357768e-05,
'andrion': 9.220839096357768e-05,
'andré': 0.00018441678192715537,
'ange': 0.00018441678192715537,
'angleterr': 9.220839096357768e-05,
'angel': 9.220839096357768e-05,
'anim': 0.0007376671277086215,
'animal': 0.00018441678192715537,
'animent': 9.220839096357768e-05,
'animos': 9.220839096357768e-05,
'anne': 9.220839096357768e-05,
'anneau': 9.220839096357768e-05,
'annonc': 0.00036883356385431073,

'anné': 0.00018441678192715537,
'anthoin': 9.220839096357768e-05,
'anthé': 9.220839096357768e-05,
'antichambr': 9.220839096357768e-05,
'antigon': 9.220839096357768e-05,
'antiochus': 9.220839096357768e-05,
'antipath': 9.220839096357768e-05,
'antiqu': 0.00018441678192715537,
'antoin': 9.220839096357768e-05,
'antre': 0.00018441678192715537,
'anté': 9.220839096357768e-05,
'anéant': 9.220839096357768e-05,
'août': 9.220839096357768e-05,
'apais': 0.0005532503457814661,
'aperçois': 9.220839096357768e-05,
'aperçoit': 9.220839096357768e-05,
'aperçu': 0.00018441678192715537,
'aperçus': 9.220839096357768e-05,
'apollon': 9.220839096357768e-05,
'apparaît': 9.220839096357768e-05,
'appareil': 0.00018441678192715537,
'apparent': 0.0005532503457814661,
'appart': 9.220839096357768e-05,
'appartiendr': 9.220839096357768e-05,
'appartiennent': 9.220839096357768e-05,
'appartient': 9.220839096357768e-05,
'apparut': 9.220839096357768e-05,
'appas': 9.220839096357768e-05,
'appel': 0.001475334255417243,
'appellent': 9.220839096357768e-05,
'appesant': 9.220839096357768e-05,
'applaud': 0.00027662517289073305,
'applic': 9.220839096357768e-05,
'appliqu': 0.00018441678192715537,
'apport': 0.0005532503457814661,
'appren': 0.00027662517289073305,
'apprend': 0.00018441678192715537,
'apprendr': 0.0004610419548178884,
'apprendron': 9.220839096357768e-05,
'apprendront': 9.220839096357768e-05,
'apprennent': 9.220839096357768e-05,
'apprenon': 9.220839096357768e-05,
'apprirent': 9.220839096357768e-05,
'appris': 0.00018441678192715537,
'apprit': 9.220839096357768e-05,
'approb': 9.220839096357768e-05,
'approch': 0.0007376671277086215,
'approchon': 9.220839096357768e-05,
'approuv': 0.0010142923005993546,
'approuvent': 9.220839096357768e-05,
'appréhend': 0.00036883356385431073,
'appréhens': 9.220839096357768e-05,
'apprêt': 0.00018441678192715537,
'appui': 0.0005532503457814661,
'appuis': 9.220839096357768e-05,
'appet': 9.220839096357768e-05,
'apre': 9.220839096357768e-05,
'apres': 9.220839096357768e-05,
'aragon': 9.220839096357768e-05,
'arax': 9.220839096357768e-05,
'araxes': 9.220839096357768e-05,
'arbas': 9.220839096357768e-05,
'arbitr': 9.220839096357768e-05,
'arbre': 0.00018441678192715537,
'arc': 0.00018441678192715537,
'arcas': 9.220839096357768e-05,
'archel': 9.220839096357768e-05,
'archelaus': 9.220839096357768e-05,
'archer': 0.00018441678192715537,

'architalass': 9.220839096357768e-05,
'arctodem': 9.220839096357768e-05,
'ardemment': 9.220839096357768e-05,
'arden': 9.220839096357768e-05,
'ardent': 0.00027662517289073305,
'ardeur': 0.00018441678192715537,
'argant': 9.220839096357768e-05,
'argen': 0.00027662517289073305,
'argent': 9.220839096357768e-05,
'argument': 9.220839096357768e-05,
'argus': 9.220839096357768e-05,
'arian': 9.220839096357768e-05,
'arist': 9.220839096357768e-05,
'aristocrat': 0.00018441678192715537,
'arland': 9.220839096357768e-05,
'arma': 9.220839096357768e-05,
'armait': 9.220839096357768e-05,
'armand': 9.220839096357768e-05,
'arme': 0.00018441678192715537,
'armer': 0.00018441678192715537,
'armeront': 9.220839096357768e-05,
'armet': 9.220839096357768e-05,
'armez': 9.220839096357768e-05,
'armé': 0.00036883356385431073,
'armen': 9.220839096357768e-05,
'aronth': 9.220839096357768e-05,
'arrach': 0.0007376671277086215,
'arrachent': 9.220839096357768e-05,
'arrachon': 9.220839096357768e-05,
'arrang': 9.220839096357768e-05,
'arrhe': 9.220839096357768e-05,
'arriv': 0.0010142923005993546,
'arrivent': 9.220839096357768e-05,
'arrivon': 9.220839096357768e-05,
'arrier': 9.220839096357768e-05,
'arrog': 0.00027662517289073305,
'arros': 0.00027662517289073305,
'arrêt': 0.001475334255417243,
'art': 0.00018441678192715537,
'artaxerc': 9.220839096357768e-05,
'articl': 9.220839096357768e-05,
'artific': 0.00018441678192715537,
'artisan': 9.220839096357768e-05,
'artois': 9.220839096357768e-05,
'artes': 9.220839096357768e-05,
'aren': 9.220839096357768e-05,
'as': 9.220839096357768e-05,
'ascalon': 9.220839096357768e-05,
'ascend': 9.220839096357768e-05,
'asi': 9.220839096357768e-05,
'asil': 9.220839096357768e-05,
'aspar': 9.220839096357768e-05,
'aspect': 0.00018441678192715537,
'aspic': 9.220839096357768e-05,
'aspir': 0.00027662517289073305,
'aspirent': 9.220839096357768e-05,
'assaill': 0.00018441678192715537,
'assassin': 0.0004610419548178884,
'assassinat': 0.00018441678192715537,
'assaut': 0.00018441678192715537,
'assembl': 0.0008298755186721991,
'asseoir': 9.220839096357768e-05,
'asserv': 0.00027662517289073305,
'assez': 9.220839096357768e-05,
'assidu': 9.220839096357768e-05,
'assidus': 9.220839096357768e-05,
'assign': 0.00018441678192715537,
'assis': 9.220839096357768e-05,
'assist': 0.0004610419548178884,

'assistent': 9.220839096357768e-05,
'associ': 9.220839096357768e-05,
'assomm': 9.220839096357768e-05,
'assort': 0.00018441678192715537,
'assoup': 0.00018441678192715537,
'assouv': 0.00036883356385431073,
'assujet': 0.00018441678192715537,
'assur': 0.0010142923005993546,
'assurent': 9.220839096357768e-05,
'assyrien': 9.220839096357768e-05,
'astre': 0.00018441678192715537,
'aster': 9.220839096357768e-05,
'atis': 9.220839096357768e-05,
'attach': 0.0008298755186721991,
'attachent': 9.220839096357768e-05,
'attaqu': 0.0008298755186721991,
'attaquent': 9.220839096357768e-05,
'attaquon': 9.220839096357768e-05,
'atteindr': 9.220839096357768e-05,
'attein': 9.220839096357768e-05,
'atteint': 0.00036883356385431073,
'attend': 0.0007376671277086215,
'attendent': 9.220839096357768e-05,
'attendon': 9.220839096357768e-05,
'attendr': 0.0004610419548178884,
'attendu': 0.00018441678192715537,
'attentat': 0.00018441678192715537,
'attent': 0.0007376671277086215,
'attest': 9.220839096357768e-05,
'attir': 0.0005532503457814661,
'attirent': 9.220839096357768e-05,
'attrait': 9.220839096357768e-05,
'attrap': 0.00027662517289073305,
'attribu': 9.220839096357768e-05,
'au': 9.220839096357768e-05,
'aucun': 0.00036883356385431073,
'audac': 9.220839096357768e-05,
'audaci': 9.220839096357768e-05,
'audienc': 9.220839096357768e-05,
'audit': 9.220839096357768e-05,
'augment': 0.0005532503457814661,
'augur': 9.220839096357768e-05,
'august': 0.00018441678192715537,
'augustin': 9.220839096357768e-05,
'aujourd': 9.220839096357768e-05,
'aujourdhui': 9.220839096357768e-05,
'aumôn': 9.220839096357768e-05,
'auparav': 9.220839096357768e-05,
'aupres': 9.220839096357768e-05,
'auquel': 9.220839096357768e-05,
'aur': 0.0007376671277086215,
'aurion': 9.220839096357768e-05,
'auron': 9.220839096357768e-05,
'auront': 9.220839096357768e-05,
'auror': 9.220839096357768e-05,
'auspic': 9.220839096357768e-05,
'auss': 9.220839096357768e-05,
'aussitôt': 9.220839096357768e-05,
'auster': 0.00018441678192715537,
'aut': 9.220839096357768e-05,
'autel': 0.00018441678192715537,
'auteur': 0.00018441678192715537,
'authent': 9.220839096357768e-05,
'automn': 9.220839096357768e-05,
'autoris': 0.00036883356385431073,
'autorisent': 9.220839096357768e-05,
'autor': 9.220839096357768e-05,
'autour': 9.220839096357768e-05,
'autr': 0.00027662517289073305,

'autrefois': 9.220839096357768e-05,
'autrui': 9.220839096357768e-05,
'aux': 9.220839096357768e-05,
'auxquel': 9.220839096357768e-05,
'avaient': 9.220839096357768e-05,
'avais': 9.220839096357768e-05,
'avait': 9.220839096357768e-05,
'aval': 9.220839096357768e-05,
'avanc': 0.0005532503457814661,
'avant': 9.220839096357768e-05,
'avantag': 0.0004610419548178884,
'avançon': 9.220839096357768e-05,
'avar': 0.00018441678192715537,
'avaric': 9.220839096357768e-05,
'avec': 9.220839096357768e-05,
'avecqu': 0.00018441678192715537,
'aven': 9.220839096357768e-05,
'aventur': 0.00018441678192715537,
'avers': 9.220839096357768e-05,
'avert': 0.0004610419548178884,
'aveu': 9.220839096357768e-05,
'aveugl': 0.0007376671277086215,
'avez': 9.220839096357768e-05,
'avid': 0.00027662517289073305,
'avi': 9.220839096357768e-05,
'avion': 9.220839096357768e-05,
'avis': 0.00036883356385431073,
'avocat': 9.220839096357768e-05,
'avoient': 9.220839096357768e-05,
'avoir': 9.220839096357768e-05,
'avon': 9.220839096357768e-05,
'avort': 9.220839096357768e-05,
'avou': 0.0006454587367450438,
'avril': 9.220839096357768e-05,
'aver': 0.00027662517289073305,
'axian': 9.220839096357768e-05,
'ayant': 0.00018441678192715537,
'ayez': 9.220839096357768e-05,
'aymon': 9.220839096357768e-05,
'ayon': 9.220839096357768e-05,
'azur': 0.00018441678192715537,
'aïness': 9.220839096357768e-05,
'aîn': 9.220839096357768e-05,
'aïeul': 0.00018441678192715537,
'aïeux': 9.220839096357768e-05,
'babylon': 9.220839096357768e-05,
'badin': 9.220839096357768e-05,
'bagag': 9.220839096357768e-05,
'bagu': 9.220839096357768e-05,
'baign': 0.00018441678192715537,
'baillif': 9.220839096357768e-05,
'bais': 0.0004610419548178884,
'baiser': 9.220839096357768e-05,
'baiss': 0.00018441678192715537,
'balanc': 0.00036883356385431073,
'balcon': 9.220839096357768e-05,
'ballet': 9.220839096357768e-05,
'band': 0.00027662517289073305,
'bandeau': 9.220839096357768e-05,
'bann': 0.0008298755186721991,
'banqui': 9.220839096357768e-05,
'barbar': 0.00036883356385431073,
'barb': 0.00018441678192715537,
'barbi': 9.220839096357768e-05,
'barbin': 9.220839096357768e-05,
'barbus': 9.220839096357768e-05,
'baron': 9.220839096357768e-05,
'barqu': 9.220839096357768e-05,
'barri': 9.220839096357768e-05,

'barsin': 9.220839096357768e-05,
'bas': 0.00018441678192715537,
'basilic': 9.220839096357768e-05,
'basion': 9.220839096357768e-05,
'bass': 0.00018441678192715537,
'bassess': 0.00018441678192715537,
'bassin': 9.220839096357768e-05,
'bast': 9.220839096357768e-05,
'baston': 9.220839096357768e-05,
'bat': 0.00018441678192715537,
'bataill': 0.00018441678192715537,
'bataillon': 9.220839096357768e-05,
'batt': 0.00018441678192715537,
'batton': 9.220839096357768e-05,
'battr': 9.220839096357768e-05,
'battu': 9.220839096357768e-05,
'battus': 9.220839096357768e-05,
'baudet': 9.220839096357768e-05,
'beau': 0.00018441678192715537,
'beaucoup': 9.220839096357768e-05,
'beaufort': 9.220839096357768e-05,
'beaut': 0.00018441678192715537,
'bel': 0.00027662517289073305,
'belliqu': 9.220839096357768e-05,
'bellon': 9.220839096357768e-05,
'beltà': 9.220839096357768e-05,
'benzamin': 9.220839096357768e-05,
'berceau': 9.220839096357768e-05,
'berg': 9.220839096357768e-05,
'berger': 0.00027662517289073305,
'bernard': 9.220839096357768e-05,
'bern': 9.220839096357768e-05,
'besogn': 9.220839096357768e-05,
'besoin': 0.00018441678192715537,
'bian': 9.220839096357768e-05,
'bibliothequ': 9.220839096357768e-05,
'bien': 0.00018441678192715537,
'bienfais': 9.220839096357768e-05,
'bienf': 9.220839096357768e-05,
'bienfait': 9.220839096357768e-05,
'bienheur': 0.00027662517289073305,
'bienséanc': 9.220839096357768e-05,
'bientôt': 9.220839096357768e-05,
'bienveil': 9.220839096357768e-05,
'bienvenu': 9.220839096357768e-05,
'bil': 9.220839096357768e-05,
'billet': 9.220839096357768e-05,
'bimeaur': 9.220839096357768e-05,
'bin': 9.220839096357768e-05,
'bizarr': 0.00018441678192715537,
'bizarrer': 9.220839096357768e-05,
'blanc': 0.00018441678192715537,
'blanch': 0.00018441678192715537,
'blancheur': 9.220839096357768e-05,
'blandomir': 9.220839096357768e-05,
'blasphem': 9.220839096357768e-05,
'bled': 9.220839096357768e-05,
'bless': 0.0010142923005993546,
'blessent': 9.220839096357768e-05,
'blessur': 0.00018441678192715537,
'bleu': 9.220839096357768e-05,
'blâmabl': 9.220839096357768e-05,
'blâm': 0.0008298755186721991,
'blâment': 9.220839096357768e-05,
'blêm': 9.220839096357768e-05,
'bocag': 9.220839096357768e-05,
'boeuf': 9.220839096357768e-05,
'boir': 9.220839096357768e-05,
'bois': 9.220839096357768e-05,

'boivent': 9.220839096357768e-05,
'bon': 0.00036883356385431073,
'bonac': 9.220839096357768e-05,
'bond': 9.220839096357768e-05,
'bonheur': 0.00018441678192715537,
'bonhomm': 9.220839096357768e-05,
'bonjour': 9.220839096357768e-05,
'bonnet': 9.220839096357768e-05,
'bonsoir': 9.220839096357768e-05,
'bont': 0.00018441678192715537,
'bord': 0.00018441678192715537,
'bordeau': 9.220839096357768e-05,
'borde1': 9.220839096357768e-05,
'born': 0.0009220839096357768,
'bornent': 9.220839096357768e-05,
'bornon': 9.220839096357768e-05,
'bouch': 0.00018441678192715537,
'bou': 9.220839096357768e-05,
'bouffon': 9.220839096357768e-05,
'boug': 9.220839096357768e-05,
'bouill': 0.00027662517289073305,
'bouillon': 0.00018441678192715537,
'boulevers': 9.220839096357768e-05,
'bouquet': 9.220839096357768e-05,
'bourg': 9.220839096357768e-05,
'bourgeois': 9.220839096357768e-05,
'bourgogn': 9.220839096357768e-05,
'bourreau': 0.00018441678192715537,
'bours': 9.220839096357768e-05,
'bout': 0.00018441678192715537,
'boutiqu': 9.220839096357768e-05,
'boec': 9.220839096357768e-05,
'bradam': 9.220839096357768e-05,
'branch': 0.00018441678192715537,
'bran1': 0.00018441678192715537,
'bras': 9.220839096357768e-05,
'brasi': 0.00018441678192715537,
'brass': 9.220839096357768e-05,
'brav': 0.00036883356385431073,
'bravour': 9.220839096357768e-05,
'bref': 9.220839096357768e-05,
'bretagn': 9.220839096357768e-05,
'bretteur': 9.220839096357768e-05,
'brid': 9.220839096357768e-05,
'brigand': 9.220839096357768e-05,
'brigu': 0.00018441678192715537,
'brill': 0.0005532503457814661,
'brillant': 9.220839096357768e-05,
'bris': 0.0004610419548178884,
'brisent': 9.220839096357768e-05,
'brison': 9.220839096357768e-05,
'bronz': 9.220839096357768e-05,
'bruit': 0.00018441678192715537,
'brun': 9.220839096357768e-05,
'brutal': 0.00036883356385431073,
'brech': 9.220839096357768e-05,
'brûl': 0.0012909174734900876,
'brûlent': 9.220839096357768e-05,
'bu': 9.220839096357768e-05,
'buisson': 9.220839096357768e-05,
'bulgar': 9.220839096357768e-05,
'bulgarien': 9.220839096357768e-05,
'bus': 9.220839096357768e-05,
'but': 9.220839096357768e-05,
'butin': 9.220839096357768e-05,
'butt': 9.220839096357768e-05,
'bât': 0.0005532503457814661,
'bâton': 9.220839096357768e-05,
'béatr': 9.220839096357768e-05,

'béatrix': 9.220839096357768e-05,
'béli': 9.220839096357768e-05,
'bénin': 9.220839096357768e-05,
'ben': 0.0006454587367450438,
'béront': 9.220839096357768e-05,
'bérénic': 9.220839096357768e-05,
'bêt': 0.00018441678192715537,
'bûch': 9.220839096357768e-05,
'bûcheron': 9.220839096357768e-05,
'caban': 9.220839096357768e-05,
'cabaret': 9.220839096357768e-05,
'cabinet': 9.220839096357768e-05,
'cabrer': 9.220839096357768e-05,
'cach': 0.0012909174734900876,
'cachent': 9.220839096357768e-05,
'cachon': 9.220839096357768e-05,
'cachot': 0.00018441678192715537,
'cadavr': 9.220839096357768e-05,
'cadenc': 9.220839096357768e-05,
'cadet': 9.220839096357768e-05,
'caduc': 9.220839096357768e-05,
'cag': 9.220839096357768e-05,
'cajol': 0.00018441678192715537,
'calam': 0.00018441678192715537,
'calar': 9.220839096357768e-05,
'calist': 9.220839096357768e-05,
'calli': 9.220839096357768e-05,
'calm': 0.0004610419548178884,
'calomn': 9.220839096357768e-05,
'calott': 9.220839096357768e-05,
'calprened': 9.220839096357768e-05,
'calsild': 9.220839096357768e-05,
'camarad': 9.220839096357768e-05,
'camill': 9.220839096357768e-05,
'camp': 0.00027662517289073305,
'campagn': 0.00018441678192715537,
'canaill': 9.220839096357768e-05,
'cand': 9.220839096357768e-05,
'canon': 9.220839096357768e-05,
'canton': 9.220839096357768e-05,
'capabl': 0.00018441678192715537,
'capadoc': 9.220839096357768e-05,
'capitain': 0.00018441678192715537,
'capital': 9.220839096357768e-05,
'capitol': 9.220839096357768e-05,
'capric': 0.00018441678192715537,
'caprici': 0.00018441678192715537,
'captif': 0.00018441678192715537,
'captiv': 0.0005532503457814661,
'captivent': 9.220839096357768e-05,
'captur': 9.220839096357768e-05,
'car': 9.220839096357768e-05,
'character': 0.00018441678192715537,
'carcam': 9.220839096357768e-05,
'carden': 9.220839096357768e-05,
'cardill': 9.220839096357768e-05,
'caress': 0.00036883356385431073,
'carlos': 9.220839096357768e-05,
'carnag': 9.220839096357768e-05,
'carquois': 9.220839096357768e-05,
'carreau': 9.220839096357768e-05,
'carri': 9.220839096357768e-05,
'carross': 9.220839096357768e-05,
'cart': 0.00018441678192715537,
'cartel': 9.220839096357768e-05,
'carthag': 9.220839096357768e-05,
'cas': 9.220839096357768e-05,
'casaqu': 9.220839096357768e-05,
'casqu': 9.220839096357768e-05,

'cassandr': 9.220839096357768e-05,
'cass': 0.00018441678192715537,
'castillan': 9.220839096357768e-05,
'castill': 9.220839096357768e-05,
'catepan': 9.220839096357768e-05,
'catherin': 9.220839096357768e-05,
'caus': 0.0011065006915629322,
'causent': 9.220839096357768e-05,
'caution': 9.220839096357768e-05,
'cavali': 0.00018441678192715537,
'ce': 0.00018441678192715537,
'cec': 9.220839096357768e-05,
'cel': 0.0006454587367450438,
'celui': 9.220839096357768e-05,
'cendr': 0.00018441678192715537,
'cen': 9.220839096357768e-05,
'censeur': 0.00018441678192715537,
'censur': 0.00018441678192715537,
'cent': 0.00018441678192715537,
'centr': 9.220839096357768e-05,
'cepend': 9.220839096357768e-05,
'cercl': 9.220839096357768e-05,
'cercueil': 0.00018441678192715537,
'cerf': 9.220839096357768e-05,
'cert': 0.00018441678192715537,
'certain': 0.0004610419548178884,
'certitud': 9.220839096357768e-05,
'cerveau': 9.220839096357768e-05,
'cervel': 9.220839096357768e-05,
'cesped': 9.220839096357768e-05,
'cess': 0.0009220839096357768,
'cessent': 9.220839096357768e-05,
'cession': 9.220839096357768e-05,
'cest': 0.00018441678192715537,
'cet': 0.00027662517289073305,
'ceux': 9.220839096357768e-05,
'ch': 9.220839096357768e-05,
'chacun': 0.00018441678192715537,
'chagrin': 0.00027662517289073305,
'chain': 9.220839096357768e-05,
'chair': 9.220839096357768e-05,
'chais': 9.220839096357768e-05,
'chaleur': 0.00018441678192715537,
'chambr': 9.220839096357768e-05,
'champ': 0.00018441678192715537,
'champion': 0.00018441678192715537,
'champêtr': 9.220839096357768e-05,
'chancel': 0.00027662517289073305,
'chanceli': 9.220839096357768e-05,
'chandel': 9.220839096357768e-05,
'chang': 0.001751959428307976,
'changent': 9.220839096357768e-05,
'changeon': 9.220839096357768e-05,
'chanson': 0.00018441678192715537,
'chant': 0.00036883356385431073,
'chanton': 9.220839096357768e-05,
'chaos': 9.220839096357768e-05,
'chapeau': 9.220839096357768e-05,
'chaqu': 9.220839096357768e-05,
'char': 0.00018441678192715537,
'charbon': 0.00018441678192715537,
'charg': 0.0006454587367450438,
'chargent': 9.220839096357768e-05,
'chariot': 9.220839096357768e-05,
'charit': 0.00036883356385431073,
'charl': 9.220839096357768e-05,
'charlimin': 9.220839096357768e-05,
'charlymin': 9.220839096357768e-05,
'charm': 0.0013831258644536654,

'charment': 9.220839096357768e-05,
'charmeur': 0.00018441678192715537,
'chartr': 9.220839096357768e-05,
'chass': 0.0009220839096357768,
'chassent': 9.220839096357768e-05,
'chasseur': 9.220839096357768e-05,
'chasson': 9.220839096357768e-05,
'chast': 0.00018441678192715537,
'chat': 0.00018441678192715537,
'chaud': 9.220839096357768e-05,
'chaum': 9.220839096357768e-05,
'chauss': 9.220839096357768e-05,
'chauv': 9.220839096357768e-05,
'chaîn': 0.00018441678192715537,
'che': 9.220839096357768e-05,
'chef': 0.00018441678192715537,
'chemin': 0.00018441678192715537,
'chemis': 9.220839096357768e-05,
'cher': 0.0010142923005993546,
'cherch': 0.00119870908252651,
'cherchent': 9.220839096357768e-05,
'cherchon': 9.220839096357768e-05,
'cheval': 0.00018441678192715537,
'chevaler': 9.220839096357768e-05,
'chevali': 0.00018441678192715537,
'cheveux': 9.220839096357768e-05,
'chevillard': 9.220839096357768e-05,
'chevill': 9.220839096357768e-05,
'chez': 9.220839096357768e-05,
'chien': 0.00027662517289073305,
'chimen': 9.220839096357768e-05,
'chimer': 0.00018441678192715537,
'choc': 9.220839096357768e-05,
'choir': 9.220839096357768e-05,
'chois': 0.0010142923005993546,
'choix': 9.220839096357768e-05,
'choqu': 0.00036883356385431073,
'choquent': 9.220839096357768e-05,
'choratel': 9.220839096357768e-05,
'chos': 0.00018441678192715537,
'chrys': 9.220839096357768e-05,
'chut': 9.220839096357768e-05,
'chypr': 9.220839096357768e-05,
'château': 0.00018441678192715537,
'châtelet': 9.220839096357768e-05,
'châti': 9.220839096357768e-05,
'chât': 0.00018441678192715537,
'chétif': 9.220839096357768e-05,
'chétiv': 9.220839096357768e-05,
'ci': 9.220839096357768e-05,
'cicéron': 9.220839096357768e-05,
'ciel': 9.220839096357768e-05,
'cieux': 9.220839096357768e-05,
'ciment': 9.220839096357768e-05,
'cimeterr': 9.220839096357768e-05,
'cimeti': 9.220839096357768e-05,
'cinq': 9.220839096357768e-05,
'cinqu': 9.220839096357768e-05,
'cinquiem': 9.220839096357768e-05,
'circonst': 0.00018441678192715537,
'cir': 9.220839096357768e-05,
'ciseau': 9.220839096357768e-05,
'citoyen': 0.00018441678192715537,
'cit': 0.00018441678192715537,
'civil': 0.0004610419548178884,
'clair': 0.0004610419548178884,
'clairmont': 9.220839096357768e-05,
'clairvill': 9.220839096357768e-05,
'clameur': 0.00018441678192715537,

'clarind': 9.220839096357768e-05,
'clarin': 9.220839096357768e-05,
'clariont': 9.220839096357768e-05,
'clart': 0.00018441678192715537,
'claud': 9.220839096357768e-05,
'clavich': 9.220839096357768e-05,
'clef': 0.00018441678192715537,
'clerg': 9.220839096357768e-05,
'clericorum': 9.220839096357768e-05,
'clidam': 9.220839096357768e-05,
'climat': 0.00018441678192715537,
'climen': 9.220839096357768e-05,
'clin': 9.220839096357768e-05,
'clindor': 9.220839096357768e-05,
'clion': 9.220839096357768e-05,
'cloch': 9.220839096357768e-05,
'clor': 0.00018441678192715537,
'cloridan': 9.220839096357768e-05,
'clos': 0.00018441678192715537,
'clous': 9.220839096357768e-05,
'club': 9.220839096357768e-05,
'cléandr': 9.220839096357768e-05,
'cléarqu': 9.220839096357768e-05,
'clémenc': 9.220839096357768e-05,
'clément': 0.00018441678192715537,
'cléodat': 9.220839096357768e-05,
'cléomir': 9.220839096357768e-05,
'cléopâtr': 9.220839096357768e-05,
'coadjuteur': 9.220839096357768e-05,
'cocard': 9.220839096357768e-05,
'coch': 9.220839096357768e-05,
'coeur': 0.00018441678192715537,
'coffr': 9.220839096357768e-05,
'coiff': 9.220839096357768e-05,
'coin': 9.220839096357768e-05,
'col': 9.220839096357768e-05,
'colchos': 9.220839096357768e-05,
'collation': 9.220839096357768e-05,
'collet': 9.220839096357768e-05,
'colli': 9.220839096357768e-05,
'colleg': 9.220839096357768e-05,
'coll': 9.220839096357768e-05,
'colon': 9.220839096357768e-05,
'color': 0.00027662517289073305,
'coler': 0.00018441678192715537,
'combat': 0.00018441678192715537,
'combatt': 0.0007376671277086215,
'combattent': 9.220839096357768e-05,
'combatton': 9.220839096357768e-05,
'combattr': 0.00027662517289073305,
'combattu': 0.00018441678192715537,
'combattus': 9.220839096357768e-05,
'combien': 9.220839096357768e-05,
'combl': 0.0008298755186721991,
'command': 0.0008298755186721991,
'comm': 0.00018441678192715537,
'commenc': 0.00119870908252651,
'commencent': 9.220839096357768e-05,
'comment': 9.220839096357768e-05,
'commençon': 9.220839096357768e-05,
'commerc': 0.00018441678192715537,
'commet': 0.00018441678192715537,
'commettr': 9.220839096357768e-05,
'commis': 9.220839096357768e-05,
'commiss': 9.220839096357768e-05,
'commod': 0.00027662517289073305,
'commun': 0.0004610419548178884,
'communaut': 9.220839096357768e-05,
'communiqu': 9.220839096357768e-05,

```

'compagn': 0.00027662517289073305,
'compagnon': 0.00018441678192715537,
'compar': 0.00027662517289073305,
'comparaison': 0.00018441678192715537,
'compass': 9.220839096357768e-05,
'compat': 9.220839096357768e-05,
'complair': 9.220839096357768e-05,
'complais': 0.00027662517289073305,
'complic': 0.00018441678192715537,
'compl': 0.00018441678192715537,
'complot': 0.00018441678192715537,
'compos': 0.0005532503457814661,
'composent': 9.220839096357768e-05,
'composit': 9.220839096357768e-05,
'comprend': 0.00018441678192715537,
'comprendr': 9.220839096357768e-05,
'compr': 9.220839096357768e-05,
'compt': 0.0004610419548178884,
'comper': 9.220839096357768e-05,
'comt': 0.00018441678192715537,
'comtess': 9.220839096357768e-05,
'comed': 0.00018441678192715537,
'con': 0.00027662517289073305,
'concern': 0.00018441678192715537,
'concert': 9.220839096357768e-05,
'concevoir': 9.220839096357768e-05,
'concitoyen': 9.220839096357768e-05,
'conclu': 9.220839096357768e-05,
'conclur': 9.220839096357768e-05,
'conclus': 0.00018441678192715537,
'conclut': 9.220839096357768e-05,
'concord': 9.220839096357768e-05,
'concurrent': 0.00018441678192715537,
'condamn': 0.0010142923005993546,
'condamnent': 9.220839096357768e-05,
'condit': 0.00018441678192715537,
'conduir': 0.00018441678192715537,
'conduis': 0.00036883356385431073,
'conduisent': 9.220839096357768e-05,
'conduit': 0.00036883356385431073,
'confess': 0.0006454587367450438,
'confesson': 9.220839096357768e-05,
'confianc': 9.220839096357768e-05,
'confident': 0.00036883356385431073,
'conf': 9.220839096357768e-05,
'confi': 0.00018441678192715537,
'confirm': 0.0005532503457814661,
'confirment': 9.220839096357768e-05,
'confisc': 9.220839096357768e-05,
'confond': 0.00027662517289073305,
'confondr': 0.00018441678192715537,
'confondu': 9.220839096357768e-05,
'conform': 0.00027662517289073305,
'confus': 0.00036883356385431073,
'conférent': 9.220839096357768e-05,
'cong': 9.220839096357768e-05,
'conjectur': 9.220839096357768e-05,
'conjoint': 0.00036883356385431073,
'conjonctur': 9.220839096357768e-05,
...}

```

Q3: How many terms with a weight strictly larger than 0 do you have in each text genre profile?

This is simply the lengths of the tf vectors (note that we stemmed the terms!).

In [45]:

```

print("Terms for comedy: ", len(tf_dict_com))
print("Terms for tragedy: ", len(tf_dict_tra))
print("Terms for tragi-comedy: ", len(tf_dict_tra_com))

```

Terms for comedy: 14179
Terms for tragedy: 8902
Terms for tragi-comedy: 5288

Q4: Select randomly 10 plays for each text genre. Represent each play by a vector.

In [46]:

```
import random

plays_com_tok_rand = [preprocess_text(play, 'French') for play in random.choices(plays_com, k=10)]
plays_tra_tok_rand = [preprocess_text(play, 'French') for play in random.choices(plays_tra, k=10)]
plays_tra_com_tok_rand = [preprocess_text(play, 'French') for play in random.choices(plays_tra_com, k=10)]
vocabulary_com_rand = extract_vocabulary(plays_com_tok_rand, min_count=2)
vocabulary_tra_rand = extract_vocabulary(plays_tra_tok_rand, min_count=2)
vocabulary_tra_com_rand = extract_vocabulary(plays_tra_com_tok_rand, min_count=2)

stemmed_vocabulary_tra_rand = [fr.stem(word) for word in vocabulary_tra_rand]
counts = dict()
for word in vocabulary_tra_rand:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_tra_rand = counts
for key, value in counts.items():
    tf_dict_tra_rand[key] = value / len(vocabulary_tra_rand)
tf_dict_tra_rand
```

Out[46]:

```
{'a': 0.00016795431642593214,
 'abaissement': 0.00016795431642593214,
 'abaisser': 0.00016795431642593214,
 'abandonnai': 0.00016795431642593214,
 'abandonne': 0.00016795431642593214,
 'abandonnent': 0.00016795431642593214,
 'abandonner': 0.00016795431642593214,
 'abandonnez': 0.00016795431642593214,
 'abandonné': 0.00016795431642593214,
 'abandonnée': 0.00016795431642593214,
 'abattre': 0.00016795431642593214,
 'abattu': 0.00016795431642593214,
 'abattue': 0.00016795431642593214,
 'abbé': 0.00016795431642593214,
 'abhorre': 0.00016795431642593214,
 'abjure': 0.00016795431642593214,
 'abjurer': 0.00016795431642593214,
 'abolir': 0.00016795431642593214,
 'abominable': 0.00016795431642593214,
 'abonde': 0.00016795431642593214,
 'abord': 0.00016795431642593214,
 'abordant': 0.00016795431642593214,
 'aborder': 0.00016795431642593214,
 'abréger': 0.00016795431642593214,
 'absence': 0.00016795431642593214,
 'absent': 0.00016795431642593214,
 'absente': 0.00016795431642593214,
 'absolu': 0.00016795431642593214,
 'absolue': 0.00016795431642593214,
 'absolument': 0.00016795431642593214,
 'absolus': 0.00016795431642593214,
 'absout': 0.00016795431642593214,
 'abus': 0.00016795431642593214,
 'abusant': 0.00016795431642593214,
 'abuse': 0.00016795431642593214,
 'abuser': 0.00016795431642593214,
 'abusez': 0.00016795431642593214,
 'abusé': 0.00016795431642593214,
 'abusée': 0.00016795431642593214,
 'abîme': 0.00016795431642593214,
```

'académie': 0.00016795431642593214,
'accablait': 0.00016795431642593214,
'accablant': 0.00016795431642593214,
'accable': 0.00016795431642593214,
'accabler': 0.00016795431642593214,
'accablez': 0.00016795431642593214,
'accablée': 0.00016795431642593214,
'accepte': 0.00016795431642593214,
'accepter': 0.00016795431642593214,
'accepterais': 0.00016795431642593214,
'accepté': 0.00016795431642593214,
'accident': 0.00016795431642593214,
'accidents': 0.00016795431642593214,
'accompagne': 0.00016795431642593214,
'accompagnent': 0.00016795431642593214,
'accompagner': 0.00016795431642593214,
'accompagné': 0.00016795431642593214,
'accomplir': 0.00016795431642593214,
'accomplisse': 0.00016795431642593214,
'accord': 0.00016795431642593214,
'accorde': 0.00016795431642593214,
'accorder': 0.00016795431642593214,
'accorderez': 0.00016795431642593214,
'accordez': 0.00016795431642593214,
'accords': 0.00016795431642593214,
'accordé': 0.00016795431642593214,
'accordée': 0.00016795431642593214,
'accours': 0.00016795431642593214,
'accourt': 0.00016795431642593214,
'accoutumé': 0.00016795431642593214,
'accoutumés': 0.00016795431642593214,
'accroire': 0.00016795431642593214,
'accroît': 0.00016795431642593214,
'accroître': 0.00016795431642593214,
'accueil': 0.00016795431642593214,
'accusant': 0.00016795431642593214,
'accusateur': 0.00016795431642593214,
'accuse': 0.00016795431642593214,
'accuser': 0.00016795431642593214,
'accusez': 0.00016795431642593214,
'accusé': 0.00016795431642593214,
'accès': 0.00016795431642593214,
'acharné': 0.00016795431642593214,
'acheter': 0.00016795431642593214,
'acheté': 0.00016795431642593214,
'acheva': 0.00016795431642593214,
'achevant': 0.00016795431642593214,
'achever': 0.00016795431642593214,
'achevez': 0.00016795431642593214,
'achevons': 0.00016795431642593214,
'achevé': 0.00016795431642593214,
'achevée': 0.00016795431642593214,
'achève': 0.00016795431642593214,
'achèvera': 0.00016795431642593214,
'achèverai': 0.00016795431642593214,
'achéron': 0.00016795431642593214,
'acier': 0.00016795431642593214,
'acomat': 0.00016795431642593214,
'acquis': 0.00016795431642593214,
'acquise': 0.00016795431642593214,
'acquitte': 0.00016795431642593214,
'acquitter': 0.00016795431642593214,
'acquérir': 0.00016795431642593214,
'acte': 0.00016795431642593214,
'actes': 0.00016795431642593214,
'acteur': 0.00016795431642593214,
'acteurs': 0.00016795431642593214,
'action': 0.00016795431642593214,
'actions': 0.00016795431642593214,

'actrice': 0.00016795431642593214,
'adieu': 0.00016795431642593214,
'adieux': 0.00016795431642593214,
'admirable': 0.00016795431642593214,
'admirables': 0.00016795431642593214,
'admirant': 0.00016795431642593214,
'admirateur': 0.00016795431642593214,
'admirateurs': 0.00016795431642593214,
'admiration': 0.00016795431642593214,
'admire': 0.00016795431642593214,
'admirer': 0.00016795431642593214,
'adorable': 0.00016795431642593214,
'adorai': 0.00016795431642593214,
'adore': 0.00016795431642593214,
'adorer': 0.00016795431642593214,
'adorez': 0.00016795431642593214,
'adoré': 0.00016795431642593214,
'adouci': 0.00016795431642593214,
'adoucie': 0.00016795431642593214,
'adoucir': 0.00016795431642593214,
'adoucit': 0.00016795431642593214,
'adraste': 0.00016795431642593214,
'adressait': 0.00016795431642593214,
'adresse': 0.00016795431642593214,
'adresseront': 0.00016795431642593214,
'adressé': 0.00016795431642593214,
'adressées': 0.00016795431642593214,
'adrienne': 0.00016795431642593214,
'adultère': 0.00016795431642593214,
'advenu': 0.00016795431642593214,
'adversaire': 0.00016795431642593214,
'adversités': 0.00016795431642593214,
'affaiblir': 0.00016795431642593214,
'affaiblit': 0.00016795431642593214,
'affaire': 0.00016795431642593214,
'affaires': 0.00016795431642593214,
'affecter': 0.00016795431642593214,
'affection': 0.00016795431642593214,
'affermi': 0.00016795431642593214,
'affermir': 0.00016795431642593214,
'affermissiez': 0.00016795431642593214,
'afflige': 0.00016795431642593214,
'affliger': 0.00016795431642593214,
'affligée': 0.00016795431642593214,
'affranchi': 0.00016795431642593214,
'affranchir': 0.00016795431642593214,
'affreuse': 0.00016795431642593214,
'affreux': 0.00016795431642593214,
'affront': 0.00016795431642593214,
'affronter': 0.00016795431642593214,
'afin': 0.00016795431642593214,
'africain': 0.00016795431642593214,
'africains': 0.00016795431642593214,
'afrique': 0.00016795431642593214,
'agir': 0.00016795431642593214,
'agit': 0.00016795431642593214,
'agite': 0.00016795431642593214,
'agité': 0.00016795431642593214,
'agitée': 0.00016795431642593214,
'agréable': 0.00016795431642593214,
'agréables': 0.00016795431642593214,
'ah': 0.00016795431642593214,
'ai': 0.00016795431642593214,
'aide': 0.00016795431642593214,
'aider': 0.00016795431642593214,
'aidez': 0.00016795431642593214,
'aie': 0.00016795431642593214,
'aient': 0.00016795431642593214,
'aigle': 0.00016795431642593214,

'aigri': 0.00016795431642593214,
'aigrit': 0.00016795431642593214,
'ailes': 0.00016795431642593214,
'aille': 0.00016795431642593214,
'ailleurs': 0.00016795431642593214,
'aima': 0.00016795431642593214,
'aimable': 0.00016795431642593214,
'aimais': 0.00016795431642593214,
'aimait': 0.00016795431642593214,
'aimant': 0.00016795431642593214,
'aime': 0.00016795431642593214,
'aient': 0.00016795431642593214,
'aimer': 0.00016795431642593214,
'aimerais': 0.00016795431642593214,
'aimerais': 0.00016795431642593214,
'aimerait': 0.00016795431642593214,
'aimeriez': 0.00016795431642593214,
'aimes': 0.00016795431642593214,
'aimez': 0.00016795431642593214,
'aimiez': 0.00016795431642593214,
'aimons': 0.00016795431642593214,
'aimâtes': 0.00016795431642593214,
'aimé': 0.00016795431642593214,
'aimée': 0.00016795431642593214,
'ainsi': 0.00016795431642593214,
'air': 0.00016795431642593214,
'airain': 0.00016795431642593214,
'airs': 0.00016795431642593214,
'aise': 0.00016795431642593214,
'aisé': 0.00016795431642593214,
'aisée': 0.00016795431642593214,
'aisément': 0.00016795431642593214,
'ait': 0.00016795431642593214,
'ajoute': 0.00016795431642593214,
'ajouter': 0.00016795431642593214,
'ajoutez': 0.00016795431642593214,
'alarme': 0.00016795431642593214,
'alarmer': 0.00016795431642593214,
'alarmes': 0.00016795431642593214,
'alarmé': 0.00016795431642593214,
'alarmée': 0.00016795431642593214,
'alcide': 0.00016795431642593214,
'alcoran': 0.00016795431642593214,
'alentit': 0.00016795431642593214,
'alentour': 0.00016795431642593214,
'alexandre': 0.00016795431642593214,
'ali': 0.00016795431642593214,
'aliment': 0.00016795431642593214,
'allaient': 0.00016795431642593214,
'allais': 0.00016795431642593214,
'allait': 0.00016795431642593214,
'aller': 0.00016795431642593214,
'allez': 0.00016795431642593214,
'alliance': 0.00016795431642593214,
'allié': 0.00016795431642593214,
'alliés': 0.00016795431642593214,
'allons': 0.00016795431642593214,
'alluma': 0.00016795431642593214,
'allume': 0.00016795431642593214,
'allument': 0.00016795431642593214,
'allumé': 0.00016795431642593214,
'allumée': 0.00016795431642593214,
'allumés': 0.00016795431642593214,
'allégeance': 0.00016795431642593214,
'alléger': 0.00016795431642593214,
'allégresse': 0.00016795431642593214,
'alors': 0.00016795431642593214,
'altier': 0.00016795431642593214,
'altière': 0.00016795431642593214,

'altéré': 0.00016795431642593214,
'amant': 0.00016795431642593214,
'amante': 0.00016795431642593214,
'amants': 0.00016795431642593214,
'amas': 0.00016795431642593214,
'amateur': 0.00016795431642593214,
'amazone': 0.00016795431642593214,
'ambassadeur': 0.00016795431642593214,
'ambitieux': 0.00016795431642593214,
'ambition': 0.00016795431642593214,
'amenait': 0.00016795431642593214,
'amende': 0.00016795431642593214,
'amener': 0.00016795431642593214,
'amenez': 0.00016795431642593214,
'amené': 0.00016795431642593214,
'amenée': 0.00016795431642593214,
'amer': 0.00016795431642593214,
'amertume': 0.00016795431642593214,
'amestris': 0.00016795431642593214,
'ami': 0.00016795431642593214,
'amis': 0.00016795431642593214,
'amitié': 0.00016795431642593214,
'ammon': 0.00016795431642593214,
'amorce': 0.00016795431642593214,
'amour': 0.00016795431642593214,
'amoureuse': 0.00016795431642593214,
'amoureuses': 0.00016795431642593214,
'amoureux': 0.00016795431642593214,
'amours': 0.00016795431642593214,
'ample': 0.00016795431642593214,
'ablement': 0.00016795431642593214,
'amurat': 0.00016795431642593214,
'amusement': 0.00016795431642593214,
'amène': 0.00016795431642593214,
'an': 0.00016795431642593214,
'analyser': 0.00016795431642593214,
'anciens': 0.00016795431642593214,
'andromaque': 0.00016795431642593214,
'ange': 0.00016795431642593214,
'anglais': 0.00016795431642593214,
'animait': 0.00016795431642593214,
'animaux': 0.00016795431642593214,
'anime': 0.00016795431642593214,
'animer': 0.00016795431642593214,
'animé': 0.00016795431642593214,
'animée': 0.00016795431642593214,
'animées': 0.00016795431642593214,
'animés': 0.00016795431642593214,
'anneau': 0.00016795431642593214,
'annonce': 0.00016795431642593214,
'annoncer': 0.00016795431642593214,
'annoncez': 0.00016795431642593214,
'annoncée': 0.00016795431642593214,
'annonçait': 0.00016795431642593214,
'année': 0.00016795431642593214,
'années': 0.00016795431642593214,
'ans': 0.00016795431642593214,
'antigone': 0.00016795431642593214,
'antiope': 0.00016795431642593214,
'antique': 0.00016795431642593214,
'antiques': 0.00016795431642593214,
'antiquité': 0.00016795431642593214,
'antoine': 0.00016795431642593214,
'antre': 0.00016795431642593214,
'août': 0.00016795431642593214,
'apaisent': 0.00016795431642593214,
'apaiser': 0.00016795431642593214,
'aperçois': 0.00016795431642593214,
'aperçoit': 0.00016795431642593214,

'aperçu': 0.00016795431642593214,
'aperçus': 0.00016795431642593214,
'aplanir': 0.00016795431642593214,
'appareil': 0.00016795431642593214,
'apparement': 0.00016795431642593214,
'apparence': 0.00016795431642593214,
'apparent': 0.00016795431642593214,
'apparents': 0.00016795431642593214,
'appartement': 0.00016795431642593214,
'appartenir': 0.00016795431642593214,
'appartient': 0.00016795431642593214,
'appas': 0.00016795431642593214,
'appela': 0.00016795431642593214,
'appelait': 0.00016795431642593214,
'appeler': 0.00016795431642593214,
'appelez': 0.00016795431642593214,
'appelle': 0.00016795431642593214,
'appellent': 0.00016795431642593214,
'appellerai': 0.00016795431642593214,
'appelé': 0.00016795431642593214,
'applaudie': 0.00016795431642593214,
'applaudir': 0.00016795431642593214,
'applaudissements': 0.00016795431642593214,
'appliquer': 0.00016795431642593214,
'apporta': 0.00016795431642593214,
'apporte': 0.00016795431642593214,
'apporter': 0.00016795431642593214,
'apporté': 0.00016795431642593214,
'apprenant': 0.00016795431642593214,
'apprend': 0.00016795431642593214,
'apprendra': 0.00016795431642593214,
'apprendrai': 0.00016795431642593214,
'apprendras': 0.00016795431642593214,
'apprendre': 0.00016795431642593214,
'apprendrez': 0.00016795431642593214,
'apprends': 0.00016795431642593214,
'apprenez': 0.00016795431642593214,
'apprenne': 0.00016795431642593214,
'apprennent': 0.00016795431642593214,
'apprentissage': 0.00016795431642593214,
'appris': 0.00016795431642593214,
'approchant': 0.00016795431642593214,
'approche': 0.00016795431642593214,
'approcher': 0.00016795431642593214,
'approchez': 0.00016795431642593214,
'approuve': 0.00016795431642593214,
'approuver': 0.00016795431642593214,
'approuvé': 0.00016795431642593214,
'appréhende': 0.00016795431642593214,
'apprête': 0.00016795431642593214,
'apprêts': 0.00016795431642593214,
'appui': 0.00016795431642593214,
'appuyer': 0.00016795431642593214,
'appât': 0.00016795431642593214,
'après': 0.00016795431642593214,
'arabe': 0.00016795431642593214,
'arabie': 0.00016795431642593214,
'arbitre': 0.00016795431642593214,
'arde': 0.00016795431642593214,
'ardemment': 0.00016795431642593214,
'ardent': 0.00016795431642593214,
'ardente': 0.00016795431642593214,
'ardents': 0.00016795431642593214,
'ardeur': 0.00016795431642593214,
'ardeurs': 0.00016795431642593214,
'argie': 0.00016795431642593214,
'argos': 0.00016795431642593214,
'ariane': 0.00016795431642593214,
'aricie': 0.00016795431642593214,

'aristote': 0.00016795431642593214,
'armait': 0.00016795431642593214,
'arme': 0.00016795431642593214,
'armer': 0.00016795431642593214,
'armes': 0.00016795431642593214,
'armet': 0.00016795431642593214,
'armez': 0.00016795431642593214,
'armé': 0.00016795431642593214,
'armée': 0.00016795431642593214,
'armées': 0.00016795431642593214,
'armés': 0.00016795431642593214,
'arnauld': 0.00016795431642593214,
'arons': 0.00016795431642593214,
'arracha': 0.00016795431642593214,
'arrachant': 0.00016795431642593214,
'arrache': 0.00016795431642593214,
'arracher': 0.00016795431642593214,
'arrachez': 0.00016795431642593214,
'arrachée': 0.00016795431642593214,
'arrivant': 0.00016795431642593214,
'arrive': 0.00016795431642593214,
'arriver': 0.00016795431642593214,
'arrivé': 0.00016795431642593214,
'arrivée': 0.00016795431642593214,
'arrêt': 0.00016795431642593214,
'arrêtant': 0.00016795431642593214,
'arrête': 0.00016795431642593214,
'arrêtent': 0.00016795431642593214,
'arrêter': 0.00016795431642593214,
'arrêtes': 0.00016795431642593214,
'arrêtez': 0.00016795431642593214,
'arrêts': 0.00016795431642593214,
'arrêté': 0.00016795431642593214,
'arrêtée': 0.00016795431642593214,
'arsace': 0.00016795431642593214,
'art': 0.00016795431642593214,
'artaban': 0.00016795431642593214,
'artaxerce': 0.00016795431642593214,
'artifice': 0.00016795431642593214,
'artifices': 0.00016795431642593214,
'arts': 0.00016795431642593214,
'as': 0.00016795431642593214,
'asdrubal': 0.00016795431642593214,
'asie': 0.00016795431642593214,
'asile': 0.00016795431642593214,
'asiles': 0.00016795431642593214,
'aspect': 0.00016795431642593214,
'aspire': 0.00016795431642593214,
'assailli': 0.00016795431642593214,
'assaillir': 0.00016795431642593214,
'assassin': 0.00016795431642593214,
'assassinat': 0.00016795431642593214,
'assassine': 0.00016795431642593214,
'assassiner': 0.00016795431642593214,
'assassins': 0.00016795431642593214,
'assaut': 0.00016795431642593214,
'assauts': 0.00016795431642593214,
'assemblage': 0.00016795431642593214,
'assemble': 0.00016795431642593214,
'assembler': 0.00016795431642593214,
'assemblé': 0.00016795431642593214,
'assemblée': 0.00016795431642593214,
'assemblés': 0.00016795431642593214,
'asservi': 0.00016795431642593214,
'asservie': 0.00016795431642593214,
'asservir': 0.00016795431642593214,
'asservit': 0.00016795431642593214,
'assez': 0.00016795431642593214,
'assise': 0.00016795431642593214,

'assistance': 0.00016795431642593214,
'assiste': 0.00016795431642593214,
'assister': 0.00016795431642593214,
'assisté': 0.00016795431642593214,
'assiège': 0.00016795431642593214,
'assiégeants': 0.00016795431642593214,
'associer': 0.00016795431642593214,
'associé': 0.00016795431642593214,
'assouvie': 0.00016795431642593214,
'assujettir': 0.00016795431642593214,
'assurance': 0.00016795431642593214,
'assure': 0.00016795431642593214,
'assurent': 0.00016795431642593214,
'assurer': 0.00016795431642593214,
'assurez': 0.00016795431642593214,
'assuré': 0.00016795431642593214,
'assurée': 0.00016795431642593214,
'assurément': 0.00016795431642593214,
'assurés': 0.00016795431642593214,
'astre': 0.00016795431642593214,
'astres': 0.00016795431642593214,
'atalide': 0.00016795431642593214,
'athalie': 0.00016795431642593214,
'athènes': 0.00016795431642593214,
'athées': 0.00016795431642593214,
'atroce': 0.00016795431642593214,
'atrocité': 0.00016795431642593214,
'attache': 0.00016795431642593214,
'attacher': 0.00016795431642593214,
'attaché': 0.00016795431642593214,
'attachée': 0.00016795431642593214,
'attachés': 0.00016795431642593214,
'attale': 0.00016795431642593214,
'attaquant': 0.00016795431642593214,
'attaque': 0.00016795431642593214,
'attaquent': 0.00016795431642593214,
'attaquer': 0.00016795431642593214,
'attaqué': 0.00016795431642593214,
'atteindre': 0.00016795431642593214,
'atteint': 0.00016795431642593214,
'atteinte': 0.00016795431642593214,
'attend': 0.00016795431642593214,
'attendais': 0.00016795431642593214,
'attendait': 0.00016795431642593214,
'attendant': 0.00016795431642593214,
'attende': 0.00016795431642593214,
'attendent': 0.00016795431642593214,
'attendez': 0.00016795431642593214,
'attendons': 0.00016795431642593214,
'attendrai': 0.00016795431642593214,
'attendre': 0.00016795431642593214,
'attendri': 0.00016795431642593214,
'attendrie': 0.00016795431642593214,
'attendrir': 0.00016795431642593214,
'attendrit': 0.00016795431642593214,
'attends': 0.00016795431642593214,
'attendu': 0.00016795431642593214,
'attendus': 0.00016795431642593214,
'attentat': 0.00016795431642593214,
'attentats': 0.00016795431642593214,
'attente': 0.00016795431642593214,
'attenter': 0.00016795431642593214,
'attention': 0.00016795431642593214,
'attentive': 0.00016795431642593214,
'atteste': 0.00016795431642593214,
'attirait': 0.00016795431642593214,
'attire': 0.00016795431642593214,
'attirer': 0.00016795431642593214,
'attiré': 0.00016795431642593214,

'attiser': 0.00016795431642593214,
'attrait': 0.00016795431642593214,
'attribuait': 0.00016795431642593214,
'au': 0.00016795431642593214,
'aucun': 0.00016795431642593214,
'aucune': 0.00016795431642593214,
'aucuns': 0.00016795431642593214,
'audace': 0.00016795431642593214,
'audacieux': 0.00016795431642593214,
'augmente': 0.00016795431642593214,
'augmenter': 0.00016795431642593214,
'augure': 0.00016795431642593214,
'auguste': 0.00016795431642593214,
'augustes': 0.00016795431642593214,
'aujourd'hui': 0.00016795431642593214,
'auparavant': 0.00016795431642593214,
'auprès': 0.00016795431642593214,
'aura': 0.00016795431642593214,
'aurai': 0.00016795431642593214,
'auraient': 0.00016795431642593214,
'aurais': 0.00016795431642593214,
'aurait': 0.00016795431642593214,
'auras': 0.00016795431642593214,
'aurez': 0.00016795431642593214,
'auriez': 0.00016795431642593214,
'aurons': 0.00016795431642593214,
'auront': 0.00016795431642593214,
'aurore': 0.00016795431642593214,
'auspices': 0.00016795431642593214,
'aussi': 0.00016795431642593214,
'aussitôt': 0.00016795431642593214,
'austère': 0.00016795431642593214,
'autant': 0.00016795431642593214,
'autel': 0.00016795431642593214,
'autels': 0.00016795431642593214,
'auteur': 0.00016795431642593214,
'auteurs': 0.00016795431642593214,
'autorise': 0.00016795431642593214,
'autorité': 0.00016795431642593214,
'autour': 0.00016795431642593214,
'autre': 0.00016795431642593214,
'autrefois': 0.00016795431642593214,
'autrement': 0.00016795431642593214,
'autres': 0.00016795431642593214,
'autrui': 0.00016795431642593214,
'aux': 0.00016795431642593214,
'avaient': 0.00016795431642593214,
'avais': 0.00016795431642593214,
'avait': 0.00016795431642593214,
'avance': 0.00016795431642593214,
'avancer': 0.00016795431642593214,
'avancerais': 0.00016795431642593214,
'avancez': 0.00016795431642593214,
'avancé': 0.00016795431642593214,
'avancée': 0.00016795431642593214,
'avant': 0.00016795431642593214,
'avantage': 0.00016795431642593214,
'avançons': 0.00016795431642593214,
'avare': 0.00016795431642593214,
'avarice': 0.00016795431642593214,
'avec': 0.00016795431642593214,
'avecque': 0.00016795431642593214,
'avecques': 0.00016795431642593214,
'avenir': 0.00016795431642593214,
'aventure': 0.00016795431642593214,
'aventures': 0.00016795431642593214,
'averti': 0.00016795431642593214,
'avertie': 0.00016795431642593214,
'avertir': 0.00016795431642593214,

'aveu': 0.00016795431642593214,
'aveuglait': 0.00016795431642593214,
'aveugle': 0.00016795431642593214,
'aveuglement': 0.00016795431642593214,
'aveugler': 0.00016795431642593214,
'aveugles': 0.00016795431642593214,
'aveuglé': 0.00016795431642593214,
'avez': 0.00016795431642593214,
'avide': 0.00016795431642593214,
'aviez': 0.00016795431642593214,
'avions': 0.00016795431642593214,
'avis': 0.00016795431642593214,
'avoir': 0.00016795431642593214,
'avait': 0.00016795431642593214,
'avons': 0.00016795431642593214,
'avoue': 0.00016795431642593214,
'avouer': 0.00016795431642593214,
'avouerais': 0.00016795431642593214,
'avouez': 0.00016795431642593214,
'avoué': 0.00016795431642593214,
'ayant': 0.00016795431642593214,
'ayez': 0.00016795431642593214,
'aïeul': 0.00016795431642593214,
'aïeux': 0.00016795431642593214,
'babylone': 0.00016795431642593214,
'baigner': 0.00016795431642593214,
'baigné': 0.00016795431642593214,
'baiser': 0.00016795431642593214,
'baisers': 0.00016795431642593214,
'bajazet': 0.00016795431642593214,
'bal': 0.00016795431642593214,
'balance': 0.00016795431642593214,
'balancer': 0.00016795431642593214,
'balancé': 0.00016795431642593214,
'bandeau': 0.00016795431642593214,
'banni': 0.00016795431642593214,
'bannie': 0.00016795431642593214,
'bannir': 0.00016795431642593214,
'bannis': 0.00016795431642593214,
'bannit': 0.00016795431642593214,
'barbare': 0.00016795431642593214,
'barbares': 0.00016795431642593214,
'barbarie': 0.00016795431642593214,
'baron': 0.00016795431642593214,
'barsine': 0.00016795431642593214,
'barthélémi': 0.00016795431642593214,
'bas': 0.00016795431642593214,
'base': 0.00016795431642593214,
'basse': 0.00016795431642593214,
'bassesse': 0.00016795431642593214,
'bat': 0.00016795431642593214,
'bataille': 0.00016795431642593214,
'batailles': 0.00016795431642593214,
'batteux': 0.00016795431642593214,
'battu': 0.00016795431642593214,
'beau': 0.00016795431642593214,
'beaucoup': 0.00016795431642593214,
'beauté': 0.00016795431642593214,
'beautés': 0.00016795431642593214,
'beauval': 0.00016795431642593214,
'beaux': 0.00016795431642593214,
'bel': 0.00016795431642593214,
'belle': 0.00016795431642593214,
'belles': 0.00016795431642593214,
'bengale': 0.00016795431642593214,
'berceau': 0.00016795431642593214,
'bernardht': 0.00016795431642593214,
'besoin': 0.00016795431642593214,
'bible': 0.00016795431642593214,

'bien': 0.00016795431642593214,
'bienfait': 0.00016795431642593214,
'bienfaits': 0.00016795431642593214,
'bienheureux': 0.00016795431642593214,
'biens': 0.00016795431642593214,
'bientôt': 0.00016795431642593214,
'bienveillance': 0.00016795431642593214,
'billet': 0.00016795431642593214,
'bine': 0.00016795431642593214,
'bizarre': 0.00016795431642593214,
'blanchi': 0.00016795431642593214,
'blasphème': 0.00016795431642593214,
'blasphémé': 0.00016795431642593214,
'blesse': 0.00016795431642593214,
'blesser': 0.00016795431642593214,
'blessure': 0.00016795431642593214,
'blessures': 0.00016795431642593214,
'blessé': 0.00016795431642593214,
'blessée': 0.00016795431642593214,
'blonde': 0.00016795431642593214,
'blonds': 0.00016795431642593214,
'blâme': 0.00016795431642593214,
'blâmer': 0.00016795431642593214,
'blême': 0.00016795431642593214,
'boileau': 0.00016795431642593214,
'bois': 0.00016795431642593214,
'bon': 0.00016795431642593214,
'bonheur': 0.00016795431642593214,
'bonne': 0.00016795431642593214,
'bonnes': 0.00016795431642593214,
'bons': 0.00016795431642593214,
'bonté': 0.00016795431642593214,
'bontés': 0.00016795431642593214,
'bord': 0.00016795431642593214,
'bords': 0.00016795431642593214,
'bornant': 0.00016795431642593214,
'borne': 0.00016795431642593214,
'borner': 0.00016795431642593214,
'bornera': 0.00016795431642593214,
'bornes': 0.00016795431642593214,
'bornée': 0.00016795431642593214,
'bouche': 0.00016795431642593214,
'bouclier': 0.00016795431642593214,
'bouillon': 0.00016795431642593214,
'bourgogne': 0.00016795431642593214,
'bourreau': 0.00016795431642593214,
'bourreaux': 0.00016795431642593214,
'bout': 0.00016795431642593214,
'boyer': 0.00016795431642593214,
'brama': 0.00016795431642593214,
'bramine': 0.00016795431642593214,
'bras': 0.00016795431642593214,
'brasier': 0.00016795431642593214,
'brave': 0.00016795431642593214,
'braver': 0.00016795431642593214,
'braves': 0.00016795431642593214,
'bref': 0.00016795431642593214,
'breuvage': 0.00016795431642593214,
'brigands': 0.00016795431642593214,
'brigue': 0.00016795431642593214,
'brigues': 0.00016795431642593214,
'brigué': 0.00016795431642593214,
'brillant': 0.00016795431642593214,
'brillante': 0.00016795431642593214,
'brille': 0.00016795431642593214,
'briller': 0.00016795431642593214,
'brise': 0.00016795431642593214,
'briser': 0.00016795431642593214,
'brisé': 0.00016795431642593214,

'britannicus': 0.00016795431642593214,
'brouilla': 0.00016795431642593214,
'bruit': 0.00016795431642593214,
'bruits': 0.00016795431642593214,
'brutale': 0.00016795431642593214,
'brutalité': 0.00016795431642593214,
'brute': 0.00016795431642593214,
'brûlait': 0.00016795431642593214,
'brûlant': 0.00016795431642593214,
'brûlants': 0.00016795431642593214,
'brûle': 0.00016795431642593214,
'brûler': 0.00016795431642593214,
'brûlez': 0.00016795431642593214,
'brûlé': 0.00016795431642593214,
'but': 0.00016795431642593214,
'butin': 0.00016795431642593214,
'butte': 0.00016795431642593214,
'byzance': 0.00016795431642593214,
'bâtis': 0.00016795431642593214,
'bénir': 0.00016795431642593214,
'bénis': 0.00016795431642593214,
'bérénice': 0.00016795431642593214,
'bête': 0.00016795431642593214,
'bêtes': 0.00016795431642593214,
'bûcher': 0.00016795431642593214,
'bûchers': 0.00016795431642593214,
'cabale': 0.00016795431642593214,
'cachais': 0.00016795431642593214,
'cachant': 0.00016795431642593214,
'cache': 0.00016795431642593214,
'cachent': 0.00016795431642593214,
'cacher': 0.00016795431642593214,
'cachet': 0.00016795431642593214,
'cachez': 0.00016795431642593214,
'caché': 0.00016795431642593214,
'cachée': 0.00016795431642593214,
'cadavre': 0.00016795431642593214,
'caliodore': 0.00016795431642593214,
'calme': 0.00016795431642593214,
'calmer': 0.00016795431642593214,
'calmé': 0.00016795431642593214,
'calomnie': 0.00016795431642593214,
'camp': 0.00016795431642593214,
'campagne': 0.00016795431642593214,
'campagnes': 0.00016795431642593214,
'camps': 0.00016795431642593214,
'candeur': 0.00016795431642593214,
'capable': 0.00016795431642593214,
'capables': 0.00016795431642593214,
'capitaines': 0.00016795431642593214,
'caprice': 0.00016795431642593214,
'caprices': 0.00016795431642593214,
'captif': 0.00016795431642593214,
'captifs': 0.00016795431642593214,
'captive': 0.00016795431642593214,
'captivité': 0.00016795431642593214,
'car': 0.00016795431642593214,
'caractère': 0.00016795431642593214,
'caractères': 0.00016795431642593214,
'caresser': 0.00016795431642593214,
'caresses': 0.00016795431642593214,
'carnage': 0.00016795431642593214,
'carrière': 0.00016795431642593214,
'carrés': 0.00016795431642593214,
'carthage': 0.00016795431642593214,
'cas': 0.00016795431642593214,
'cause': 0.00016795431642593214,
'causer': 0.00016795431642593214,
'causes': 0.00016795431642593214,

'cause': 0.00016795431642593214,
'caverne': 0.00016795431642593214,
'ce': 0.00016795431642593214,
'ceci': 0.00016795431642593214,
'cela': 0.00016795431642593214,
'celer': 0.00016795431642593214,
'celle': 0.00016795431642593214,
'celles': 0.00016795431642593214,
'celui': 0.00016795431642593214,
'cendre': 0.00016795431642593214,
'censeur': 0.00016795431642593214,
'censure': 0.00016795431642593214,
'cent': 0.00016795431642593214,
'cents': 0.00016795431642593214,
'cependant': 0.00016795431642593214,
'cercueil': 0.00016795431642593214,
'cerfs': 0.00016795431642593214,
'certain': 0.00016795431642593214,
'certaine': 0.00016795431642593214,
'certainement': 0.00016795431642593214,
'certaines': 0.00016795431642593214,
'certains': 0.00016795431642593214,
'certes': 0.00016795431642593214,
'ces': 0.00016795431642593214,
'cessais': 0.00016795431642593214,
'cesse': 0.00016795431642593214,
'cesser': 0.00016795431642593214,
'cessez': 0.00016795431642593214,
'cet': 0.00016795431642593214,
'cette': 0.00016795431642593214,
'ceux': 0.00016795431642593214,
'chacun': 0.00016795431642593214,
'chagrin': 0.00016795431642593214,
'chagrins': 0.00016795431642593214,
'chaleur': 0.00016795431642593214,
'chambre': 0.00016795431642593214,
'chameaux': 0.00016795431642593214,
'champ': 0.00016795431642593214,
'champmeslé': 0.00016795431642593214,
'champs': 0.00016795431642593214,
'chancelant': 0.00016795431642593214,
'change': 0.00016795431642593214,
'changeant': 0.00016795431642593214,
'changement': 0.00016795431642593214,
'changer': 0.00016795431642593214,
'changez': 0.00016795431642593214,
'changé': 0.00016795431642593214,
'changée': 0.00016795431642593214,
'chant': 0.00016795431642593214,
'chaque': 0.00016795431642593214,
'char': 0.00016795431642593214,
'charge': 0.00016795431642593214,
'chargea': 0.00016795431642593214,
'charger': 0.00016795431642593214,
'chargez': 0.00016795431642593214,
'chargé': 0.00016795431642593214,
'chargée': 0.00016795431642593214,
'chargés': 0.00016795431642593214,
'charmant': 0.00016795431642593214,
'charmante': 0.00016795431642593214,
'charmants': 0.00016795431642593214,
'charme': 0.00016795431642593214,
'charmer': 0.00016795431642593214,
'charmes': 0.00016795431642593214,
'charmé': 0.00016795431642593214,
'chasse': 0.00016795431642593214,
'chasser': 0.00016795431642593214,
'chasseur': 0.00016795431642593214,
'chassez': 0.00016795431642593214,

'chassé': 0.00016795431642593214,
'chaste': 0.00016795431642593214,
'chastes': 0.00016795431642593214,
'chasteté': 0.00016795431642593214,
'chaîne': 0.00016795431642593214,
'chaînes': 0.00016795431642593214,
'chef': 0.00016795431642593214,
'chefs': 0.00016795431642593214,
'chemin': 0.00016795431642593214,
'chemins': 0.00016795431642593214,
'cher': 0.00016795431642593214,
'cherchais': 0.00016795431642593214,
'cherchait': 0.00016795431642593214,
'cherchant': 0.00016795431642593214,
'cherche': 0.00016795431642593214,
'cherchent': 0.00016795431642593214,
'chercher': 0.00016795431642593214,
'cherchera': 0.00016795431642593214,
'cherchez': 0.00016795431642593214,
'cherchons': 0.00016795431642593214,
'cherché': 0.00016795431642593214,
'chers': 0.00016795431642593214,
'chevaux': 0.00016795431642593214,
'cheveux': 0.00016795431642593214,
'chez': 0.00016795431642593214,
'chimère': 0.00016795431642593214,
'chimères': 0.00016795431642593214,
'chimérique': 0.00016795431642593214,
'choc': 0.00016795431642593214,
'choeur': 0.00016795431642593214,
'choir': 0.00016795431642593214,
'choisi': 0.00016795431642593214,
'choisir': 0.00016795431642593214,
'choisis': 0.00016795431642593214,
'choisissons': 0.00016795431642593214,
'choix': 0.00016795431642593214,
'choquer': 0.00016795431642593214,
'choqué': 0.00016795431642593214,
'chose': 0.00016795431642593214,
'choses': 0.00016795431642593214,
'chute': 0.00016795431642593214,
'châtier': 0.00016795431642593214,
'châtiment': 0.00016795431642593214,
'châtiments': 0.00016795431642593214,
'chère': 0.00016795431642593214,
'chères': 0.00016795431642593214,
'chéri': 0.00016795431642593214,
'chérie': 0.00016795431642593214,
'chérir': 0.00016795431642593214,
'chérisse': 0.00016795431642593214,
'chérissons': 0.00016795431642593214,
'chérit': 0.00016795431642593214,
'ci': 0.00016795431642593214,
'ciel': 0.00016795431642593214,
'cieux': 0.00016795431642593214,
'cimetière': 0.00016795431642593214,
'cinq': 0.00016795431642593214,
'cinquième': 0.00016795431642593214,
'circonstance': 0.00016795431642593214,
'circonstances': 0.00016795431642593214,
'cirte': 0.00016795431642593214,
'citer': 0.00016795431642593214,
'citoyen': 0.00016795431642593214,
'citoyens': 0.00016795431642593214,
'cité': 0.00016795431642593214,
'cités': 0.00016795431642593214,
'civile': 0.00016795431642593214,
'clair': 0.00016795431642593214,
'clairement': 0.00016795431642593214,

```

'clairon': 0.00016795431642593214,
'clairs': 0.00016795431642593214,
'clameurs': 0.00016795431642593214,
'clarté': 0.00016795431642593214,
'clartés': 0.00016795431642593214,
'climat': 0.00016795431642593214,
'climats': 0.00016795431642593214,
'clos': 0.00016795431642593214,
'close': 0.00016795431642593214,
'clytemnestre': 0.00016795431642593214,
'clémence': 0.00016795431642593214,
'cléments': 0.00016795431642593214,
'cléopâtre': 0.00016795431642593214,
'coeur': 0.00016795431642593214,
'coeurs': 0.00016795431642593214,
'cohortes': 0.00016795431642593214,
'collaborateur': 0.00016795431642593214,
'collatin': 0.00016795431642593214,
'collège': 0.00016795431642593214,
'colère': 0.00016795431642593214,
'combat': 0.00016795431642593214,
'combats': 0.00016795431642593214,
'combattants': 0.00016795431642593214,
'combatte': 0.00016795431642593214,
'combattent': 0.00016795431642593214,
'combattit': 0.00016795431642593214,
'combattre': 0.00016795431642593214,
'combattu': 0.00016795431642593214,
'combattue': 0.00016795431642593214,
'combien': 0.00016795431642593214,
'comble': 0.00016795431642593214,
'comblé': 0.00016795431642593214,
'comique': 0.00016795431642593214,
'commande': 0.00016795431642593214,
'commandement': 0.00016795431642593214,
'commandements': 0.00016795431642593214,
'commander': 0.00016795431642593214,
'commandez': 0.00016795431642593214,
'commandées': 0.00016795431642593214,
'comme': 0.00016795431642593214,
'commence': 0.00016795431642593214,
'commencement': 0.00016795431642593214,
'commencer': 0.00016795431642593214,
'commencez': 0.00016795431642593214,
'commencé': 0.00016795431642593214,
'commencée': 0.00016795431642593214,
'comment': 0.00016795431642593214,
'commençant': 0.00016795431642593214,
'commerce': 0.00016795431642593214,
'commet': 0.00016795431642593214,
'commettre': 0.00016795431642593214,
'commis': 0.00016795431642593214,
'commun': 0.00016795431642593214,
'communauté': 0.00016795431642593214,
'commune': 0.00016795431642593214,
'communes': 0.00016795431642593214,
'communs': 0.00016795431642593214,
'compagne': 0.00016795431642593214,
'compagnons': 0.00016795431642593214,
'comparable': 0.00016795431642593214,
'comparaison': 0.00016795431642593214,
'comparer': 0.00016795431642593214,
'comparée': 0.00016795431642593214,
...}

```

In [47]:

```

stemmed_vocabulary_com_rand = [fr.stem(word) for word in vocabulary_com_rand]
counts = dict()
for word in vocabulary_com_rand:

```

```

    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_com_rand = counts
for key, value in counts.items():
    tf_dict_com_rand[key] = value / len(vocabulary_com_rand)
tf_dict_com_rand

```

```

Out[47]: {'a': 0.0002098635886673662,
'abaisse': 0.0002098635886673662,
'abandonne': 0.0002098635886673662,
'abandonner': 0.0002098635886673662,
'abattre': 0.0002098635886673662,
'abattu': 0.0002098635886673662,
'abattue': 0.0002098635886673662,
'abhorre': 0.0002098635886673662,
'abondance': 0.0002098635886673662,
'abord': 0.0002098635886673662,
'aborder': 0.0002098635886673662,
'aboumékàrès': 0.0002098635886673662,
'abri': 0.0002098635886673662,
'absence': 0.0002098635886673662,
'absent': 0.0002098635886673662,
'absente': 0.0002098635886673662,
'absenter': 0.0002098635886673662,
'absents': 0.0002098635886673662,
'absolu': 0.0002098635886673662,
'absolument': 0.0002098635886673662,
'abuse': 0.0002098635886673662,
'abuser': 0.0002098635886673662,
'abuserais': 0.0002098635886673662,
'académie': 0.0002098635886673662,
'accabler': 0.0002098635886673662,
'accablée': 0.0002098635886673662,
'accepter': 0.0002098635886673662,
'accepterai': 0.0002098635886673662,
'acceptez': 0.0002098635886673662,
'accident': 0.0002098635886673662,
'accidents': 0.0002098635886673662,
'accommoder': 0.0002098635886673662,
'accommoderai': 0.0002098635886673662,
'accompagne': 0.0002098635886673662,
'accompagner': 0.0002098635886673662,
'accomplie': 0.0002098635886673662,
'accomplir': 0.0002098635886673662,
'accomplissons': 0.0002098635886673662,
'accord': 0.0002098635886673662,
'accorde': 0.0002098635886673662,
'accorder': 0.0002098635886673662,
'accordez': 0.0002098635886673662,
'accords': 0.0002098635886673662,
'accroire': 0.0002098635886673662,
'accuse': 0.0002098635886673662,
'accuser': 0.0002098635886673662,
'accusez': 0.0002098635886673662,
'accès': 0.0002098635886673662,
'achevais': 0.0002098635886673662,
'achever': 0.0002098635886673662,
'achevez': 0.0002098635886673662,
'achevé': 0.0002098635886673662,
'achevée': 0.0002098635886673662,
'achille': 0.0002098635886673662,
'achève': 0.0002098635886673662,
'acquis': 0.0002098635886673662,
'acquise': 0.0002098635886673662,
'acquitter': 0.0002098635886673662,

```

'acquérir': 0.0002098635886673662,
'acte': 0.0002098635886673662,
'acteur': 0.0002098635886673662,
'acteurs': 0.0002098635886673662,
'action': 0.0002098635886673662,
'actions': 0.0002098635886673662,
'actrice': 0.0002098635886673662,
'actuellement': 0.0002098635886673662,
'ad': 0.0002098635886673662,
'adieu': 0.0002098635886673662,
'admirable': 0.0002098635886673662,
'admirait': 0.0002098635886673662,
'admirateur': 0.0002098635886673662,
'admire': 0.0002098635886673662,
'admirer': 0.0002098635886673662,
'adorable': 0.0002098635886673662,
'adrateur': 0.0002098635886673662,
'adore': 0.0002098635886673662,
'adorer': 0.0002098635886673662,
'adorée': 0.0002098635886673662,
'adoucit': 0.0002098635886673662,
'adresse': 0.0002098635886673662,
'adroit': 0.0002098635886673662,
'adroitement': 0.0002098635886673662,
'advenir': 0.0002098635886673662,
'affaire': 0.0002098635886673662,
'affaires': 0.0002098635886673662,
'affectent': 0.0002098635886673662,
'affection': 0.0002098635886673662,
'afflige': 0.0002098635886673662,
'affligée': 0.0002098635886673662,
'affreuse': 0.0002098635886673662,
'affreux': 0.0002098635886673662,
'affront': 0.0002098635886673662,
'affronteur': 0.0002098635886673662,
'affronts': 0.0002098635886673662,
'afin': 0.0002098635886673662,
'age': 0.0002098635886673662,
'agir': 0.0002098635886673662,
'agit': 0.0002098635886673662,
'agité': 0.0002098635886673662,
'agitée': 0.0002098635886673662,
'agréable': 0.0002098635886673662,
'agréablement': 0.0002098635886673662,
'agréables': 0.0002098635886673662,
'agréé': 0.0002098635886673662,
'agréer': 0.0002098635886673662,
'agréez': 0.0002098635886673662,
'agrément': 0.0002098635886673662,
'agréments': 0.0002098635886673662,
'ah': 0.0002098635886673662,
'ahi': 0.0002098635886673662,
'ahy': 0.0002098635886673662,
'ai': 0.0002098635886673662,
'aide': 0.0002098635886673662,
'aider': 0.0002098635886673662,
'aie': 0.0002098635886673662,
'aies': 0.0002098635886673662,
'aile': 0.0002098635886673662,
'ailes': 0.0002098635886673662,
'aille': 0.0002098635886673662,
'ailleurs': 0.0002098635886673662,
'aimable': 0.0002098635886673662,
'aimables': 0.0002098635886673662,
'aimai': 0.0002098635886673662,
'aimais': 0.0002098635886673662,
'aimant': 0.0002098635886673662,
'aime': 0.0002098635886673662,
'aient': 0.0002098635886673662,

'aimer': 0.0002098635886673662,
'aimerais': 0.0002098635886673662,
'aimerais': 0.0002098635886673662,
'aimerez': 0.0002098635886673662,
'aimez': 0.0002098635886673662,
'aimiez': 0.0002098635886673662,
'aimons': 0.0002098635886673662,
'aimé': 0.0002098635886673662,
'aimée': 0.0002098635886673662,
'aimés': 0.0002098635886673662,
'ainsi': 0.0002098635886673662,
'air': 0.0002098635886673662,
'airs': 0.0002098635886673662,
'aise': 0.0002098635886673662,
'aisé': 0.0002098635886673662,
'aisée': 0.0002098635886673662,
'aisément': 0.0002098635886673662,
'ait': 0.0002098635886673662,
'ajoute': 0.0002098635886673662,
'ajustement': 0.0002098635886673662,
'alcidon': 0.0002098635886673662,
'alentour': 0.0002098635886673662,
'alerte': 0.0002098635886673662,
'alexandre': 0.0002098635886673662,
'alla': 0.0002098635886673662,
'allaient': 0.0002098635886673662,
'allais': 0.0002098635886673662,
'allait': 0.0002098635886673662,
'allant': 0.0002098635886673662,
'aller': 0.0002098635886673662,
'allez': 0.0002098635886673662,
'alliance': 0.0002098635886673662,
'allie': 0.0002098635886673662,
'allier': 0.0002098635886673662,
'allons': 0.0002098635886673662,
'allègue': 0.0002098635886673662,
'allé': 0.0002098635886673662,
'allégresse': 0.0002098635886673662,
'alléguer': 0.0002098635886673662,
'allégé': 0.0002098635886673662,
'almanach': 0.0002098635886673662,
'almanzor': 0.0002098635886673662,
'alors': 0.0002098635886673662,
'alzire': 0.0002098635886673662,
'amant': 0.0002098635886673662,
'amante': 0.0002098635886673662,
'amants': 0.0002098635886673662,
'amas': 0.0002098635886673662,
'amazone': 0.0002098635886673662,
'ambitieux': 0.0002098635886673662,
'amende': 0.0002098635886673662,
'amener': 0.0002098635886673662,
'ami': 0.0002098635886673662,
'amidor': 0.0002098635886673662,
'amie': 0.0002098635886673662,
'amies': 0.0002098635886673662,
'amis': 0.0002098635886673662,
'amitié': 0.0002098635886673662,
'amour': 0.0002098635886673662,
'amourettes': 0.0002098635886673662,
'amoureuse': 0.0002098635886673662,
'amoureux': 0.0002098635886673662,
'amours': 0.0002098635886673662,
'ampoulés': 0.0002098635886673662,
'amuse': 0.0002098635886673662,
'amusent': 0.0002098635886673662,
'amuser': 0.0002098635886673662,
'amusons': 0.0002098635886673662,
'amène': 0.0002098635886673662,

'an': 0.0002098635886673662,
'ange': 0.0002098635886673662,
'animal': 0.0002098635886673662,
'anime': 0.0002098635886673662,
'années': 0.0002098635886673662,
'ans': 0.0002098635886673662,
'antiques': 0.0002098635886673662,
'antiquité': 0.0002098635886673662,
'antistrophe': 0.0002098635886673662,
'antoine': 0.0002098635886673662,
'antre': 0.0002098635886673662,
'apaisé': 0.0002098635886673662,
'apercevoir': 0.0002098635886673662,
'aperçois': 0.0002098635886673662,
'aperçoit': 0.0002098635886673662,
'aperçus': 0.0002098635886673662,
'apocryphe': 0.0002098635886673662,
'apollon': 0.0002098635886673662,
'apparemment': 0.0002098635886673662,
'apparence': 0.0002098635886673662,
'apparences': 0.0002098635886673662,
'appartement': 0.0002098635886673662,
'appartient': 0.0002098635886673662,
'appas': 0.0002098635886673662,
'appeler': 0.0002098635886673662,
'appelez': 0.0002098635886673662,
'appellations': 0.0002098635886673662,
'appelle': 0.0002098635886673662,
'applaudir': 0.0002098635886673662,
'apporte': 0.0002098635886673662,
'apporter': 0.0002098635886673662,
'apprend': 0.0002098635886673662,
'apprendra': 0.0002098635886673662,
'apprendre': 0.0002098635886673662,
'apprends': 0.0002098635886673662,
'apprenez': 0.0002098635886673662,
'appris': 0.0002098635886673662,
'apprise': 0.0002098635886673662,
'apprivoiser': 0.0002098635886673662,
'approchant': 0.0002098635886673662,
'approche': 0.0002098635886673662,
'approcher': 0.0002098635886673662,
'approchez': 0.0002098635886673662,
'approuve': 0.0002098635886673662,
'approuvera': 0.0002098635886673662,
'approuvez': 0.0002098635886673662,
'appréhende': 0.0002098635886673662,
'appréhendez': 0.0002098635886673662,
'apprête': 0.0002098635886673662,
'apprêts': 0.0002098635886673662,
'apprêté': 0.0002098635886673662,
'appui': 0.0002098635886673662,
'appétit': 0.0002098635886673662,
'après': 0.0002098635886673662,
'arbres': 0.0002098635886673662,
'arc': 0.0002098635886673662,
'ardent': 0.0002098635886673662,
'ardente': 0.0002098635886673662,
'ardeur': 0.0002098635886673662,
'ardeurs': 0.0002098635886673662,
'argent': 0.0002098635886673662,
'argine': 0.0002098635886673662,
'arlequin': 0.0002098635886673662,
'arme': 0.0002098635886673662,
'armes': 0.0002098635886673662,
'armide': 0.0002098635886673662,
'armé': 0.0002098635886673662,
'armée': 0.0002098635886673662,
'armées': 0.0002098635886673662,

'arracher': 0.0002098635886673662,
'arrivant': 0.0002098635886673662,
'arrive': 0.0002098635886673662,
'arriver': 0.0002098635886673662,
'arrivera': 0.0002098635886673662,
'arrivé': 0.0002098635886673662,
'arrose': 0.0002098635886673662,
'arrêt': 0.0002098635886673662,
'arrête': 0.0002098635886673662,
'arrêter': 0.0002098635886673662,
'arrêteraient': 0.0002098635886673662,
'arrêtez': 0.0002098635886673662,
'arrêtons': 0.0002098635886673662,
'arrêté': 0.0002098635886673662,
'art': 0.0002098635886673662,
'artabaze': 0.0002098635886673662,
'arts': 0.0002098635886673662,
'as': 0.0002098635886673662,
'asie': 0.0002098635886673662,
'asile': 0.0002098635886673662,
'aspect': 0.0002098635886673662,
'aspire': 0.0002098635886673662,
'assaisonne': 0.0002098635886673662,
'assassine': 0.0002098635886673662,
'assassiner': 0.0002098635886673662,
'assassins': 0.0002098635886673662,
'assaut': 0.0002098635886673662,
'assemblez': 0.0002098635886673662,
'assemblée': 0.0002098635886673662,
'assemblés': 0.0002098635886673662,
'asseoir': 0.0002098635886673662,
'asservir': 0.0002098635886673662,
'asseyons': 0.0002098635886673662,
'assez': 0.0002098635886673662,
'assidu': 0.0002098635886673662,
'assiette': 0.0002098635886673662,
'assiégé': 0.0002098635886673662,
'assomme': 0.0002098635886673662,
'assommer': 0.0002098635886673662,
'assoupi': 0.0002098635886673662,
'assouvi': 0.0002098635886673662,
'assurance': 0.0002098635886673662,
'assure': 0.0002098635886673662,
'assurer': 0.0002098635886673662,
'assurément': 0.0002098635886673662,
'astre': 0.0002098635886673662,
'astrolabe': 0.0002098635886673662,
'astrologue': 0.0002098635886673662,
'athène': 0.0002098635886673662,
'atropos': 0.0002098635886673662,
'attache': 0.0002098635886673662,
'attaché': 0.0002098635886673662,
'attachée': 0.0002098635886673662,
'attaque': 0.0002098635886673662,
'attaquer': 0.0002098635886673662,
'atteinte': 0.0002098635886673662,
'attend': 0.0002098635886673662,
'attendais': 0.0002098635886673662,
'attendait': 0.0002098635886673662,
'attendant': 0.0002098635886673662,
'attendent': 0.0002098635886673662,
'attendez': 0.0002098635886673662,
'attendions': 0.0002098635886673662,
'attendre': 0.0002098635886673662,
'attendris': 0.0002098635886673662,
'attendrissez': 0.0002098635886673662,
'attends': 0.0002098635886673662,
'attente': 0.0002098635886673662,
'attenter': 0.0002098635886673662,

'attention': 0.0002098635886673662,
'atteste': 0.0002098635886673662,
'attirail': 0.0002098635886673662,
'attire': 0.0002098635886673662,
'attraits': 0.0002098635886673662,
'attraper': 0.0002098635886673662,
'attrapée': 0.0002098635886673662,
'attribuer': 0.0002098635886673662,
'attributs': 0.0002098635886673662,
'au': 0.0002098635886673662,
'aucun': 0.0002098635886673662,
'aucune': 0.0002098635886673662,
'aucunement': 0.0002098635886673662,
'audace': 0.0002098635886673662,
'audacieux': 0.0002098635886673662,
'audit': 0.0002098635886673662,
'auditeurs': 0.0002098635886673662,
'augmenter': 0.0002098635886673662,
'augure': 0.0002098635886673662,
'auguste': 0.0002098635886673662,
'aujordy': 0.0002098635886673662,
'aujourd': 0.0002098635886673662,
'aujourd'hui': 0.0002098635886673662,
'auparavant': 0.0002098635886673662,
'auprès': 0.0002098635886673662,
'aura': 0.0002098635886673662,
'aurai': 0.0002098635886673662,
'auraient': 0.0002098635886673662,
'aurais': 0.0002098635886673662,
'aurait': 0.0002098635886673662,
'auras': 0.0002098635886673662,
'aurez': 0.0002098635886673662,
'auriez': 0.0002098635886673662,
'aurions': 0.0002098635886673662,
'aurons': 0.0002098635886673662,
'auront': 0.0002098635886673662,
'aurore': 0.0002098635886673662,
'aussi': 0.0002098635886673662,
'aussitôt': 0.0002098635886673662,
'austère': 0.0002098635886673662,
'autant': 0.0002098635886673662,
'autel': 0.0002098635886673662,
'auteur': 0.0002098635886673662,
'auteurs': 0.0002098635886673662,
'automne': 0.0002098635886673662,
'autoriser': 0.0002098635886673662,
'autorité': 0.0002098635886673662,
'autour': 0.0002098635886673662,
'autre': 0.0002098635886673662,
'autrefois': 0.0002098635886673662,
'autrement': 0.0002098635886673662,
'autres': 0.0002098635886673662,
'autrui': 0.0002098635886673662,
'aux': 0.0002098635886673662,
'avaient': 0.0002098635886673662,
'avais': 0.0002098635886673662,
'avait': 0.0002098635886673662,
'avance': 0.0002098635886673662,
'avancez': 0.0002098635886673662,
'avant': 0.0002098635886673662,
'avantage': 0.0002098635886673662,
'avantageux': 0.0002098635886673662,
'avare': 0.0002098635886673662,
'avec': 0.0002098635886673662,
'avecque': 0.0002098635886673662,
'avenir': 0.0002098635886673662,
'aventure': 0.0002098635886673662,
'aventures': 0.0002098635886673662,
'avenue': 0.0002098635886673662,

'averti': 0.0002098635886673662,
'avertie': 0.0002098635886673662,
'avertir': 0.0002098635886673662,
'avertis': 0.0002098635886673662,
'aveu': 0.0002098635886673662,
'avez': 0.0002098635886673662,
'avidité': 0.0002098635886673662,
'aviez': 0.0002098635886673662,
'avons': 0.0002098635886673662,
'avis': 0.0002098635886673662,
'avisé': 0.0002098635886673662,
'avisée': 0.0002098635886673662,
'avisés': 0.0002098635886673662,
'avocat': 0.0002098635886673662,
'avoir': 0.0002098635886673662,
'avons': 0.0002098635886673662,
'avoue': 0.0002098635886673662,
'avouer': 0.0002098635886673662,
'avouerez': 0.0002098635886673662,
'ayant': 0.0002098635886673662,
'ayez': 0.0002098635886673662,
'ayons': 0.0002098635886673662,
'azur': 0.0002098635886673662,
'aîné': 0.0002098635886673662,
'aînés': 0.0002098635886673662,
'aïeux': 0.0002098635886673662,
'babillard': 0.0002098635886673662,
'bachique': 0.0002098635886673662,
'badinage': 0.0002098635886673662,
'bagage': 0.0002098635886673662,
'bagatelle': 0.0002098635886673662,
'bagatelles': 0.0002098635886673662,
'bagdad': 0.0002098635886673662,
'baguenaudière': 0.0002098635886673662,
'bah': 0.0002098635886673662,
'baignez': 0.0002098635886673662,
'baise': 0.0002098635886673662,
'baiser': 0.0002098635886673662,
'baisse': 0.0002098635886673662,
'bal': 0.0002098635886673662,
'balachou': 0.0002098635886673662,
'balada': 0.0002098635886673662,
'balance': 0.0002098635886673662,
'balancer': 0.0002098635886673662,
'baleine': 0.0002098635886673662,
'balle': 0.0002098635886673662,
'bande': 0.0002098635886673662,
'bandeau': 0.0002098635886673662,
'bannir': 0.0002098635886673662,
'bannis': 0.0002098635886673662,
'bannissant': 0.0002098635886673662,
'bannissons': 0.0002098635886673662,
'baragouin': 0.0002098635886673662,
'barbarie': 0.0002098635886673662,
'barbe': 0.0002098635886673662,
'barbier': 0.0002098635886673662,
'barcelone': 0.0002098635886673662,
'bas': 0.0002098635886673662,
'basses': 0.0002098635886673662,
'bastonnade': 0.0002098635886673662,
'bat': 0.0002098635886673662,
'bataille': 0.0002098635886673662,
'batailles': 0.0002098635886673662,
'battrai': 0.0002098635886673662,
'battre': 0.0002098635886673662,
'battu': 0.0002098635886673662,
'bavarde': 0.0002098635886673662,
'bavolet': 0.0002098635886673662,
'beau': 0.0002098635886673662,

'beaucoup': 0.0002098635886673662,
'beauté': 0.0002098635886673662,
'beautés': 0.0002098635886673662,
'beaux': 0.0002098635886673662,
'bec': 0.0002098635886673662,
'bel': 0.0002098635886673662,
'belle': 0.0002098635886673662,
'belles': 0.0002098635886673662,
'ben': 0.0002098635886673662,
'benêt': 0.0002098635886673662,
'berceau': 0.0002098635886673662,
'berger': 0.0002098635886673662,
'bergère': 0.0002098635886673662,
'besogne': 0.0002098635886673662,
'besoin': 0.0002098635886673662,
'biaise': 0.0002098635886673662,
'bian': 0.0002098635886673662,
'bibliothèque': 0.0002098635886673662,
'bien': 0.0002098635886673662,
'bienfaiteur': 0.0002098635886673662,
'bienheureuse': 0.0002098635886673662,
'bienheureux': 0.0002098635886673662,
'biens': 0.0002098635886673662,
'bientôt': 0.0002098635886673662,
'billet': 0.0002098635886673662,
'billets': 0.0002098635886673662,
'bizarre': 0.0002098635886673662,
'bière': 0.0002098635886673662,
'blaise': 0.0002098635886673662,
'blanc': 0.0002098635886673662,
'blancs': 0.0002098635886673662,
'blessure': 0.0002098635886673662,
'blessé': 0.0002098635886673662,
'blessée': 0.0002098635886673662,
'blâme': 0.0002098635886673662,
'blâmer': 0.0002098635886673662,
'blâmez': 0.0002098635886673662,
'blâmée': 0.0002098635886673662,
'boeufs': 0.0002098635886673662,
'boire': 0.0002098635886673662,
'bois': 0.0002098635886673662,
'bon': 0.0002098635886673662,
'bonheur': 0.0002098635886673662,
'bonhomme': 0.0002098635886673662,
'bonjour': 0.0002098635886673662,
'bonne': 0.0002098635886673662,
'bonnement': 0.0002098635886673662,
'bonnes': 0.0002098635886673662,
'bons': 0.0002098635886673662,
'bonsoir': 0.0002098635886673662,
'bonté': 0.0002098635886673662,
'bontés': 0.0002098635886673662,
'bord': 0.0002098635886673662,
'bords': 0.0002098635886673662,
'bornes': 0.0002098635886673662,
'bouche': 0.0002098635886673662,
'bouchon': 0.0002098635886673662,
'boudoir': 0.0002098635886673662,
'boue': 0.0002098635886673662,
'bouge': 0.0002098635886673662,
'bouillant': 0.0002098635886673662,
'bourg': 0.0002098635886673662,
'bourgeois': 0.0002098635886673662,
'bourgeoise': 0.0002098635886673662,
'bourreau': 0.0002098635886673662,
'bourse': 0.0002098635886673662,
'bout': 0.0002098635886673662,
'bouteille': 0.0002098635886673662,
'boutique': 0.0002098635886673662,

'bouvillon': 0.0002098635886673662,
'boîte': 0.0002098635886673662,
'branche': 0.0002098635886673662,
'bras': 0.0002098635886673662,
'brassiez': 0.0002098635886673662,
'bravant': 0.0002098635886673662,
'brave': 0.0002098635886673662,
'braver': 0.0002098635886673662,
'bravez': 0.0002098635886673662,
'bravoure': 0.0002098635886673662,
'bravé': 0.0002098635886673662,
'bride': 0.0002098635886673662,
'briffe': 0.0002098635886673662,
'brillant': 0.0002098635886673662,
'brillante': 0.0002098635886673662,
'briller': 0.0002098635886673662,
'brise': 0.0002098635886673662,
'briser': 0.0002098635886673662,
'brisé': 0.0002098635886673662,
'bronze': 0.0002098635886673662,
'bruit': 0.0002098635886673662,
'brun': 0.0002098635886673662,
'brusque': 0.0002098635886673662,
'brûlais': 0.0002098635886673662,
'brûlante': 0.0002098635886673662,
'brûle': 0.0002098635886673662,
'brûler': 0.0002098635886673662,
'bureau': 0.0002098635886673662,
'burin': 0.0002098635886673662,
'but': 0.0002098635886673662,
'buvez': 0.0002098635886673662,
'buvons': 0.0002098635886673662,
'bâiller': 0.0002098635886673662,
'bâti': 0.0002098635886673662,
'bâtir': 0.0002098635886673662,
'bâton': 0.0002098635886673662,
'bénirez': 0.0002098635886673662,
'bénis': 0.0002098635886673662,
'béziers': 0.0002098635886673662,
'bête': 0.0002098635886673662,
'bûcher': 0.0002098635886673662,
'cabarets': 0.0002098635886673662,
'cabinet': 0.0002098635886673662,
'cache': 0.0002098635886673662,
'cacher': 0.0002098635886673662,
'cachette': 0.0002098635886673662,
'cachez': 0.0002098635886673662,
'cachons': 0.0002098635886673662,
'caché': 0.0002098635886673662,
'cachés': 0.0002098635886673662,
'cadence': 0.0002098635886673662,
'cadet': 0.0002098635886673662,
'cadi': 0.0002098635886673662,
'cadre': 0.0002098635886673662,
'caisse': 0.0002098635886673662,
'calcule': 0.0002098635886673662,
'calife': 0.0002098635886673662,
'calliope': 0.0002098635886673662,
'calme': 0.0002098635886673662,
'calmez': 0.0002098635886673662,
'camarade': 0.0002098635886673662,
'camp': 0.0002098635886673662,
'campagne': 0.0002098635886673662,
'camusat': 0.0002098635886673662,
'canaille': 0.0002098635886673662,
'canailles': 0.0002098635886673662,
'canaux': 0.0002098635886673662,
'canon': 0.0002098635886673662,
'capable': 0.0002098635886673662,

'capables': 0.0002098635886673662,
'capitaine': 0.0002098635886673662,
'capitan': 0.0002098635886673662,
'caprice': 0.0002098635886673662,
'caprices': 0.0002098635886673662,
'captifs': 0.0002098635886673662,
'captive': 0.0002098635886673662,
'captiver': 0.0002098635886673662,
'caquet': 0.0002098635886673662,
'car': 0.0002098635886673662,
'caractère': 0.0002098635886673662,
'caressait': 0.0002098635886673662,
'caresser': 0.0002098635886673662,
'caresses': 0.0002098635886673662,
'carnage': 0.0002098635886673662,
'carreau': 0.0002098635886673662,
'carrière': 0.0002098635886673662,
'carrosse': 0.0002098635886673662,
'carte': 0.0002098635886673662,
'cartes': 0.0002098635886673662,
'carthage': 0.0002098635886673662,
'cas': 0.0002098635886673662,
'cassandre': 0.0002098635886673662,
'cassette': 0.0002098635886673662,
'cassé': 0.0002098635886673662,
'cathos': 0.0002098635886673662,
'causa': 0.0002098635886673662,
'causait': 0.0002098635886673662,
'causant': 0.0002098635886673662,
'cause': 0.0002098635886673662,
'causent': 0.0002098635886673662,
'causer': 0.0002098635886673662,
'causes': 0.0002098635886673662,
'causé': 0.0002098635886673662,
'cavalier': 0.0002098635886673662,
'cavaliers': 0.0002098635886673662,
'caverne': 0.0002098635886673662,
'ce': 0.0002098635886673662,
'ceci': 0.0002098635886673662,
'cela': 0.0002098635886673662,
'celais': 0.0002098635886673662,
'celer': 0.0002098635886673662,
'celle': 0.0002098635886673662,
'celles': 0.0002098635886673662,
'celui': 0.0002098635886673662,
'cendre': 0.0002098635886673662,
'cent': 0.0002098635886673662,
'centre': 0.0002098635886673662,
'cents': 0.0002098635886673662,
'cependant': 0.0002098635886673662,
'cerf': 0.0002098635886673662,
'certain': 0.0002098635886673662,
'certaine': 0.0002098635886673662,
'certainement': 0.0002098635886673662,
'certaines': 0.0002098635886673662,
'certains': 0.0002098635886673662,
'certes': 0.0002098635886673662,
'cerveau': 0.0002098635886673662,
'cervelle': 0.0002098635886673662,
'ces': 0.0002098635886673662,
'cesse': 0.0002098635886673662,
'cesser': 0.0002098635886673662,
'cessez': 0.0002098635886673662,
'cest': 0.0002098635886673662,
'cet': 0.0002098635886673662,
'cette': 0.0002098635886673662,
'ceux': 0.0002098635886673662,
'chacun': 0.0002098635886673662,
'chacune': 0.0002098635886673662,

'chagrin': 0.0002098635886673662,
'chair': 0.0002098635886673662,
'chaise': 0.0002098635886673662,
'chaleur': 0.0002098635886673662,
'chambre': 0.0002098635886673662,
'champ': 0.0002098635886673662,
'champmeslé': 0.0002098635886673662,
'champs': 0.0002098635886673662,
'chandelle': 0.0002098635886673662,
'change': 0.0002098635886673662,
'changement': 0.0002098635886673662,
'changer': 0.0002098635886673662,
'changez': 0.0002098635886673662,
'changé': 0.0002098635886673662,
'chanson': 0.0002098635886673662,
'chansons': 0.0002098635886673662,
'chant': 0.0002098635886673662,
'chante': 0.0002098635886673662,
'chanter': 0.0002098635886673662,
'chantre': 0.0002098635886673662,
'chanté': 0.0002098635886673662,
'chaos': 0.0002098635886673662,
'chapeau': 0.0002098635886673662,
'chapons': 0.0002098635886673662,
'chaque': 0.0002098635886673662,
'charge': 0.0002098635886673662,
'charger': 0.0002098635886673662,
'chargé': 0.0002098635886673662,
'charitable': 0.0002098635886673662,
'charité': 0.0002098635886673662,
'charmant': 0.0002098635886673662,
'charmante': 0.0002098635886673662,
'charmants': 0.0002098635886673662,
'charme': 0.0002098635886673662,
'charment': 0.0002098635886673662,
'charmer': 0.0002098635886673662,
'charmes': 0.0002098635886673662,
'charmé': 0.0002098635886673662,
'chasse': 0.0002098635886673662,
'chasser': 0.0002098635886673662,
'chastes': 0.0002098635886673662,
'chat': 0.0002098635886673662,
'chaud': 0.0002098635886673662,
'chaînes': 0.0002098635886673662,
'chef': 0.0002098635886673662,
'chefs': 0.0002098635886673662,
'chemin': 0.0002098635886673662,
'cheminée': 0.0002098635886673662,
'cher': 0.0002098635886673662,
'cherchais': 0.0002098635886673662,
'cherche': 0.0002098635886673662,
'chercher': 0.0002098635886673662,
'cherchez': 0.0002098635886673662,
'cherchons': 0.0002098635886673662,
'cherché': 0.0002098635886673662,
'chers': 0.0002098635886673662,
'cheval': 0.0002098635886673662,
'chevalier': 0.0002098635886673662,
'chevaux': 0.0002098635886673662,
'cheveux': 0.0002098635886673662,
'chez': 0.0002098635886673662,
'chien': 0.0002098635886673662,
'chienne': 0.0002098635886673662,
'chiens': 0.0002098635886673662,
'chimère': 0.0002098635886673662,
'chimérique': 0.0002098635886673662,
'choir': 0.0002098635886673662,
'choisi': 0.0002098635886673662,
'choisir': 0.0002098635886673662,

'choississans': 0.0002098635886673662,
'choix': 0.0002098635886673662,
'chose': 0.0002098635886673662,
'choses': 0.0002098635886673662,
'chromatique': 0.0002098635886673662,
'chute': 0.0002098635886673662,
'château': 0.0002098635886673662,
'châtier': 0.0002098635886673662,
'chère': 0.0002098635886673662,
'chères': 0.0002098635886673662,
'chéri': 0.0002098635886673662,
'chérir': 0.0002098635886673662,
'chéris': 0.0002098635886673662,
'chérit': 0.0002098635886673662,
'chétif': 0.0002098635886673662,
'ci': 0.0002098635886673662,
'cidalise': 0.0002098635886673662,
'ciel': 0.0002098635886673662,
'cieux': 0.0002098635886673662,
'cinq': 0.0002098635886673662,
'cinquième': 0.0002098635886673662,
'circé': 0.0002098635886673662,
'ciseaux': 0.0002098635886673662,
'civile': 0.0002098635886673662,
'civilité': 0.0002098635886673662,
'claire': 0.0002098635886673662,
'clameurs': 0.0002098635886673662,
'clef': 0.0002098635886673662,
'climante': 0.0002098635886673662,
'climats': 0.0002098635886673662,
'clio': 0.0002098635886673662,
'clitandre': 0.0002098635886673662,
'clorimant': 0.0002098635886673662,
'cloris': 0.0002098635886673662,
'cléopâtre': 0.0002098635886673662,
'coeur': 0.0002098635886673662,
'coeurs': 0.0002098635886673662,
'coins': 0.0002098635886673662,
'col': 0.0002098635886673662,
'colette': 0.0002098635886673662,
'colin': 0.0002098635886673662,
'colonnes': 0.0002098635886673662,
'colère': 0.0002098635886673662,
'combat': 0.0002098635886673662,
'combats': 0.0002098635886673662,
'combattre': 0.0002098635886673662,
'combien': 0.0002098635886673662,
'comble': 0.0002098635886673662,
'combler': 0.0002098635886673662,
'comique': 0.0002098635886673662,
'commande': 0.0002098635886673662,
'commander': 0.0002098635886673662,
'commandé': 0.0002098635886673662,
'comme': 0.0002098635886673662,
'commence': 0.0002098635886673662,
'commencement': 0.0002098635886673662,
'commencer': 0.0002098635886673662,
'commencez': 0.0002098635886673662,
'comment': 0.0002098635886673662,
'commerce': 0.0002098635886673662,
'commettrais': 0.0002098635886673662,
'commettre': 0.0002098635886673662,
'commis': 0.0002098635886673662,
'commission': 0.0002098635886673662,
'commode': 0.0002098635886673662,
'commun': 0.0002098635886673662,
'commune': 0.0002098635886673662,
'communes': 0.0002098635886673662,
'compagnie': 0.0002098635886673662,

'comparable': 0.0002098635886673662,
'comparaison': 0.0002098635886673662,
'comparer': 0.0002098635886673662,
'comparée': 0.0002098635886673662,
'compas': 0.0002098635886673662,
'complaire': 0.0002098635886673662,
'complaisance': 0.0002098635886673662,
'compliment': 0.0002098635886673662,
'compliments': 0.0002098635886673662,
'complot': 0.0002098635886673662,
'compose': 0.0002098635886673662,
'composer': 0.0002098635886673662,
'composé': 0.0002098635886673662,
'composés': 0.0002098635886673662,
'comprendre': 0.0002098635886673662,
'comprends': 0.0002098635886673662,
'compris': 0.0002098635886673662,
'compte': 0.0002098635886673662,
'compter': 0.0002098635886673662,
'comptez': 0.0002098635886673662,
'comtesse': 0.0002098635886673662,
'comus': 0.0002098635886673662,
'comédie': 0.0002098635886673662,
'comédien': 0.0002098635886673662,
'comédienne': 0.0002098635886673662,
'comédiens': 0.0002098635886673662,
'comédies': 0.0002098635886673662,
'concert': 0.0002098635886673662,
'concertez': 0.0002098635886673662,
'concerté': 0.0002098635886673662,
'concevoir': 0.0002098635886673662,
'conclure': 0.0002098635886673662,
'conclus': 0.0002098635886673662,
'condamner': 0.0002098635886673662,
'condamnez': 0.0002098635886673662,
'condition': 0.0002098635886673662,
'conduire': 0.0002098635886673662,
'conduisent': 0.0002098635886673662,
'conduit': 0.0002098635886673662,
'conduite': 0.0002098635886673662,
'confesse': 0.0002098635886673662,
'confessez': 0.0002098635886673662,
'confession': 0.0002098635886673662,
'confiance': 0.0002098635886673662,
'confidence': 0.0002098635886673662,
'confiscation': 0.0002098635886673662,
'confond': 0.0002098635886673662,
'confondre': 0.0002098635886673662,
'confondu': 0.0002098635886673662,
'conformer': 0.0002098635886673662,
'confrères': 0.0002098635886673662,
'confus': 0.0002098635886673662,
'confusion': 0.0002098635886673662,
'congé': 0.0002098635886673662,
'congédié': 0.0002098635886673662,
'congédiés': 0.0002098635886673662,
'conjoint': 0.0002098635886673662,
'conjonction': 0.0002098635886673662,
'conjugale': 0.0002098635886673662,
'conjure': 0.0002098635886673662,
'conjurent': 0.0002098635886673662,
'connais': 0.0002098635886673662,
'connaissais': 0.0002098635886673662,
'connaissance': 0.0002098635886673662,
'connaissant': 0.0002098635886673662,
'connaissiez': 0.0002098635886673662,
'connaît': 0.0002098635886673662,
'connaître': 0.0002098635886673662,
'connaîtrez': 0.0002098635886673662,

'connu': 0.0002098635886673662,
'connues': 0.0002098635886673662,
'conquis': 0.0002098635886673662,
'conquérant': 0.0002098635886673662,
'conquérants': 0.0002098635886673662,
'conquêtes': 0.0002098635886673662,
'consacré': 0.0002098635886673662,
'conscience': 0.0002098635886673662,
'conseil': 0.0002098635886673662,
'conseillers': 0.0002098635886673662,
'consens': 0.0002098635886673662,
'consentement': 0.0002098635886673662,
'consentez': 0.0002098635886673662,
'consentir': 0.0002098635886673662,
'conserve': 0.0002098635886673662,
'conserver': 0.0002098635886673662,
'considère': 0.0002098635886673662,
'considérer': 0.0002098635886673662,
'considérez': 0.0002098635886673662,
'considéré': 0.0002098635886673662,
'considérés': 0.0002098635886673662,
'console': 0.0002098635886673662,
'consoler': 0.0002098635886673662,
'consolez': 0.0002098635886673662,
'consommez': 0.0002098635886673662,
'constance': 0.0002098635886673662,
'consulter': 0.0002098635886673662,
'consultons': 0.0002098635886673662,
'conséquence': 0.0002098635886673662,
'conséquent': 0.0002098635886673662,
'conte': 0.0002098635886673662,
'contemplant': 0.0002098635886673662,
'contemple': 0.0002098635886673662,
'contempler': 0.0002098635886673662,
'content': 0.0002098635886673662,
'contente': 0.0002098635886673662,
'contenter': 0.0002098635886673662,
'contents': 0.0002098635886673662,
'conter': 0.0002098635886673662,
'contera': 0.0002098635886673662,
'contes': 0.0002098635886673662,
'contester': 0.0002098635886673662,
'contez': 0.0002098635886673662,
'continuer': 0.0002098635886673662,
'contracter': 0.0002098635886673662,
'contraindre': 0.0002098635886673662,
'contraint': 0.0002098635886673662,
'contrainte': 0.0002098635886673662,
'contraints': 0.0002098635886673662,
'contraire': 0.0002098635886673662,
'contrat': 0.0002098635886673662,
'contre': 0.0002098635886673662,
'contredire': 0.0002098635886673662,
'contredis': 0.0002098635886673662,
'contretemps': 0.0002098635886673662,
'contrevenants': 0.0002098635886673662,
'conté': 0.0002098635886673662,
'convainc': 0.0002098635886673662,
'convaincue': 0.0002098635886673662,
'convenir': 0.0002098635886673662,
'conversation': 0.0002098635886673662,
'convie': 0.0002098635886673662,
'conviens': 0.0002098635886673662,
'convier': 0.0002098635886673662,
'conçois': 0.0002098635886673662,
'conçoit': 0.0002098635886673662,
'conçu': 0.0002098635886673662,
'copie': 0.0002098635886673662,
'coq': 0.0002098635886673662,

```

'coquette': 0.0002098635886673662,
'coquettes': 0.0002098635886673662,
'coquin': 0.0002098635886673662,
'coquins': 0.0002098635886673662,
'coral': 0.0002098635886673662,
'corde': 0.0002098635886673662,
'corner': 0.0002098635886673662,
'cornes': 0.0002098635886673662,
'corps': 0.0002098635886673662,
'correspondance': 0.0002098635886673662,
'couche': 0.0002098635886673662,
'coucher': 0.0002098635886673662,
'couchés': 0.0002098635886673662,
'couler': 0.0002098635886673662,
'couleur': 0.0002098635886673662,
'couleurs': 0.0002098635886673662,
'coup': 0.0002098635886673662,
'coupable': 0.0002098635886673662,
'coupant': 0.0002098635886673662,
'coups': 0.0002098635886673662,
'cour': 0.0002098635886673662,
'courage': 0.0002098635886673662,
'courait': 0.0002098635886673662,
'courante': 0.0002098635886673662,
'courent': 0.0002098635886673662,
'courez': 0.0002098635886673662,
'courir': 0.0002098635886673662,
'couronne': 0.0002098635886673662,
'couronner': 0.0002098635886673662,
'courons': 0.0002098635886673662,
'courroux': 0.0002098635886673662,
'cours': 0.0002098635886673662,
'course': 0.0002098635886673662,
'court': 0.0002098635886673662,
'courtisans': 0.0002098635886673662,
'courtois': 0.0002098635886673662,
'courtoise': 0.0002098635886673662,
'courtoisie': 0.0002098635886673662,
'coursu': 0.0002098635886673662,
'cousine': 0.0002098635886673662,
'coutume': 0.0002098635886673662,
'couvert': 0.0002098635886673662,
'couverts': 0.0002098635886673662,
'couvre': 0.0002098635886673662,
'couvrez': 0.0002098635886673662,
...}

```

In [48]:

```

stemmed_vocabulary_tra_com_rand = [fr.stem(word) for word in vocabulary_tra_com_rand]
counts = dict()
for word in vocabulary_tra_com_rand:
    if word in counts:
        counts[word] += 1
    else:
        counts[word] = 1

tf_dict_tra_com_rand = counts
for key, value in counts.items():
    tf_dict_tra_com_rand[key] = value / len(vocabulary_tra_com_rand)
tf_dict_tra_com_rand

```

Out[48]:

```

{'a': 0.00016949152542372882,
'abaisse': 0.00016949152542372882,
'abaissement': 0.00016949152542372882,
'abaisser': 0.00016949152542372882,
'abandonnant': 0.00016949152542372882,
'abandonne': 0.00016949152542372882,
'abandonner': 0.00016949152542372882,
'abandonnez': 0.00016949152542372882,

```

'abandonné': 0.00016949152542372882,
'abattre': 0.00016949152542372882,
'abattu': 0.00016949152542372882,
'abattus': 0.00016949152542372882,
'abhorre': 0.00016949152542372882,
'abois': 0.00016949152542372882,
'abominable': 0.00016949152542372882,
'abonde': 0.00016949152542372882,
'abord': 0.00016949152542372882,
'aborder': 0.00016949152542372882,
'abri': 0.00016949152542372882,
'abrégé': 0.00016949152542372882,
'abrégé': 0.00016949152542372882,
'absence': 0.00016949152542372882,
'absent': 0.00016949152542372882,
'absolu': 0.00016949152542372882,
'absolue': 0.00016949152542372882,
'absolument': 0.00016949152542372882,
'absolus': 0.00016949152542372882,
'abuser': 0.00016949152542372882,
'abusé': 0.00016949152542372882,
'abusée': 0.00016949152542372882,
'abîme': 0.00016949152542372882,
'accable': 0.00016949152542372882,
'accabler': 0.00016949152542372882,
'accablé': 0.00016949152542372882,
'acceptant': 0.00016949152542372882,
'accepter': 0.00016949152542372882,
'acceptez': 0.00016949152542372882,
'accepté': 0.00016949152542372882,
'accident': 0.00016949152542372882,
'accidents': 0.00016949152542372882,
'accommoder': 0.00016949152542372882,
'accompagne': 0.00016949152542372882,
'accompagné': 0.00016949152542372882,
'accompli': 0.00016949152542372882,
'accomplir': 0.00016949152542372882,
'accomplis': 0.00016949152542372882,
'accomplisse': 0.00016949152542372882,
'accomplissez': 0.00016949152542372882,
'accord': 0.00016949152542372882,
'accorde': 0.00016949152542372882,
'accorder': 0.00016949152542372882,
'accorderez': 0.00016949152542372882,
'accordez': 0.00016949152542372882,
'accoutumé': 0.00016949152542372882,
'accoutumée': 0.00016949152542372882,
'accroît': 0.00016949152542372882,
'accroître': 0.00016949152542372882,
'accueil': 0.00016949152542372882,
'accuse': 0.00016949152542372882,
'accuser': 0.00016949152542372882,
'accusez': 0.00016949152542372882,
'accusons': 0.00016949152542372882,
'accès': 0.00016949152542372882,
'acheter': 0.00016949152542372882,
'achetez': 0.00016949152542372882,
'acheté': 0.00016949152542372882,
'acheva': 0.00016949152542372882,
'achever': 0.00016949152542372882,
'achevez': 0.00016949152542372882,
'achevons': 0.00016949152542372882,
'achevé': 0.00016949152542372882,
'achevée': 0.00016949152542372882,
'achève': 0.00016949152542372882,
'acquiérs': 0.00016949152542372882,
'acquiert': 0.00016949152542372882,
'acquis': 0.00016949152542372882,
'acquise': 0.00016949152542372882,

'acquies': 0.00016949152542372882,
'acquitte': 0.00016949152542372882,
'acquitter': 0.00016949152542372882,
'acquitté': 0.00016949152542372882,
'acquière': 0.00016949152542372882,
'acquérir': 0.00016949152542372882,
'acte': 0.00016949152542372882,
'acteurs': 0.00016949152542372882,
'action': 0.00016949152542372882,
'actions': 0.00016949152542372882,
'adieu': 0.00016949152542372882,
'adieux': 0.00016949152542372882,
'admet': 0.00016949152542372882,
'admets': 0.00016949152542372882,
'admettre': 0.00016949152542372882,
'admirable': 0.00016949152542372882,
'admire': 0.00016949152542372882,
'admirer': 0.00016949152542372882,
'adoption': 0.00016949152542372882,
'adorable': 0.00016949152542372882,
'adorables': 0.00016949152542372882,
'adorant': 0.00016949152542372882,
'adore': 0.00016949152542372882,
'adoren': 0.00016949152542372882,
'adores': 0.00016949152542372882,
'adoré': 0.00016949152542372882,
'adouci': 0.00016949152542372882,
'adoucir': 0.00016949152542372882,
'adoucissent': 0.00016949152542372882,
'adresse': 0.00016949152542372882,
'adressent': 0.00016949152542372882,
'adresser': 0.00016949152542372882,
'adressez': 0.00016949152542372882,
'adroit': 0.00016949152542372882,
'adroite': 0.00016949152542372882,
'adroitement': 0.00016949152542372882,
'adversaire': 0.00016949152542372882,
'adversaires': 0.00016949152542372882,
'affaiblir': 0.00016949152542372882,
'affaire': 0.00016949152542372882,
'affaires': 0.00016949152542372882,
'affection': 0.00016949152542372882,
'affections': 0.00016949152542372882,
'affermi': 0.00016949152542372882,
'afferries': 0.00016949152542372882,
'affermer': 0.00016949152542372882,
'affliction': 0.00016949152542372882,
'afflige': 0.00016949152542372882,
'affligea': 0.00016949152542372882,
'affligeas': 0.00016949152542372882,
'affligerait': 0.00016949152542372882,
'affligée': 0.00016949152542372882,
'affranchie': 0.00016949152542372882,
'affranchir': 0.00016949152542372882,
'affreuse': 0.00016949152542372882,
'affreux': 0.00016949152542372882,
'affront': 0.00016949152542372882,
'affronter': 0.00016949152542372882,
'affronts': 0.00016949152542372882,
'affrontée': 0.00016949152542372882,
'afin': 0.00016949152542372882,
'agencé': 0.00016949152542372882,
'agent': 0.00016949152542372882,
'agents': 0.00016949152542372882,
'agir': 0.00016949152542372882,
'agisse': 0.00016949152542372882,
'agissez': 0.00016949152542372882,
'agit': 0.00016949152542372882,
'agrandi': 0.00016949152542372882,

'agrandir': 0.00016949152542372882,
'agrandit': 0.00016949152542372882,
'agréable': 0.00016949152542372882,
'agréer': 0.00016949152542372882,
'agréerez': 0.00016949152542372882,
'agréez': 0.00016949152542372882,
'aguerris': 0.00016949152542372882,
'ah': 0.00016949152542372882,
'ai': 0.00016949152542372882,
'aidant': 0.00016949152542372882,
'aide': 0.00016949152542372882,
'aider': 0.00016949152542372882,
'aie': 0.00016949152542372882,
'aient': 0.00016949152542372882,
'aigle': 0.00016949152542372882,
'aigreur': 0.00016949152542372882,
'aigri': 0.00016949152542372882,
'ailes': 0.00016949152542372882,
'aille': 0.00016949152542372882,
'ailleurs': 0.00016949152542372882,
'aima': 0.00016949152542372882,
'aimable': 0.00016949152542372882,
'aimables': 0.00016949152542372882,
'aimai': 0.00016949152542372882,
'aimais': 0.00016949152542372882,
'aimait': 0.00016949152542372882,
'aimant': 0.00016949152542372882,
'aimas': 0.00016949152542372882,
'aime': 0.00016949152542372882,
'aiment': 0.00016949152542372882,
'aimer': 0.00016949152542372882,
'aimerais': 0.00016949152542372882,
'aimerais': 0.00016949152542372882,
'aimes': 0.00016949152542372882,
'aimez': 0.00016949152542372882,
'aimiez': 0.00016949152542372882,
'aimons': 0.00016949152542372882,
'aimât': 0.00016949152542372882,
'aimé': 0.00016949152542372882,
'aimée': 0.00016949152542372882,
'ainsi': 0.00016949152542372882,
'air': 0.00016949152542372882,
'airs': 0.00016949152542372882,
'aise': 0.00016949152542372882,
'aisé': 0.00016949152542372882,
'aisée': 0.00016949152542372882,
'aisément': 0.00016949152542372882,
'aisés': 0.00016949152542372882,
'ait': 0.00016949152542372882,
'ajoute': 0.00016949152542372882,
'ajouter': 0.00016949152542372882,
'ajoutez': 0.00016949152542372882,
'ajustant': 0.00016949152542372882,
'alarme': 0.00016949152542372882,
'alarmes': 0.00016949152542372882,
'alarmée': 0.00016949152542372882,
'albe': 0.00016949152542372882,
'albion': 0.00016949152542372882,
'alcale': 0.00016949152542372882,
'alcide': 0.00016949152542372882,
'alcione': 0.00016949152542372882,
'alentit': 0.00016949152542372882,
'alentour': 0.00016949152542372882,
'alexandre': 0.00016949152542372882,
'aliment': 0.00016949152542372882,
'allais': 0.00016949152542372882,
'allait': 0.00016949152542372882,
'aller': 0.00016949152542372882,
'allez': 0.00016949152542372882,

'alliance': 0.00016949152542372882,
'allie': 0.00016949152542372882,
'alliez': 0.00016949152542372882,
'allié': 0.00016949152542372882,
'alliés': 0.00016949152542372882,
'allomice': 0.00016949152542372882,
'allongerai': 0.00016949152542372882,
'allons': 0.00016949152542372882,
'alluma': 0.00016949152542372882,
'allume': 0.00016949152542372882,
'allumées': 0.00016949152542372882,
'allèchements': 0.00016949152542372882,
'allégeance': 0.00016949152542372882,
'alléger': 0.00016949152542372882,
'allégorie': 0.00016949152542372882,
'allégresse': 0.00016949152542372882,
'alonne': 0.00016949152542372882,
'alors': 0.00016949152542372882,
'alphénor': 0.00016949152542372882,
'altesse': 0.00016949152542372882,
'altière': 0.00016949152542372882,
'altère': 0.00016949152542372882,
'altérant': 0.00016949152542372882,
'alve': 0.00016949152542372882,
'amadis': 0.00016949152542372882,
'amant': 0.00016949152542372882,
'amante': 0.00016949152542372882,
'amants': 0.00016949152542372882,
'amas': 0.00016949152542372882,
'ambassadeur': 0.00016949152542372882,
'ambassadeurs': 0.00016949152542372882,
'ambitieuses': 0.00016949152542372882,
'ambitieux': 0.00016949152542372882,
'ambition': 0.00016949152542372882,
'amenant': 0.00016949152542372882,
'amende': 0.00016949152542372882,
'amendes': 0.00016949152542372882,
'amener': 0.00016949152542372882,
'amenez': 0.00016949152542372882,
'amertume': 0.00016949152542372882,
'ami': 0.00016949152542372882,
'amie': 0.00016949152542372882,
'amiral': 0.00016949152542372882,
'amis': 0.00016949152542372882,
'amitié': 0.00016949152542372882,
'amitiés': 0.00016949152542372882,
'amour': 0.00016949152542372882,
'amoureuse': 0.00016949152542372882,
'amoureuses': 0.00016949152542372882,
'amoureux': 0.00016949152542372882,
'amours': 0.00016949152542372882,
'ample': 0.00016949152542372882,
'ablement': 0.00016949152542372882,
'amuser': 0.00016949152542372882,
'amène': 0.00016949152542372882,
'amères': 0.00016949152542372882,
'amélie': 0.00016949152542372882,
'anaxandre': 0.00016949152542372882,
'anche': 0.00016949152542372882,
'andalousie': 0.00016949152542372882,
'andrigene': 0.00016949152542372882,
'andrion': 0.00016949152542372882,
'ange': 0.00016949152542372882,
'angleterre': 0.00016949152542372882,
'animal': 0.00016949152542372882,
'anime': 0.00016949152542372882,
'animé': 0.00016949152542372882,
'animée': 0.00016949152542372882,
'annoncent': 0.00016949152542372882,

'annoncez': 0.00016949152542372882,
'année': 0.00016949152542372882,
'années': 0.00016949152542372882,
'ans': 0.00016949152542372882,
'antiquité': 0.00016949152542372882,
'antoine': 0.00016949152542372882,
'août': 0.00016949152542372882,
'apaiser': 0.00016949152542372882,
'apaisez': 0.00016949152542372882,
'apaisé': 0.00016949152542372882,
'apaisée': 0.00016949152542372882,
'apaisés': 0.00016949152542372882,
'aperçois': 0.00016949152542372882,
'aperçu': 0.00016949152542372882,
'apollon': 0.00016949152542372882,
'appareil': 0.00016949152542372882,
'apparement': 0.00016949152542372882,
'apparence': 0.00016949152542372882,
'apparences': 0.00016949152542372882,
'apparent': 0.00016949152542372882,
'appartement': 0.00016949152542372882,
'appartient': 0.00016949152542372882,
'appas': 0.00016949152542372882,
'appelait': 0.00016949152542372882,
'appeler': 0.00016949152542372882,
'appelez': 0.00016949152542372882,
'appelle': 0.00016949152542372882,
'applaudissement': 0.00016949152542372882,
'appliquer': 0.00016949152542372882,
'apporte': 0.00016949152542372882,
'apprenant': 0.00016949152542372882,
'apprend': 0.00016949152542372882,
'apprendra': 0.00016949152542372882,
'apprendrai': 0.00016949152542372882,
'apprendre': 0.00016949152542372882,
'apprendrez': 0.00016949152542372882,
'apprends': 0.00016949152542372882,
'apprenez': 0.00016949152542372882,
'apprenne': 0.00016949152542372882,
'apprenons': 0.00016949152542372882,
'apprirent': 0.00016949152542372882,
'appris': 0.00016949152542372882,
'approche': 0.00016949152542372882,
'approcher': 0.00016949152542372882,
'approchez': 0.00016949152542372882,
'approchons': 0.00016949152542372882,
'approchés': 0.00016949152542372882,
'approuve': 0.00016949152542372882,
'approuver': 0.00016949152542372882,
'approuvez': 0.00016949152542372882,
'appréhende': 0.00016949152542372882,
'appréhender': 0.00016949152542372882,
'apprête': 0.00016949152542372882,
'appui': 0.00016949152542372882,
'appuis': 0.00016949152542372882,
'appuya': 0.00016949152542372882,
'appuyait': 0.00016949152542372882,
'appuyant': 0.00016949152542372882,
'appuyer': 0.00016949152542372882,
'appuyé': 0.00016949152542372882,
'après': 0.00016949152542372882,
'aquilons': 0.00016949152542372882,
'aragon': 0.00016949152542372882,
'arbitre': 0.00016949152542372882,
'arbres': 0.00016949152542372882,
'arcas': 0.00016949152542372882,
'archers': 0.00016949152542372882,
'architalassie': 0.00016949152542372882,
'arcs': 0.00016949152542372882,

'arctodeme': 0.00016949152542372882,
'ardemment': 0.00016949152542372882,
'ardent': 0.00016949152542372882,
'ardente': 0.00016949152542372882,
'ardents': 0.00016949152542372882,
'ardeur': 0.00016949152542372882,
'ardeurs': 0.00016949152542372882,
'argent': 0.00016949152542372882,
'armand': 0.00016949152542372882,
'arme': 0.00016949152542372882,
'armer': 0.00016949152542372882,
'armerai': 0.00016949152542372882,
'armes': 0.00016949152542372882,
'armet': 0.00016949152542372882,
'armez': 0.00016949152542372882,
'armé': 0.00016949152542372882,
'armée': 0.00016949152542372882,
'arrache': 0.00016949152542372882,
'arracher': 0.00016949152542372882,
'arraché': 0.00016949152542372882,
'arrhes': 0.00016949152542372882,
'arrivant': 0.00016949152542372882,
'arrive': 0.00016949152542372882,
'arriver': 0.00016949152542372882,
'arrivera': 0.00016949152542372882,
'arrivé': 0.00016949152542372882,
'arrivée': 0.00016949152542372882,
'arrière': 0.00016949152542372882,
'arrogance': 0.00016949152542372882,
'arrogant': 0.00016949152542372882,
'arrêt': 0.00016949152542372882,
'arrêtant': 0.00016949152542372882,
'arrête': 0.00016949152542372882,
'arrêter': 0.00016949152542372882,
'arrêtez': 0.00016949152542372882,
'arrêts': 0.00016949152542372882,
'arrêté': 0.00016949152542372882,
'arrêtée': 0.00016949152542372882,
'arrêtés': 0.00016949152542372882,
'art': 0.00016949152542372882,
'artifice': 0.00016949152542372882,
'artifices': 0.00016949152542372882,
'artisan': 0.00016949152542372882,
'arts': 0.00016949152542372882,
'as': 0.00016949152542372882,
'ascendant': 0.00016949152542372882,
'asile': 0.00016949152542372882,
'aspect': 0.00016949152542372882,
'aspects': 0.00016949152542372882,
'aspire': 0.00016949152542372882,
'aspirer': 0.00016949152542372882,
'aspirez': 0.00016949152542372882,
'assassin': 0.00016949152542372882,
'assassinat': 0.00016949152542372882,
'assassinats': 0.00016949152542372882,
'assassine': 0.00016949152542372882,
'assassiner': 0.00016949152542372882,
'assemble': 0.00016949152542372882,
'assemblez': 0.00016949152542372882,
'asservi': 0.00016949152542372882,
'asservie': 0.00016949152542372882,
'asservir': 0.00016949152542372882,
'assez': 0.00016949152542372882,
'assis': 0.00016949152542372882,
'assistance': 0.00016949152542372882,
'assiste': 0.00016949152542372882,
'assister': 0.00016949152542372882,
'associé': 0.00016949152542372882,
'assomme': 0.00016949152542372882,

'assorti': 0.00016949152542372882,
'assoupir': 0.00016949152542372882,
'assouppissement': 0.00016949152542372882,
'assouvie': 0.00016949152542372882,
'assouvir': 0.00016949152542372882,
'assujettie': 0.00016949152542372882,
'assujettir': 0.00016949152542372882,
'assurance': 0.00016949152542372882,
'assurant': 0.00016949152542372882,
'assure': 0.00016949152542372882,
'assurer': 0.00016949152542372882,
'assuré': 0.00016949152542372882,
'assurément': 0.00016949152542372882,
'astre': 0.00016949152542372882,
'astres': 0.00016949152542372882,
'astérie': 0.00016949152542372882,
'attache': 0.00016949152542372882,
'attachement': 0.00016949152542372882,
'attachent': 0.00016949152542372882,
'attaqua': 0.00016949152542372882,
'attaque': 0.00016949152542372882,
'attaquer': 0.00016949152542372882,
'attaquez': 0.00016949152542372882,
'attaquons': 0.00016949152542372882,
'attaqué': 0.00016949152542372882,
'atteindre': 0.00016949152542372882,
'atteint': 0.00016949152542372882,
'atteinte': 0.00016949152542372882,
'atteintes': 0.00016949152542372882,
'attend': 0.00016949152542372882,
'attendais': 0.00016949152542372882,
'attendant': 0.00016949152542372882,
'attendez': 0.00016949152542372882,
'attendons': 0.00016949152542372882,
'attendre': 0.00016949152542372882,
'attendri': 0.00016949152542372882,
'attendrie': 0.00016949152542372882,
'attendrissez': 0.00016949152542372882,
'attends': 0.00016949152542372882,
'attendu': 0.00016949152542372882,
'attentat': 0.00016949152542372882,
'attentats': 0.00016949152542372882,
'attente': 0.00016949152542372882,
'attenter': 0.00016949152542372882,
'attentes': 0.00016949152542372882,
'attentif': 0.00016949152542372882,
'attention': 0.00016949152542372882,
'atteste': 0.00016949152542372882,
'attire': 0.00016949152542372882,
'attirent': 0.00016949152542372882,
'attraits': 0.00016949152542372882,
'attrape': 0.00016949152542372882,
'au': 0.00016949152542372882,
'aucun': 0.00016949152542372882,
'aucune': 0.00016949152542372882,
'aucunement': 0.00016949152542372882,
'audace': 0.00016949152542372882,
'audacieux': 0.00016949152542372882,
'audience': 0.00016949152542372882,
'audit': 0.00016949152542372882,
'augmente': 0.00016949152542372882,
'augmenter': 0.00016949152542372882,
'augmenterai': 0.00016949152542372882,
'augure': 0.00016949152542372882,
'auguste': 0.00016949152542372882,
'aujourd'hui': 0.00016949152542372882,
'aumône': 0.00016949152542372882,
'auparavant': 0.00016949152542372882,
'auprès': 0.00016949152542372882,

'aura': 0.00016949152542372882,
'aurai': 0.00016949152542372882,
'auraient': 0.00016949152542372882,
'aurais': 0.00016949152542372882,
'aurait': 0.00016949152542372882,
'auras': 0.00016949152542372882,
'aurez': 0.00016949152542372882,
'auriez': 0.00016949152542372882,
'aurons': 0.00016949152542372882,
'auront': 0.00016949152542372882,
'aurore': 0.00016949152542372882,
'aussi': 0.00016949152542372882,
'aussitôt': 0.00016949152542372882,
'austère': 0.00016949152542372882,
'austères': 0.00016949152542372882,
'autant': 0.00016949152542372882,
'autels': 0.00016949152542372882,
'auteur': 0.00016949152542372882,
'auteurs': 0.00016949152542372882,
'autorise': 0.00016949152542372882,
'autoriser': 0.00016949152542372882,
'autorité': 0.00016949152542372882,
'autour': 0.00016949152542372882,
'autre': 0.00016949152542372882,
'autrefois': 0.00016949152542372882,
'autrement': 0.00016949152542372882,
'autres': 0.00016949152542372882,
'autrui': 0.00016949152542372882,
'aux': 0.00016949152542372882,
'avaient': 0.00016949152542372882,
'avais': 0.00016949152542372882,
'avait': 0.00016949152542372882,
'avance': 0.00016949152542372882,
'avancer': 0.00016949152542372882,
'avancez': 0.00016949152542372882,
'avancé': 0.00016949152542372882,
'avancées': 0.00016949152542372882,
'avant': 0.00016949152542372882,
'avantage': 0.00016949152542372882,
'avantages': 0.00016949152542372882,
'avantageux': 0.00016949152542372882,
'avançons': 0.00016949152542372882,
'avare': 0.00016949152542372882,
'avec': 0.00016949152542372882,
'avecque': 0.00016949152542372882,
'avecques': 0.00016949152542372882,
'avenir': 0.00016949152542372882,
'aventure': 0.00016949152542372882,
'aventures': 0.00016949152542372882,
'aversion': 0.00016949152542372882,
'avertie': 0.00016949152542372882,
'avertir': 0.00016949152542372882,
'aveu': 0.00016949152542372882,
'aveugle': 0.00016949152542372882,
'aveuglement': 0.00016949152542372882,
'aveuglément': 0.00016949152542372882,
'avez': 0.00016949152542372882,
'aviez': 0.00016949152542372882,
'avons': 0.00016949152542372882,
'avis': 0.00016949152542372882,
'aviser': 0.00016949152542372882,
'aviserez': 0.00016949152542372882,
'avocat': 0.00016949152542372882,
'avoient': 0.00016949152542372882,
'avoir': 0.00016949152542372882,
'avons': 0.00016949152542372882,
'avorter': 0.00016949152542372882,
'avoue': 0.00016949152542372882,
'avouer': 0.00016949152542372882,

'avoueraï': 0.00016949152542372882,
'avouez': 0.00016949152542372882,
'avoué': 0.00016949152542372882,
'avérer': 0.00016949152542372882,
'avéré': 0.00016949152542372882,
'ayant': 0.00016949152542372882,
'aye': 0.00016949152542372882,
'ayez': 0.00016949152542372882,
'aymon': 0.00016949152542372882,
'ayons': 0.00016949152542372882,
'aïeux': 0.00016949152542372882,
'badin': 0.00016949152542372882,
'baise': 0.00016949152542372882,
'baiser': 0.00016949152542372882,
'baisers': 0.00016949152542372882,
'balance': 0.00016949152542372882,
'balancer': 0.00016949152542372882,
'balcon': 0.00016949152542372882,
'bande': 0.00016949152542372882,
'bandeau': 0.00016949152542372882,
'bander': 0.00016949152542372882,
'banni': 0.00016949152542372882,
'bannie': 0.00016949152542372882,
'bannir': 0.00016949152542372882,
'bannis': 0.00016949152542372882,
'bannit': 0.00016949152542372882,
'barbare': 0.00016949152542372882,
'barbe': 0.00016949152542372882,
'barbes': 0.00016949152542372882,
'barbier': 0.00016949152542372882,
'barbus': 0.00016949152542372882,
'barcelone': 0.00016949152542372882,
'barque': 0.00016949152542372882,
'bas': 0.00016949152542372882,
'base': 0.00016949152542372882,
'basilic': 0.00016949152542372882,
'basion': 0.00016949152542372882,
'basse': 0.00016949152542372882,
'bassesse': 0.00016949152542372882,
'bassin': 0.00016949152542372882,
'baste': 0.00016949152542372882,
'bastions': 0.00016949152542372882,
'bat': 0.00016949152542372882,
'batailles': 0.00016949152542372882,
'bats': 0.00016949152542372882,
'battez': 0.00016949152542372882,
'battons': 0.00016949152542372882,
'battre': 0.00016949152542372882,
'battu': 0.00016949152542372882,
'baudet': 0.00016949152542372882,
'beau': 0.00016949152542372882,
'beaucoup': 0.00016949152542372882,
'beaufort': 0.00016949152542372882,
'beauté': 0.00016949152542372882,
'beautés': 0.00016949152542372882,
'beaux': 0.00016949152542372882,
'bel': 0.00016949152542372882,
'belle': 0.00016949152542372882,
'belles': 0.00016949152542372882,
'bernard': 0.00016949152542372882,
'besoin': 0.00016949152542372882,
'besoins': 0.00016949152542372882,
'bien': 0.00016949152542372882,
'bienfait': 0.00016949152542372882,
'bienfaits': 0.00016949152542372882,
'bienheureux': 0.00016949152542372882,
'biens': 0.00016949152542372882,
'bientôt': 0.00016949152542372882,
'bile': 0.00016949152542372882,

'billet': 0.00016949152542372882,
'bizarre': 0.00016949152542372882,
'bizarrerie': 0.00016949152542372882,
'blanc': 0.00016949152542372882,
'blanchir': 0.00016949152542372882,
'blessait': 0.00016949152542372882,
'blesse': 0.00016949152542372882,
'blesser': 0.00016949152542372882,
'blessure': 0.00016949152542372882,
'blessures': 0.00016949152542372882,
'blessé': 0.00016949152542372882,
'blessée': 0.00016949152542372882,
'blessés': 0.00016949152542372882,
'blâmable': 0.00016949152542372882,
'blâme': 0.00016949152542372882,
'blâmer': 0.00016949152542372882,
'blâmez': 0.00016949152542372882,
'blâmions': 0.00016949152542372882,
'blâmé': 0.00016949152542372882,
'blême': 0.00016949152542372882,
'boire': 0.00016949152542372882,
'bois': 0.00016949152542372882,
'bon': 0.00016949152542372882,
'bonheur': 0.00016949152542372882,
'bonhomme': 0.00016949152542372882,
'bonne': 0.00016949152542372882,
'bonnes': 0.00016949152542372882,
'bons': 0.00016949152542372882,
'bonté': 0.00016949152542372882,
'bontés': 0.00016949152542372882,
'bord': 0.00016949152542372882,
'bordeaux': 0.00016949152542372882,
'bordelais': 0.00016949152542372882,
'bords': 0.00016949152542372882,
'borne': 0.00016949152542372882,
'borner': 0.00016949152542372882,
'bornez': 0.00016949152542372882,
'borné': 0.00016949152542372882,
'bornées': 0.00016949152542372882,
'bouche': 0.00016949152542372882,
'bouches': 0.00016949152542372882,
'boue': 0.00016949152542372882,
'bouffon': 0.00016949152542372882,
'bouillant': 0.00016949152542372882,
'bouillants': 0.00016949152542372882,
'bouillon': 0.00016949152542372882,
'bourgogne': 0.00016949152542372882,
'bourreaux': 0.00016949152542372882,
'bourse': 0.00016949152542372882,
'bout': 0.00016949152542372882,
'branle': 0.00016949152542372882,
'branler': 0.00016949152542372882,
'bras': 0.00016949152542372882,
'brasier': 0.00016949152542372882,
'brasiers': 0.00016949152542372882,
'brave': 0.00016949152542372882,
'braver': 0.00016949152542372882,
'braves': 0.00016949152542372882,
'bravoure': 0.00016949152542372882,
'bref': 0.00016949152542372882,
'bretteurs': 0.00016949152542372882,
'brigue': 0.00016949152542372882,
'brigues': 0.00016949152542372882,
'brillant': 0.00016949152542372882,
'brillants': 0.00016949152542372882,
'brillant': 0.00016949152542372882,
'briller': 0.00016949152542372882,
'briser': 0.00016949152542372882,
'broncher': 0.00016949152542372882,

'bruit': 0.00016949152542372882,
'bruits': 0.00016949152542372882,
'brun': 0.00016949152542372882,
'brutal': 0.00016949152542372882,
'brûla': 0.00016949152542372882,
'brûlai': 0.00016949152542372882,
'brûlant': 0.00016949152542372882,
'brûlants': 0.00016949152542372882,
'brûle': 0.00016949152542372882,
'brûler': 0.00016949152542372882,
'but': 0.00016949152542372882,
'butin': 0.00016949152542372882,
'butte': 0.00016949152542372882,
'bâtir': 0.00016949152542372882,
'bâtis': 0.00016949152542372882,
'bâton': 0.00016949152542372882,
'béatrix': 0.00016949152542372882,
'béatrix': 0.00016949152542372882,
'bénignes': 0.00016949152542372882,
'bénins': 0.00016949152542372882,
'bénis': 0.00016949152542372882,
'cabinet': 0.00016949152542372882,
'cabrer': 0.00016949152542372882,
'cacha': 0.00016949152542372882,
'cachaient': 0.00016949152542372882,
'cachait': 0.00016949152542372882,
'cache': 0.00016949152542372882,
'cachent': 0.00016949152542372882,
'cacher': 0.00016949152542372882,
'cachera': 0.00016949152542372882,
'cachez': 0.00016949152542372882,
'cachons': 0.00016949152542372882,
'caché': 0.00016949152542372882,
'cachés': 0.00016949152542372882,
'cage': 0.00016949152542372882,
'calamités': 0.00016949152542372882,
'calaris': 0.00016949152542372882,
'calme': 0.00016949152542372882,
'calmer': 0.00016949152542372882,
'calmez': 0.00016949152542372882,
'camarade': 0.00016949152542372882,
'camp': 0.00016949152542372882,
'campagne': 0.00016949152542372882,
'campagnes': 0.00016949152542372882,
'camps': 0.00016949152542372882,
'candeur': 0.00016949152542372882,
'canton': 0.00016949152542372882,
'capable': 0.00016949152542372882,
'capables': 0.00016949152542372882,
'capitaine': 0.00016949152542372882,
'capitale': 0.00016949152542372882,
'caprice': 0.00016949152542372882,
'caprices': 0.00016949152542372882,
'capricieuse': 0.00016949152542372882,
'captieux': 0.00016949152542372882,
'captif': 0.00016949152542372882,
'captifs': 0.00016949152542372882,
'captive': 0.00016949152542372882,
'captivent': 0.00016949152542372882,
'captivité': 0.00016949152542372882,
'captivé': 0.00016949152542372882,
'car': 0.00016949152542372882,
'caractère': 0.00016949152542372882,
'cardenie': 0.00016949152542372882,
'cardille': 0.00016949152542372882,
'cardillon': 0.00016949152542372882,
'caressait': 0.00016949152542372882,
'caresse': 0.00016949152542372882,
'caresses': 0.00016949152542372882,

'carlos': 0.00016949152542372882,
'carnage': 0.00016949152542372882,
'carrière': 0.00016949152542372882,
'carrosse': 0.00016949152542372882,
'cartel': 0.00016949152542372882,
'cas': 0.00016949152542372882,
'casque': 0.00016949152542372882,
'cassandre': 0.00016949152542372882,
'cassée': 0.00016949152542372882,
'castillan': 0.00016949152542372882,
'castillans': 0.00016949152542372882,
'castille': 0.00016949152542372882,
'catalan': 0.00016949152542372882,
'causa': 0.00016949152542372882,
'causait': 0.00016949152542372882,
'causant': 0.00016949152542372882,
'cause': 0.00016949152542372882,
'causer': 0.00016949152542372882,
'causera': 0.00016949152542372882,
'causes': 0.00016949152542372882,
'causèrent': 0.00016949152542372882,
'causé': 0.00016949152542372882,
'cavalier': 0.00016949152542372882,
'ce': 0.00016949152542372882,
'ceci': 0.00016949152542372882,
'cela': 0.00016949152542372882,
'celer': 0.00016949152542372882,
'celez': 0.00016949152542372882,
'celle': 0.00016949152542372882,
'celles': 0.00016949152542372882,
'celui': 0.00016949152542372882,
'cendre': 0.00016949152542372882,
'cent': 0.00016949152542372882,
'cents': 0.00016949152542372882,
'cependant': 0.00016949152542372882,
'cercle': 0.00016949152542372882,
'cercueil': 0.00016949152542372882,
'cerf': 0.00016949152542372882,
'certain': 0.00016949152542372882,
'certaine': 0.00016949152542372882,
'certains': 0.00016949152542372882,
'certes': 0.00016949152542372882,
'ces': 0.00016949152542372882,
'cespede': 0.00016949152542372882,
'cessait': 0.00016949152542372882,
'cessant': 0.00016949152542372882,
'cesse': 0.00016949152542372882,
'cesser': 0.00016949152542372882,
'cessez': 0.00016949152542372882,
'cessons': 0.00016949152542372882,
'cet': 0.00016949152542372882,
'cette': 0.00016949152542372882,
'ceux': 0.00016949152542372882,
'chacun': 0.00016949152542372882,
'chagrin': 0.00016949152542372882,
'chaleur': 0.00016949152542372882,
'chambre': 0.00016949152542372882,
'champ': 0.00016949152542372882,
'champions': 0.00016949152542372882,
'champs': 0.00016949152542372882,
'chancelant': 0.00016949152542372882,
'change': 0.00016949152542372882,
'changeant': 0.00016949152542372882,
'changement': 0.00016949152542372882,
'changeons': 0.00016949152542372882,
'changer': 0.00016949152542372882,
'changeât': 0.00016949152542372882,
'changé': 0.00016949152542372882,
'changée': 0.00016949152542372882,

'chansons': 0.00016949152542372882,
'chapeau': 0.00016949152542372882,
'chaque': 0.00016949152542372882,
'char': 0.00016949152542372882,
'charbon': 0.00016949152542372882,
'charge': 0.00016949152542372882,
'chargeant': 0.00016949152542372882,
'charger': 0.00016949152542372882,
'charges': 0.00016949152542372882,
'chargé': 0.00016949152542372882,
'chargés': 0.00016949152542372882,
'charitables': 0.00016949152542372882,
'charité': 0.00016949152542372882,
'charlimin': 0.00016949152542372882,
'charlymin': 0.00016949152542372882,
'charmant': 0.00016949152542372882,
'charmante': 0.00016949152542372882,
'charmants': 0.00016949152542372882,
'charme': 0.00016949152542372882,
'charment': 0.00016949152542372882,
'charmer': 0.00016949152542372882,
'charmes': 0.00016949152542372882,
'charmeur': 0.00016949152542372882,
'charmé': 0.00016949152542372882,
'charmée': 0.00016949152542372882,
'chasse': 0.00016949152542372882,
'chasser': 0.00016949152542372882,
'chasserai': 0.00016949152542372882,
'chassez': 0.00016949152542372882,
'chassé': 0.00016949152542372882,
'chassée': 0.00016949152542372882,
'chaste': 0.00016949152542372882,
'chastes': 0.00016949152542372882,
'chats': 0.00016949152542372882,
'chaude': 0.00016949152542372882,
'chaîne': 0.00016949152542372882,
'chaînes': 0.00016949152542372882,
'chef': 0.00016949152542372882,
'chemin': 0.00016949152542372882,
'chemins': 0.00016949152542372882,
'cher': 0.00016949152542372882,
'cherchai': 0.00016949152542372882,
'cherchais': 0.00016949152542372882,
'cherche': 0.00016949152542372882,
'chercher': 0.00016949152542372882,
'chercherons': 0.00016949152542372882,
'cherchez': 0.00016949152542372882,
'cherchiez': 0.00016949152542372882,
'cherchons': 0.00016949152542372882,
'cherché': 0.00016949152542372882,
'chers': 0.00016949152542372882,
'cheval': 0.00016949152542372882,
'chevalier': 0.00016949152542372882,
'chevaliers': 0.00016949152542372882,
'chevaux': 0.00016949152542372882,
'cheveux': 0.00016949152542372882,
'chevillard': 0.00016949152542372882,
'cheville': 0.00016949152542372882,
'chez': 0.00016949152542372882,
'chien': 0.00016949152542372882,
'choc': 0.00016949152542372882,
'choir': 0.00016949152542372882,
'choisi': 0.00016949152542372882,
'choisie': 0.00016949152542372882,
'choisir': 0.00016949152542372882,
'choisit': 0.00016949152542372882,
'choix': 0.00016949152542372882,
'choque': 0.00016949152542372882,
'choquent': 0.00016949152542372882,

'choquer': 0.00016949152542372882,
'choquez': 0.00016949152542372882,
'choratele': 0.00016949152542372882,
'chose': 0.00016949152542372882,
'choses': 0.00016949152542372882,
'chute': 0.00016949152542372882,
'château': 0.00016949152542372882,
'châtier': 0.00016949152542372882,
'châtiment': 0.00016949152542372882,
'châtiments': 0.00016949152542372882,
'chère': 0.00016949152542372882,
'chères': 0.00016949152542372882,
'chéri': 0.00016949152542372882,
'chérie': 0.00016949152542372882,
'chérir': 0.00016949152542372882,
'chéris': 0.00016949152542372882,
'chérit': 0.00016949152542372882,
'chétif': 0.00016949152542372882,
'chétives': 0.00016949152542372882,
'ci': 0.00016949152542372882,
'ciel': 0.00016949152542372882,
'cieux': 0.00016949152542372882,
'cimeterre': 0.00016949152542372882,
'cinq': 0.00016949152542372882,
'cinquante': 0.00016949152542372882,
'cinquième': 0.00016949152542372882,
'circonstance': 0.00016949152542372882,
'citoyens': 0.00016949152542372882,
'civile': 0.00016949152542372882,
'civilité': 0.00016949152542372882,
'civilités': 0.00016949152542372882,
'clair': 0.00016949152542372882,
'claire': 0.00016949152542372882,
'clairement': 0.00016949152542372882,
'clairs': 0.00016949152542372882,
'clandestine': 0.00016949152542372882,
'clarté': 0.00016949152542372882,
'clartés': 0.00016949152542372882,
'claviche': 0.00016949152542372882,
'clef': 0.00016949152542372882,
'clefs': 0.00016949152542372882,
'climats': 0.00016949152542372882,
'clion': 0.00016949152542372882,
'cloris': 0.00016949152542372882,
'close': 0.00016949152542372882,
'clous': 0.00016949152542372882,
'coadjuteur': 0.00016949152542372882,
'coche': 0.00016949152542372882,
'coeur': 0.00016949152542372882,
'coeurs': 0.00016949152542372882,
'coiffer': 0.00016949152542372882,
'col': 0.00016949152542372882,
'colline': 0.00016949152542372882,
'colorer': 0.00016949152542372882,
'colère': 0.00016949152542372882,
'combat': 0.00016949152542372882,
'combats': 0.00016949152542372882,
'combattant': 0.00016949152542372882,
'combattiez': 0.00016949152542372882,
'combattit': 0.00016949152542372882,
'combattons': 0.00016949152542372882,
'combattraî': 0.00016949152542372882,
'combattre': 0.00016949152542372882,
'combattu': 0.00016949152542372882,
'combattue': 0.00016949152542372882,
'combien': 0.00016949152542372882,
'comble': 0.00016949152542372882,
'combler': 0.00016949152542372882,
'commande': 0.00016949152542372882,

```
'commandement': 0.00016949152542372882,
'commander': 0.00016949152542372882,
'comme': 0.00016949152542372882,
'commence': 0.00016949152542372882,
'commencement': 0.00016949152542372882,
'commencer': 0.00016949152542372882,
'commencez': 0.00016949152542372882,
'commencé': 0.00016949152542372882,
'comment': 0.00016949152542372882,
'commença': 0.00016949152542372882,
'commerce': 0.00016949152542372882,
'commettre': 0.00016949152542372882,
'commis': 0.00016949152542372882,
'commise': 0.00016949152542372882,
'commission': 0.00016949152542372882,
'commode': 0.00016949152542372882,
'commun': 0.00016949152542372882,
'communauté': 0.00016949152542372882,
'commune': 0.00016949152542372882,
'communs': 0.00016949152542372882,
'compagnie': 0.00016949152542372882,
'compagnon': 0.00016949152542372882,
'compagnons': 0.00016949152542372882,
'comparables': 0.00016949152542372882,
'comparaison': 0.00016949152542372882,
'compassion': 0.00016949152542372882,
...}
```

Q5: For each text genre and play, how many terms with a weight strictly larger than 0 do you have in the vector?

This is again simply the lengths of the tf vectors as all words occur at least twice.

In [49]:

```
print("Terms for comedy: ", len(tf_dict_com))
print("Terms for tragedy: ", len(tf_dict_tra))
print("Terms for tragi-comedy: ", len(tf_dict_tra_com))
```

```
Terms for comedy: 14179
Terms for tragedy: 8902
Terms for tragi-comedy: 5288
```

Q6: For each text genre and play, how many terms with a weight strictly equal to 1 do you have in the vector?

With tf, none of them will have a weight of 1, but we can look at the max values:

In [50]:

```
# print(max(tf_dict_com.items(), key=operator.itemgetter(1))[0])
# max(tf_dict_com.values())
[k for k,v in tf_dict_com.items() if v == max(tf_dict_com.values())]
```

Out[50]: ['jou']

In [51]:

```
# max(tf_dict_tra.items(), key=operator.itemgetter(1))[0]
# max(tf_dict_tra.values())
[k for k,v in tf_dict_tra.items() if v == max(tf_dict_tra.values())]
```

Out[51]: ['jou']

In [52]:

```
# max(tf_dict_tra_com.items(), key=operator.itemgetter(1))[0]
# max(tf_dict_tra_com.values())
[k for k,v in tf_dict_tra_com.items() if v == max(tf_dict_tra_com.values())]
```

Out[52]: ['don']