

Midterm

October 29, 2023

```
[1]: print("""
Riker Wachtler
29 October 2023
NLP Midterm Project
https://github.com/RikerW/midterm-nlp
""")
```

Riker Wachtler
29 October 2023
NLP Midterm Project
<https://github.com/RikerW/midterm-nlp>

```
[2]: # relevant imports/initializations, loosely sorted by category

import nltk
from nltk.corpus import stopwords
from nltk.tokenize import sent_tokenize, word_tokenize
from nltk.probability import FreqDist
nltk.download('punkt')
nltk.download('stopwords')
stop_words = set(stopwords.words('english'))

from sklearn.feature_extraction.text import TfidfVectorizer
import numpy as np
import re

from pypdf import PdfReader

import matplotlib.pyplot as plt
%matplotlib inline

from wordcloud import WordCloud

from textblob import TextBlob

import spacy
```

```
from spacy import displacy
nlp = spacy.load('en_core_web_sm')
```

```
[nltk_data] Downloading package punkt to
[nltk_data]   C:\Users\riker\AppData\Roaming\nltk_data...
[nltk_data]   Package punkt is already up-to-date!
[nltk_data] Downloading package stopwords to
[nltk_data]   C:\Users\riker\AppData\Roaming\nltk_data...
[nltk_data]   Package stopwords is already up-to-date!
```

```
[3]: print("""
This is demonstrating various Natural Language Processing techniques we've
↳learned so far in this class.

I got a small sample of some of my friend Conor's poetry, taken from some
↳portfolio submission he did a bit ago. It's 8 pages total, though I had to
↳skip the 7th since it's doesn't have any real text on it.

I read the text from a PDF directly via pypdf, including the titles for each
↳poem & the name at the start, and then do a little immediate clean-up for
↳some of the punctuation. It's mostly just putting periods at the end of
↳stanzas without punctuation, since otherwise the tokenizer gets a little
↳confused on what to break on - I didn't have to do this, but it seemed more
↳conducive to interesting data.

After, the sentences are tokenized and some basic exploratory processing is
↳done. Following that, the text is cleaned up - non-alpha words are purged,
↳all words are cast to lowercase, etc.

Then, TF-IDF vectorization is done to get some of the most "important"
↳sentences in the text - the top 3 are printed. It's a little biased towards
↳longer sentences but it's not like the text isn't full of some of those.

Following the TF-IDF, less technical things are done - a WordCloud on both the
↳original and cleaned up text, translating the original text to spanish,
↳using displacy to identify important words in the original/cleaned up text,
↳and then finally POS tagging & FreqDist counting of words in the original/
↳cleaned up texts.
""")
```

This is demonstrating various Natural Language Processing techniques we've learned so far in this class.

I got a small sample of some of my friend Conor's poetry, taken from some portfolio submission he did a bit ago. It's 8 pages total, though I had to skip the 7th since it's doesn't have any real text on it.

I read the text from a PDF directly via pypdf, including the titles for each poem & the name at the start, and then do a little immediate clean-up for some of the punctuation. It's mostly just putting periods at the end of stanzas without punctuation, since otherwise the tokenizer gets a little confused on what to break on - I didn't have to do this, but it seemed more conducive to interesting data.

After, the sentences are tokenized and some basic exploratory processing is done. Following that, the text is cleaned up - non-alpha words are purged, all words are cast to lowercase, etc.

Then, TF-IDF vectorization is done to get some of the most "important" sentences in the text - the top 3 are printed. It's a little biased towards longer sentences but it's not like the text isn't full of some of those.

Following the TF-IDF, less technical things are done - a WordCloud on both the original and cleaned up text, translating the original text to spanish, using displacy to identify important words in the original/cleaned up text, and then finally POS tagging & FreqDist counting of words in the original/cleaned up texts.

```
[4]: # Read the poetry from a PDF - ignore the 7th page, it has an image that
      ↳ inspired one of the poems, so skip it

reader = PdfReader("data/obrien_portfolio.pdf")
text = ""
for page in enumerate(reader.pages):
    if (page[0] == 6):
        continue; # that page has an image & url on it, skip it
    text += page[1].extract_text()[:-4] + ".\n\n"

text = text[:-4]

# Process the data a bit - cleaning up quotes because I felt like it, turning
↳ em-dashes into periods
# and then putting periods after stanzas without punctuation.
# due to the nature of this poetry, otherwise certain stanzas
# have no punctuation whatsoever so they all get lumped into one "sentence"
# lastly, each line gets its funky whitespace stripped, and any empty lines get
↳ skipped. len==1 catches haiku numbers

text = re.sub(r'[""]', r'', text)
text = re.sub(r"[-]", r".", text)
text = re.sub(r"([!?.,:;]) \n\n", r"\1.\n", text)
text_arr = [line.strip() for line in text.split("\n") if len(line.strip()) > 1]
text = "\n".join(text_arr)
```

```
print(text)
```

Poetry Portfolio

Conor O'Brien

Haiku

wet feet, hot-sand caked
tracing yesterday's footprints
'til winter takes hold.
as the sun sinks low
we chase parasite squirrels
while hummingbirds drink.
dusk air, dance freely!
navigate new jersey streets,
be my melody.
a captive pine stands
shaking, tall amid its kin,
now ours for Christmas.
in yellow canoe,
paddles stuck in mud and grime,
we pause for a breath.
heaven's tsunami
crashes over high mountains.
fitting for a walk.

Blank Verse

these my thoughts.
anonymized by arbored trails near roads
unkempt i lumber far away from these
my thoughts of you and what we could have been;
unstill, you pull me back, and still i trudge
onwards past the empty lots, reflecting:
if you could name a place where i belong
then maybe i could reach out to the past
and give myself some tips; that maybe then
i'd find a straighter path to where i am
supposed to go from here.
if you could lead us now and teach me how
to dance upon the path of those regrets,
then maybe i could come to life and live
a little freer than before; my friend,
each day's a taxing exercise in how
to navigate the maze of spike-like glass
which spiderwebs all that i'll ever do,
where one mistake is fatal in my mind.
i lift my eyes to taste the air awhile
and avoid asphalt asphyxiation
alongside sidewalk streetlamps burning
behind the forest's cover, like eyes on fire,

a procession of tin-can saints, ever fixed
in place to guide the lost at night to home;
no sign of you there in that train, you struck
a different path out through the wilder weeds.
thus, resolute, i turn to go, and find
that others wander now and here as well,
beneath the gibbous sky; never alone,
they howl delight with souls also like eyes
that brim to match the sun; to them, to wander
is merely transit.

my feet no lighter than before, i grasp
a momentary hope in hand: that if
one day we meet each other's eyes again,
maybe we could wander a bit as two,
and lose our minds to cloudless winds offshore.

Surrealism

in the through it all.

she speaks to me, sweet Vanity,
looking straight through the vanity
i turn to face the mirror's speech
alarmed, awake, she averts her gaze.

"keep your eyes to yourself,
they're sticky and gross,"
she claims as she drags herself
away, wringing her neck and giggling,
careening into the sky,
convulsing with laughter.

or was it despair?

she never told me which.

i shrug and stand, slamming shut
the three-fold mirrors, its rosy glass
melting slowly to the floor
before congealing into shards.

having had enough of Vanity's
looking glass apartment,
i sink through her floorboards
passing through the ground.

"my eyes are really, panoramic,"

i think to myself,

beholding inky dirt

punctuated by peculiar pictures:

behold, the plum was a sock.

or was it a peach?.

darned with vain juices

mailing itself à Paris.

there, i watch the women go and come,
talking to michelangelo's son,

"when did sweet alfred leave?

the hushed one poets unweave?".
outside, a turtle rests,
smiling on a damp vuitton with receipt
one woman had discarded
in her pursuit of the chi.
Blues/Rap
where you lay.
heart attack, play it back,
blood contract won't retract,
masses act, violent tact,
news reacts, homes ransacked.
flash refracts on smashed glass,
one shot cracks.
family maimed, mourning fame,
TVs play Deaths' new claims;
all the same, morning came,
found our cries strained.
i could ask why you trashed
our home back then.
you always blamed the past,
ate the scraps from newscasts,
too scared to catch somber jazz
and rehash his urn's ash.
open pad, this i'd add:
i wish you saw a shrink.
your rage amassed can't outlast
life's lambaste, grey forecast,
grief comes back, comes back fast,
smacks you with marred flashbacks.
truth's all there, you don't care,
took that dare, shunned the world.
loved your pain, blood tears rain,
laughed insane, dad's to blame,
colt's disdain, roulette trained,
bullseye aim shook your frame.
ringing air, bleeding hair,
sirens blare where you lay.
Slang
real talk.
i dunno really
it's hard to think about
when bullshitting myself's
easier than breathing.
i guess i just
hate it here sometimes
goddamn the world
and everything in it.
sorry for the GD

i'm kinda heated
there's just so much
to resent.
i don't want no binge love
no pregamed fling
no conked out high
nunathat.
i'm not slang enough to hang
y'all got pocket aces
i gotta all-in on my three and six
when i don't sit at the table often enough.
shoving zen to the backburner
ain't no way to live
yet ain't got much time
to do much else.
ya feel me?.
Ekphrastic
just beyond the tree.
the house lights betray that father's away,
skimming the graves for gold;
his daughters know well that their father can't tell
the riches from the mold.
lesser trees burn midnight green,
their leaves plume up between
a greener roof and the sickle moon.
they're all missed by peerless Cypress
who can't resist but bisect the abyss,
posing as old Babel's ghost.
although the front door never offers more
than a view of that vain tree,
the children prefer the wispier firs
which still know honesty.
hay fields wave the carriage away
down the road paved with rivulet clay
to father and his friend.
they hail its driver and
father faux limps over on his shovel,
smelling of booze.
he asks the driver for medical supplies.
they respond they had none.
father smiles to his friend.
his friend smiles back
and aims his pistol..
Epistolary
to the wolves & birds.
may your corpse rot forever.
you'll never know how
for years i'd wake each day

with a matchstick tribute
 to the ashtray with your name
 stapled to the back
 seventeen times,
 one for each year you've haunted me.
 you thought i'd be better off alone
 (like the lone wolf you are)
 that i should be proud that
 i could count on one hand
 the times i remember crying,
 as if stone-facedness were
 a reasonable coping mechanism.
 i suppose
 we must bury our past selves.
 i pray you remain dead;
 prithee, fuck off.
 let me feel the breeze on my face,
 under my shirt and in my nostrils;
 let me stand a moment here and
 take note of those who sing to me:
 who are you all?
 perched proud,
 diving and chirruping.
 i want to know your names,
 where you came from,
 and where you go when it's not so cold.
 not that we don't mind a bit of winter.
 could you take me with you?
 i can't fly, but
 dammit! i'll

```

[5]: # Tokenize the sentences :)
      # The info is the closest I have to EDA, since there's not too many preliminary
      ↪statistics for text
sentences = sent_tokenize(text)
print(f"Total number of sentences: {len(sentences)}")
print(f"Longest sentence: {max(sentences, key=len)}")
print(f"Shortest sentence: {min(sentences, key=len)}")
print()
print("All sentences:")
for s in sentences: print(s)
  
```

Total number of sentences: 70
 Longest sentence: anonymized by arbored trails near roads
 unkempt i lumber far away from these
 my thoughts of you and what we could have been;
 unstill, you pull me back, and still i trudge
 onwards past the empty lots, reflecting:

if you could name a place where i belong
then maybe i could reach out to the past
and give myself some tips; that maybe then
i'd find a straighter path to where i am
supposed to go from here.
Shortest sentence: i'll

All sentences:

Poetry Portfolio

Conor O'Brien

Haiku

wet feet, hot-sand caked
tracing yesterday's footprints
'til winter takes hold.
as the sun sinks low
we chase parasite squirrels
while hummingbirds drink.
dusk air, dance freely!
navigate new jersey streets,
be my melody.
a captive pine stands
shaking, tall amid its kin,
now ours for Christmas.
in yellow canoe,
paddles stuck in mud and grime,
we pause for a breath.
heaven's tsunami
crashes over high mountains.
fitting for a walk.

Blank Verse

these my thoughts.
anonymized by arbored trails near roads
unkempt i lumber far away from these
my thoughts of you and what we could have been;
unstill, you pull me back, and still i trudge
onwards past the empty lots, reflecting:
if you could name a place where i belong
then maybe i could reach out to the past
and give myself some tips; that maybe then
i'd find a straighter path to where i am
supposed to go from here.
if you could lead us now and teach me how
to dance upon the path of those regrets,
then maybe i could come to life and live
a little freer than before; my friend,
each day's a taxing exercise in how
to navigate the maze of spike-like glass
which spiderwebs all that i'll ever do,

where one mistake is fatal in my mind.
i lift my eyes to taste the air awhile
and avoid asphalt asphyxiation
alongside sidewalk streetlamps burning
behind the forest's cover, like eyes on fire,
a procession of tin-can saints, ever fixed
in place to guide the lost at night to home;
no sign of you there in that train, you struck
a different path out through the wilder weeds.
thus, resolute, i turn to go, and find
that others wander now and here as well,
beneath the gibbous sky; never alone,
they howl delight with souls also like eyes
that brim to match the sun; to them, to wander
is merely transit.
my feet no lighter than before, i grasp
a momentary hope in hand: that if
one day we meet each other's eyes again,
maybe we could wander a bit as two,
and lose our minds to cloudless winds offshore.

Surrealism

in the through it all.
she speaks to me, sweet Vanity,
looking straight through the vanity
i turn to face the mirror's speech
alarmed, awake, she averts her gaze.
"keep your eyes to yourself,
they're sticky and gross,"
she claims as she drags herself
away, wringing her neck and giggling,
careening into the sky,
convulsing with laughter.
or was it despair?
she never told me which.
i shrug and stand, slamming shut
the three-fold mirrors, its rosy glass
melting slowly to the floor
before congealing into shards.
having had enough of Vanity's
looking glass apartment,
i sink through her floorboards
passing through the ground.
"my eyes are really, panoramic,"
i think to myself,
beholding inky dirt
punctuated by peculiar pictures:
behold, the plum was a sock.
or was it a peach?.

darned with vain juices
mailing itself à Paris.
there, i watch the women go and come,
talking to michelangelo's son,
"when did sweet alfred leave?
the hushed one poets unweave?".
outside, a turtle rests,
smiling on a damp vuitton with receipt
one woman had discarded
in her pursuit of the chi.
Blues/Rap
where you lay.
heart attack, play it back,
blood contract won't retract,
masses act, violent tact,
news reacts, homes ransacked.
flash refracts on smashed glass,
one shot cracks.
family maimed, mourning fame,
TVs play Deaths' new claims;
all the same, morning came,
found our cries strained.
i could ask why you trashed
our home back then.
you always blamed the past,
ate the scraps from newscasts,
too scared to catch somber jazz
and rehash his urn's ash.
open pad, this i'd add:
i wish you saw a shrink.
your rage amassed can't outlast
life's lambaste, grey forecast,
grief comes back, comes back fast,
smacks you with marred flashbacks.
truth's all there, you don't care,
took that dare, shunned the world.
loved your pain, blood tears rain,
laughed insane, dad's to blame,
colt's disdain, roulette trained,
bullseye aim shook your frame.
ringing air, bleeding hair,
sirens blare where you lay.
Slang
real talk.
i dunno really
it's hard to think about
when bullshitting myself's
easier than breathing.

i guess i just
hate it here sometimes
goddamn the world
and everything in it.
sorry for the GD
i'm kinda heated
there's just so much
to resent.
i don't want no binge love
no pregamed fling
no conked out high
nunathat.
i'm not slang enough to hang
y'all got pocket aces
i gotta all-in on my three and six
when i don't sit at the table often enough.
shoving zen to the backburner
ain't no way to live
yet ain't got much time
to do much else.
ya feel me?.

Ekphrastic
just beyond the tree.
the house lights betray that father's away,
skimming the graves for gold;
his daughters know well that their father can't tell
the riches from the mold.
lesser trees burn midnight green,
their leaves plume up between
a greener roof and the sickle moon.
they're all missed by peerless Cypress
who can't resist but bisect the abyss,
posing as old Babel's ghost.
although the front door never offers more
than a view of that vain tree,
the children prefer the wispier firs
which still know honesty.
hay fields wave the carriage away
down the road paved with rivulet clay
to father and his friend.
they hail its driver and
father faux limps over on his shovel,
smelling of booze.
he asks the driver for medical supplies.
they respond they had none.
father smiles to his friend.
his friend smiles back
and aims his pistol..

Epistolary
to the wolves & birds.
may your corpse rot forever.
you'll never know how
for years i'd wake each day
with a matchstick tribute
to the ashtray with your name
stapled to the back
seventeen times,
one for each year you've haunted me.
you thought i'd be better off alone
(like the lone wolf you are)
that i should be proud that
i could count on one hand
the times i remember crying,
as if stone-facedness were
a reasonable coping mechanism.
i suppose
we must bury our past selves.
i pray you remain dead;
prithee, fuck off.
let me feel the breeze on my face,
under my shirt and in my nostrils;
let me stand a moment here and
take note of those who sing to me:
who are you all?
perched proud,
diving and chirruping.
i want to know your names,
where you came from,
and where you go when it's not so cold.
not that we don't mind a bit of winter.
could you take me with you?
i can't fly, but
dammit!
i'll

```
[6]: # Define preprocessing function, which uses re.sub to clean non-alphanumeric  

      ↳ characters out of the text,  

      # then cast to lowercase & skip any other non-alpha words that got missed  

      ↳ somehow  

      def preprocess_text(text):  

          words = word_tokenize(text)  

          cleaned_words = word_tokenize(re.sub(r'^a-zA-Z', ' ', text))  

          cleaned_words = [word.lower() for word in cleaned_words if word.isalnum()  

          ↳ and word not in stop_words]
```

```

    return ' '.join(cleaned_words)

# Create our cleaned up list of sentences & text, for use with TF-IDF
→vectorization and then later fun visuals
prepped_sentences = list(map(preprocess_text, sentences))
cleaned_text = '\n'.join(prepped_sentences)
print(cleaned_text)

```

poetry portfolio conor o brien haiku wet feet hot sand caked tracing yesterday
 footprints til winter takes hold
 sun sinks low chase parasite squirrels hummingbirds drink
 dusk air dance freely
 navigate new jersey streets melody
 captive pine stands shaking tall amid kin christmas
 yellow canoe paddles stuck mud grime pause breath
 heaven tsunami crashes high mountains
 fitting walk
 blank verse thoughts
 anonymized arbored trails near roads unkempt lumber far away thoughts could
 unstill pull back still trudge onwards past empty lots reflecting could name
 place belong maybe could reach past give tips maybe find straighter path
 supposed go
 could lead us teach dance upon path regrets maybe could come life live little
 freer friend day taxing exercise navigate maze spike like glass spiderwebs ever
 one mistake fatal mind
 lift eyes taste air awhile avoid asphalt asphyxiation alongside sidewalk
 streetlamps burning behind forest cover like eyes fire procession tin saints
 ever fixed place guide lost night home sign train struck different path wilder
 weeds
 thus resolute turn go find others wander well beneath gibbous sky never alone
 howl delight souls also like eyes brim match sun wander merely transit
 feet lighter grasp momentary hope hand one day meet eyes maybe could wander bit
 two lose minds cloudless winds offshore
 surrealism
 speaks sweet vanity looking straight vanity turn face mirror speech alarmed
 awake averts gaze
 keep eyes sticky gross claims drags away wringing neck giggling careening sky
 convulsing laughter
 despair
 never told
 shrug stand slamming shut three fold mirrors rosy glass melting slowly floor
 congealing shards
 enough vanity looking glass apartment sink floorboards passing ground
 eyes really panoramic think beholding inky dirt punctuated peculiar pictures
 behold plum sock
 peach
 darned vain juices mailing paris
 watch women go come talking michelangelo son sweet alfred leave

hushed one poets unweave
outside turtle rests smiling damp vuitton receipt one woman discarded pursuit
chi
blues rap lay
heart attack play back blood contract retract masses act violent tact news
reacts homes ransacked
flash refracts smashed glass one shot cracks
family maimed mourning fame tvs play deaths new claims morning came found cries
strained
could ask trashed home back
always blamed past ate scraps newscasts scared catch somber jazz rehash urn ash
open pad add wish saw shrink
rage amassed outlast life lambaste grey forecast grief comes back comes back
fast smacks marred flashbacks
truth care took dare shunned world
loved pain blood tears rain laughed insane dad blame colt disdain roulette
trained bullseye aim shook frame
ringing air bleeding hair sirens blare lay
slang real talk
dunno really hard think bullshitting easier breathing
guess hate sometimes goddamn world everything
sorry gd kinda heated much resent
want binge love pregamed fling conked high nunathat
slang enough hang got pocket aces got ta three six sit table often enough
shoving zen backburner way live yet got much time much else
ya feel
ekphrastic beyond tree
house lights betray father away skimming graves gold daughters know well father
tell riches mold
lesser trees burn midnight green leaves plume greener roof sickle moon
missed peerless cypress resist bisect abyss posing old babel ghost
although front door never offers view vain tree children prefer wispiers
still know honesty
hay fields wave carriage away road paved rivulet clay father friend
hail driver father faux limps shovel smelling booze
asks driver medical supplies
respond none
father smiles friend
friend smiles back aims pistol
epistolary wolves birds
may corpse rot forever
never know years wake day matchstick tribute ashtray name stapled back seventeen
times one year haunted
thought better alone like lone wolf proud could count one hand times remember
crying stone facedness reasonable coping mechanism
suppose must bury past selves
pray remain dead prithee fuck
let feel breeze face shirt nostrils let stand moment take note sing

```
perched proud diving chirruping
want know names came go cold
mind bit winter
could take
fly dammit
```

```
[7]: # TF-IDF vectorize this stuff - specifically the tokenized sentences, but not
      ↪ cleaned up ones
      # We're going to use this to find the most common words in conor's poetry :)
      tfidf = TfidfVectorizer()
      tfidf_matrix = tfidf.fit_transform(sentences)
      tfidf_scores = tfidf_matrix.sum(axis=1)
      print(tfidf_scores)
```

```
[4.12138604]
[3.39360051]
[1.99536585]
[2.62484596]
[3.44703519]
[3.39095175]
[2.44721742]
[1.71053604]
[2.21301886]
[7.17849045]
[7.03251175]
[6.8683727 ]
[5.74123405]
[5.79385864]
[2.37086184]
[4.2049492 ]
[5.04024229]
[1.98379383]
[2.21832478]
[4.40523922]
[3.74388884]
[4.38223042]
[1.98379383]
[2.6264567 ]
[3.88633045]
[2.14171083]
[4.2660167 ]
[2.18772957]
[4.10340509]
[2.79770193]
[4.18516571]
[2.95257446]
[4.37880074]
```



```

[2.79090968]
[4.05926931]
[3.36092022]
[4.30209011]
[2.95761001]
[1.73044403]
[3.55159953]
[3.3239961 ]
[3.36385363]
[3.08804878]
[4.55835356]
[3.64989445]
[1.70988049]
[2.15561497]
[4.46691743]
[3.92231032]
[4.29156801]
[4.54024232]
[4.03686076]
[3.81096794]
[2.56417358]
[1.94784766]
[2.19174694]
[2.53357566]
[2.09954889]
[2.22116776]
[5.06826641]
[5.35969289]
[2.62718096]
[2.60587583]
[4.67051387]
[2.18290234]
[3.86027924]
[2.80157834]
[2.19603407]
[1.98864222]
[1.      ]]

```

```

[8]: top_sentence_indices = np.argsort(tfidf_scores, axis=0)[-3:]
    for index in top_sentence_indices:
        val = sentences[index[0,0]]
        print(val, "\n")

```

```

i lift my eyes to taste the air awhile
and avoid asphalt asphyxiation
alongside sidewalk streetlamps burning
behind the forest's cover, like eyes on fire,
a procession of tin-can saints, ever fixed

```

in place to guide the lost at night to home;
no sign of you there in that train, you struck
a different path out through the wilder weeds.

if you could lead us now and teach me how
to dance upon the path of those regrets,
then maybe i could come to life and live
a little freer than before; my friend,
each day's a taxing exercise in how
to navigate the maze of spike-like glass
which spiderwebs all that i'll ever do,
where one mistake is fatal in my mind.

anonymized by arbored trails near roads
unkempt i lumber far away from these
my thoughts of you and what we could have been;
unstill, you pull me back, and still i trudge
onwards past the empty lots, reflecting:
if you could name a place where i belong
then maybe i could reach out to the past
and give myself some tips; that maybe then
i'd find a straighter path to where i am
supposed to go from here.

```
[9]: # Word cloud for original text
wordcloud = WordCloud(width=800, height=400, background_color="black").
    ↪generate(text)

# Display the word cloud using matplotlib
plt.figure(figsize=(10, 5))
plt.imshow(wordcloud, interpolation="bilinear")
plt.axis("off")
plt.show()
```



```
[11]: # Translate the text to Spanish
blob = TextBlob(text)
translated_blob = blob.translate('en', 'es')
print(translated_blob)
```

Cartera de poesía
Conor O'Brien
Haiku
Pies mojados, capado de arena caliente
rastrear las huellas de ayer
'Hasta el invierno se apodera.
A medida que el sol se hunde
perseguiamos ardillas de parásito
mientras los colibríes beben.
¡Aire anochecer, baila libremente!
navegar por las calles de Nueva Jersey,
ser mi melodía.
Un pino cautivo se encuentra
temblando, alto en medio de sus parientes,
Ahora el nuestro para Navidad.
en canoa amarilla,
paletas atrapadas en barro y mugre,
Hacemos una pausa por respirar.
Tsunami del cielo
se bloquea sobre las altas montañas.
apropiado para caminar.
Verso en blanco
Estos mis pensamientos.
anonimizado por senderos arboriosos cerca de las carreteras
desagradable, me quedo lejos de estos
Mis pensamientos sobre ti y lo que podríamos haber sido;
Desconectado, me retiras, y aún así me puse caminando
En adelante, más allá de los lotes vacíos, reflejando:
Si pudieras nombrar un lugar donde pertenezco
Entonces tal vez podría comunicarme con el pasado
Y dame algunos consejos; que tal vez entonces
Encontraría un camino más recto hacia donde estoy
se supone que debe ir desde aquí.
Si pudieras guiarnos ahora y enseñarme cómo
bailar en el camino de esos arrepentimientos,
Entonces tal vez podría dar vida y vivir
un poco más libre que antes; mi amigo,
Cada día es un ejercicio exigente sobre cómo
para navegar el laberinto de vidrio de espiga
¿Qué parpadeo de araña todo lo que haré?
Donde un error es fatal en mi mente.
levanto los ojos para probar el aire un rato
y evitar asfixia de asfalto

Junto a las farolas de la acera ardiendo
 Detrás de la cubierta del bosque, como ojos en fuego,
 una procesión de santos de lata, siempre fijo
 en su lugar para guiar a los perdidos por la noche a casa;
 No hay señal de ti allí en ese tren, golpeaste
 Un camino diferente a través de las malas hierbas más salvajes.
 Por lo tanto, resuelto, me giro para ir y encuentro
 que otros deambulan ahora y aquí también
 Debajo del cielo giboso; nunca solo,
 Aullan deleite con almas también como los ojos
 ese borde para que coincida con el sol; a ellos, para pasear
 es simplemente tránsito.
 mis pies no más ligero que antes, agarro
 Una esperanza momentánea en la mano: que si
 Un día nos volvemos a ver a los ojos,
 Tal vez podríamos deambular un poco como dos,
 y perder la mente por los vientos sin nubes en la costa.
 Surrealismo
 a través de todo.
 Ella me habla, dulce vanidad,
 mirando directamente a través del tocador
 Me giro para enfrentar el discurso del espejo
 alarmada, despierta, evita la mirada.
 "Mantén tus ojos para ti mismo,
 Son pegajosos y asquerosos "
 Ella afirma mientras se arrastra
 lejos, retorciéndose el cuello y riendo,
 cargándose en el cielo,
 convulsionando de risa.
 ¿O fue desesperado?
 Ella nunca me dijo cuál.
 Me encogí de hombros y me quedo con golpes de cerrado
 los espejos triple, su vidrio rosado
 derretirse lentamente al piso
 antes de congelar en fragmentos.
 Habiendo tenido suficiente de Vanity's
 apartamento de vidrio,
 Me hundo en las tablas de su piso
 pasando por el suelo.
 "Mis ojos son realmente panorámicos",
 Pienso para mí,
 Contemplado de tierra de tinta
 puntuado por imágenes peculiares:
 He aquí, la ciruela era un calcetín.
 ¿O fue un durazno?
 maldito con vanos jugos
 enviando por correo a París.
 Allí veo a las mujeres irse y venir

hablando con el hijo de Miguel Ángel,
"¿Cuándo se fue el dulce Alfred?
¿Los poetas silenciosos desencadenaron? ".
Afuera, una tortuga descansa,
sonriendo en un Vuitton húmedo con recibo
Una mujer había descartado
en su búsqueda del chi.
Blues/rap
Donde te acostas.
Ataque al corazón, reproducirlo,
El contrato de sangre no se retraerá,
Ley de masas, tacto violento,
Las noticias reaccionan, casas saqueadas.
Flash se refracta sobre vidrios rotos,
Un disparo se grieta.
Familia mutilada, fama de luto,
Los televisores tocan las nuevas afirmaciones de las muertes;
de todos modos, llegó la mañana,
Encontraron nuestros gritos tensados.
Podría preguntar por qué destrozaste
Nuestra casa en aquel entonces.
Siempre culpaste al pasado
Comió los restos de los noticieros,
Demasiado asustado para atrapar a Sombrar Jazz
y rehacer la ceniza de su urna.
Almohadilla abierta, esto agregaría:
Ojalá vieras un encogimiento.
Tu ira amasada no puede durar más allá
Lambaste de la vida, pronóstico gris,
El dolor regresa, regresa rápido,
Te golpea con flashbacks de vías.
la verdad está todo ahí, no te importa,
Tomó ese desafío, rechazó el mundo.
Me encantó tu dolor, la sangre de las lágrimas, la lluvia,
se rió loco, papá tiene la culpa,
desdén de Colt, ruleta entrenada,
Bullseye AIM sacudió tu marco.
Aire que suena, pelo sangrante,
Sirens Bloure donde yacía.
Jerga
charla honesta.
No sé realmente
Es difícil pensar en
Cuando me está golpeando a mí mismo
más fácil que respirar.
Supongo que yo solo
Odio aquí a veces
maldita sea el mundo

y todo en él.
Perdón por el GD
Estoy un poco calentado
Hay tanto
a resentir.
No quiero ningún amor de amor
Sin fling pregamado
No se cazó alto
Nunathat.
No soy lo suficientemente argot como para colgar
Todos ustedes tienen ases de bolsillo
Tengo que todo en mis tres y seis
Cuando no me siento en la mesa con la frecuencia suficiente.
Empujando a Zen al retorno
No hay forma de vivir
Sin embargo, no tiene mucho tiempo
hacer mucho más.
¿Me sientes?.
Ekphrastic
justo más allá del árbol.
las luces de la casa traicionan que el padre está lejos,
Hechiciendo las tumbas por oro;
sus hijas saben bien que su padre no puede decir
Las riquezas del molde.
Los árboles menores se queman a la medianoche verde,
sus hojas se llenan entre
Un techo más verde y la luna falciforme.
Todos son extrañados por Peerless Cypress
quien no puede resistirse pero bisecar el abismo,
Posando como el fantasma del viejo Babel.
Aunque la puerta principal nunca ofrece más
que una vista de ese vano árbol,
los niños prefieren los abetos más obedientes
que todavía saben la honestidad.
Los campos de heno agitan el carruaje
en el camino pavimentado con arcilla de rivulet
al padre y su amigo.
Ellos saludan a su conductor y
El padre falsa se interpone en su pala,
oliendo a alcohol.
Le pregunta al conductor suministros médicos.
Responden que no tuvieron ninguno.
El padre le sonríe a su amigo.
Su amigo le devuelve la sonrisa
y apunta su pistola ...
Epistolar
a los lobos y pájaros.
Que tu cadáver se pudra para siempre.

Nunca sabrás como
 Durante años me despertaba cada día
 con un tributo a fósforo
 al cenicero con tu nombre
 Grapado en la parte posterior
 diecisiete veces,
 Uno por cada año me has perseguido.
 Pensaste que estaría mejor solo
 (como el lobo solitario que eres)
 que debería estar orgulloso de que
 Podría contar con una mano
 Las veces que recuerdo llorar,
 Como si la caracol de piedra fuera
 un mecanismo de afrontamiento razonable.
 Supongo
 Debemos enterrarnos a nosotros mismos.
 Rezo para que permanezcas muerto;
 Prithee, joder.
 Déjame sentir la brisa en mi cara
 debajo de mi camisa y en mis fosas nasales;
 Déjame soportar un momento aquí y
 Tome nota de los que me cantan:
 ¿Quiénes son todos?
 encaramado orgulloso,
 Buceo y encogido.
 Quiero saber tus nombres,
 De donde vienes,
 Y a dónde vas cuando no hace tanto frío.
 No es que no nos importe un poco de invierno.
 ¿Podrías llevarme contigo?
 No puedo volar, pero
 ¡Maldita sea! enfermo

```

[12]: colors = {'PERSON': 'linear-gradient(90deg, #aa9cfc, #fc9ce7)',
               'DATE': 'linear-gradient(90deg, #34ebb4, #9c34eb)',
               'ORG': 'linear-gradient(90deg, #eb3d34, #ebe834)',
               'GPE': 'linear-gradient(90deg, #eb34d9, #bd0636)',
               'TIME': 'linear-gradient(90deg, #9bbd06, #06bd89)',
               'NORP': 'linear-gradient(90deg, #067dbd, #1206bd)',
               'CARDINAL': 'linear-gradient(90deg, #06b4bd, #de0942)',

               }

options = {'ents': ['PERSON', 'DATE', 'ORG', 'GPE', 'TIME', 'NORP', 'CARDINAL'], 'colors': colors}
  
```



```
[13]: # Display the original text, and label
doc = nlp(text)
displacy.render(doc, style='ent', jupyter=True, options = options)
```

<IPython.core.display.HTML object>

```
[14]: doc = nlp(cleaned_text)
displacy.render(doc, style='ent', jupyter=True, options = options)
```

<IPython.core.display.HTML object>

```
[15]: words = word_tokenize(text)
print(f"Total words in original text: {len(words)}")
print("Parts of speech for each word (dups removed):")
for i in sorted(list(set(nltk.pos_tag(words))): print(i)
```

Total words in original text: 1405

Parts of speech for each word (dups removed):

```
('!', '.')
```

```
('&', 'CC')
```

```
('\"', '\"')
```

```
('(', '(')
```

```
(')', ')')
```

```
(',' , ',')
```

```
('.', '.')
```

```
('..', 'NNP')
```

```
(':', ':')
```

```
(';', ':')
```

```
('?', '.')
```

```
('Babel', 'NNP')
```

```
('Blank', 'NNP')
```

```
('Blues/Rap', 'NNP')
```

```
('Brien', 'NNP')
```

```
('Christmas', 'NNP')
```

```
('Conor', 'NNP')
```

```
('Cypress', 'NNP')
```

```
('Deaths', 'NNP')
```

```
('Ekphrastic', 'JJ')
```

```
('Epistolary', 'NNP')
```

```
('GD', 'NNP')
```

```
('Haiku', 'NNP')
```

```
('O', 'NNP')
```

```
('Paris', 'NNP')
```

```
('Poetry', 'NNP')
```

```
('Portfolio', 'NNP')
```

```
('Slang', 'NNP')
```

```
('Surrealism', 'NN')
```

```
('TVs', 'NNP')
```

```
('Vanity', 'NNP')
```

('Verse', 'NNP')
('``', '``')
('a', 'DT')
('about', 'IN')
('abyss', 'NN')
('aces', 'NNS')
('act', 'NN')
('add', 'NN')
('again', 'RB')
('aim', 'NN')
('aims', 'VBZ')
('ain', 'JJ')
('ain', 'NN')
('air', 'NN')
('alarmed', 'VBD')
('alfred', 'JJ')
('all', 'DT')
('all', 'PDT')
('all-in', 'NN')
('alone', 'RB')
('alongside', 'IN')
('also', 'RB')
('although', 'IN')
('always', 'RB')
('am', 'VBP')
('amassed', 'VBD')
('amid', 'IN')
('and', 'CC')
('anonymized', 'VBN')
('apartment', 'NN')
('arbored', 'JJ')
('are', 'VBP')
('as', 'IN')
('as', 'RB')
('ash', 'NN')
('ashtray', 'NN')
('ask', 'VB')
('asks', 'VBZ')
('asphalt', 'NN')
('asphyxiation', 'NN')
('at', 'IN')
('ate', 'VB')
('attack', 'NN')
('averts', 'VBZ')
('avoid', 'JJ')
('awake', 'NN')
('away', 'RB')
('awhile', 'NN')

('back', 'RB')
('backburner', 'NN')
('be', 'VB')
('been', 'VBN')
('before', 'IN')
('before', 'RB')
('behind', 'IN')
('behold', 'NN')
('beholding', 'VBG')
('belong', 'VBP')
('beneath', 'IN')
('betray', 'NN')
('better', 'JJR')
('between', 'IN')
('beyond', 'IN')
('binge', 'NN')
('birds', 'NNS')
('bisect', 'VBP')
('bit', 'NN')
('blame', 'VB')
('blamed', 'VBD')
('blare', 'NN')
('bleeding', 'VBG')
('blood', 'NN')
('booze', 'NN')
('breath', 'NN')
('breathing', 'VBG')
('breeze', 'NN')
('brim', 'VBP')
('bullseye', 'JJ')
('bullshitting', 'VBG')
('burn', 'VBP')
('burning', 'VBG')
('bury', 'VB')
('but', 'CC')
('by', 'IN')
('caked', 'VBD')
('came', 'VBD')
('can', 'MD')
('canoe', 'NN')
('captive', 'JJ')
('care', 'NN')
('careening', 'VBG')
('carriage', 'NN')
('catch', 'VB')
('chase', 'VBP')
('chi', 'NN')
('children', 'NNS')

('chirruping', 'VBG')
('claims', 'NNS')
('claims', 'VBZ')
('clay', 'NN')
('cloudless', 'VB')
('cold', 'JJ')
('colt', 'VB')
('come', 'VB')
('comes', 'VBZ')
('congealing', 'VBG')
('conked', 'VBN')
('contract', 'NN')
('convulsing', 'VBG')
('coping', 'NN')
('corpse', 'NN')
('could', 'MD')
('count', 'VB')
('cover', 'NN')
('cracks', 'NNS')
('crashes', 'NNS')
('cries', 'NNS')
('crying', 'VBG')
('d', 'JJ')
('d', 'NN')
('dad', 'NN')
('dammit', 'NN')
('damp', 'JJ')
('dance', 'NN')
('dance', 'VB')
('dare', 'NN')
('darned', 'VBN')
('daughters', 'NNS')
('day', 'NN')
('dead', 'JJ')
('delight', 'JJ')
('despair', 'NN')
('did', 'VBD')
('different', 'JJ')
('dirt', 'NN')
('discarded', 'VBN')
('disdain', 'NN')
('diving', 'VBG')
('do', 'VB')
('do', 'VBP')
('don', 'VBP')
('door', 'NN')
('down', 'IN')
('drags', 'VBZ')

('drink', 'VBP')
('driver', 'NN')
('dunno', 'VBP')
('dusk', 'NN')
('each', 'DT')
('easier', 'JJR')
('else', 'RB')
('empty', 'JJ')
('enough', 'NN')
('enough', 'RB')
('ever', 'RB')
('everything', 'NN')
('exercise', 'NN')
('eyes', 'NNS')
('face', 'NN')
('face', 'VB')
('fame', 'NN')
('family', 'NN')
('far', 'RB')
('fast', 'JJ')
('fatal', 'JJ')
('father', 'NN')
('faux', 'NN')
('feel', 'VB')
('feet', 'NNS')
('fields', 'NNS')
('find', 'VB')
('find', 'VBP')
('fire', 'NN')
('firs', 'NNS')
('fitting', 'VBG')
('fixed', 'VBN')
('flash', 'JJ')
('flashbacks', 'NNS')
('fling', 'VBG')
('floor', 'NN')
('floorboards', 'NNS')
('fly', 'NN')
('footprints', 'NNS')
('for', 'IN')
('forecast', 'NN')
('forest', 'JJS')
('forever', 'RB')
('found', 'VBD')
('frame', 'NN')
('freely', 'RB')
('freer', 'NN')
('friend', 'NN')

('from', 'IN')
('front', 'JJ')
('fuck', 'VBD')
('gaze', 'NN')
('ghost', 'NN')
('gibbous', 'JJ')
('giggling', 'NN')
('give', 'VB')
('glass', 'NN')
('go', 'VB')
('go', 'VBP')
('goddamn', 'VB')
('gold', 'NN')
('got', 'VBD')
('grasp', 'VBP')
('graves', 'NNS')
('green', 'NN')
('greener', 'NN')
('grey', 'NN')
('grief', 'NN')
('grime', 'NN')
('gross', 'JJ')
('ground', 'NN')
('guess', 'NN')
('guide', 'VB')
('had', 'VBD')
('hail', 'VBP')
('hair', 'NN')
('hand', 'NN')
('hang', 'VB')
('hard', 'JJ')
('hate', 'VB')
('haunted', 'VBN')
('have', 'VB')
('having', 'VBG')
('hay', 'NN')
('he', 'PRP')
('heart', 'NN')
('heated', 'VBN')
('heaven', 'JJ')
('her', 'PRP\$')
('here', 'RB')
('herself', 'PRP')
('high', 'JJ')
('his', 'PRP\$')
('hold', 'RB')
('home', 'NN')
('homes', 'NNS')

('honesty', 'NN')
('hope', 'NN')
('hot-sand', 'NN')
('house', 'NN')
('how', 'WRB')
('howl', 'VBP')
('hummingbirds', 'NNS')
('hushed', 'JJ')
('i', 'JJ')
('i', 'NN')
('i', 'VB')
('i', 'VBP')
('i', 'VBZ')
('if', 'IN')
('in', 'IN')
('inky', 'JJ')
('insane', 'NN')
('into', 'IN')
('is', 'VBZ')
('it', 'PRP')
('its', 'PRP\$')
('itself', 'PRP')
('jazz', 'NN')
('jersey', 'NN')
('juices', 'NNS')
('just', 'RB')
('keep', 'VB')
('kin', 'NN')
('kinda', 'NN')
('know', 'VB')
('know', 'VBP')
('lambaste', 'NN')
('laughed', 'VBD')
('laughter', 'NN')
('lay', 'VBP')
('lead', 'VB')
('leave', 'VB')
('leaves', 'NNS')
('lesser', 'JJR')
('let', 'VB')
('life', 'NN')
('lift', 'NN')
('lighter', 'JJR')
('lights', 'NNS')
('like', 'IN')
('limps', 'VBZ')
('little', 'JJ')
('live', 'VB')

('ll', 'JJ')
('ll', 'NN')
('lone', 'NN')
('looking', 'VBG')
('lose', 'VB')
('lost', 'NN')
('lots', 'NNS')
('love', 'VBZ')
('loved', 'VBN')
('low', 'IN')
('lumber', 'VBP')
('m', 'NN')
('m', 'PRP')
('mailing', 'VBG')
('maimed', 'VBD')
('marred', 'JJ')
('masses', 'VBZ')
('match', 'VB')
('matchstick', 'JJ')
('may', 'MD')
('maybe', 'RB')
('maze', 'NN')
('me', 'PRP')
('mechanism', 'NN')
('medical', 'JJ')
('meet', 'VBP')
('melody', 'NN')
('melting', 'VBG')
('merely', 'RB')
('michelangelo', 'VB')
('midnight', 'NN')
('mind', 'NN')
('minds', 'NNS')
('mirror', 'NN')
('mirrors', 'NNS')
('missed', 'VBN')
('mistake', 'NN')
('mold', 'NN')
('moment', 'NN')
('momentary', 'JJ')
('moon', 'NN')
('more', 'JJR')
('morning', 'NN')
('mountains', 'NNS')
('mourning', 'VBG')
('much', 'JJ')
('much', 'RB')
('mud', 'NN')

('must', 'MD')
('my', 'PRP\$')
('myself', 'PRP')
('name', 'NN')
('name', 'VB')
('names', 'NNS')
('navigate', 'JJ')
('navigate', 'VB')
('near', 'IN')
('neck', 'NN')
('never', 'RB')
('new', 'JJ')
('news', 'NN')
('newscasts', 'NNS')
('night', 'NN')
('no', 'DT')
('none', 'NN')
('nostrils', 'NNS')
('not', 'RB')
('note', 'NN')
('now', 'RB')
('nunathat', 'RB')
('of', 'IN')
('off', 'IN')
('off', 'RP')
('offers', 'VBZ')
('offshore', 'RB')
('often', 'RB')
('old', 'JJ')
('on', 'IN')
('one', 'CD')
('onwards', 'NNS')
('open', 'JJ')
('or', 'CC')
('other', 'JJ')
('others', 'NNS')
('our', 'PRP\$')
('ours', 'VBZ')
('out', 'IN')
('out', 'RP')
('outlast', 'JJ')
('outside', 'IN')
('over', 'IN')
('over', 'RP')
('pad', 'NN')
('paddles', 'NNS')
('pain', 'NN')
('panoramic', 'JJ')

('parasite', 'JJ')
('passing', 'VBG')
('past', 'IN')
('past', 'JJ')
('past', 'NN')
('path', 'NN')
('pause', 'VBP')
('paved', 'VBD')
('peach', 'NN')
('peculiar', 'JJ')
('peerless', 'JJ')
('perched', 'VBN')
('pictures', 'NNS')
('pine', 'NN')
('pistol', 'NN')
('place', 'NN')
('play', 'VB')
('play', 'VBP')
('plum', 'NN')
('plume', 'VBP')
('pocket', 'JJ')
('poets', 'NNS')
('posing', 'VBG')
('pray', 'NN')
('prefer', 'VBP')
('pregamed', 'JJ')
('prithiee', 'NN')
('procession', 'NN')
('proud', 'JJ')
('proud', 'NN')
('pull', 'VBP')
('punctuated', 'VBN')
('pursuit', 'NN')
('rage', 'NN')
('rain', 'VBP')
('ransacked', 'VBD')
('re', 'JJ')
('re', 'VB')
('reach', 'VB')
('reacts', 'NNS')
('real', 'JJ')
('really', 'RB')
('reasonable', 'JJ')
('receipt', 'NN')
('reflecting', 'VBG')
('refracts', 'NNS')
('regrets', 'NNS')
('rehash', 'VB')

('remain', 'VBP')
('remember', 'VBP')
('resent', 'VB')
('resist', 'NN')
('resolute', 'NN')
('respond', 'VBP')
('rests', 'VBZ')
('retract', 'NN')
('riches', 'NNS')
('ringing', 'VBG')
('rivulet', 'NN')
('road', 'NN')
('roads', 'NNS')
('roof', 'NN')
('rosy', 'JJ')
('rot', 'NN')
('roulette', 'NN')
('s', 'JJ')
('s', 'NN')
('s', 'PRP')
('s', 'RB')
('s', 'VBD')
('s', 'VBZ')
('saints', 'NNS')
('same', 'JJ')
('saw', 'VBD')
('scared', 'VBN')
('scraps', 'NNS')
('selves', 'NNS')
('seventeen', 'JJ')
('shaking', 'VBG')
('shards', 'NNS')
('she', 'PRP')
('shirt', 'NN')
('shook', 'VB')
('shot', 'NN')
('should', 'MD')
('shovel', 'NN')
('shoving', 'VBG')
('shrink', 'NN')
('shrug', 'JJ')
('shunned', 'VBD')
('shut', 'VBD')
('sickle', 'NN')
('sidewalk', 'NN')
('sign', 'NN')
('sing', 'VBP')
('sink', 'VBP')

('sinks', 'NNS')
('sirens', 'VBZ')
('sit', 'NN')
('six', 'CD')
('skimming', 'VBG')
('sky', 'NN')
('slamming', 'VBG')
('slang', 'VBD')
('slowly', 'RB')
('smacks', 'NNS')
('smashed', 'JJ')
('smelling', 'VBG')
('smiles', 'VBZ')
('smiling', 'VBG')
('so', 'RB')
('sock', 'NN')
('somber', 'JJ')
('some', 'DT')
('sometimes', 'RB')
('son', 'NN')
('sorry', 'NN')
('souls', 'NNS')
('speaks', 'VBZ')
('speech', 'NN')
('spiderwebs', 'VBP')
('spike-like', 'JJ')
('squirrels', 'NNS')
('stand', 'NN')
('stand', 'VB')
('stands', 'VBZ')
('stapled', 'VBD')
('sticky', 'NN')
('still', 'RB')
('stone-facedness', 'JJ')
('straight', 'VBN')
('straighter', 'JJ')
('strained', 'VBN')
('streetlamps', 'NNS')
('streets', 'NNS')
('struck', 'VBP')
('stuck', 'VBP')
('sun', 'NN')
('supplies', 'NNS')
('suppose', 'VBP')
('supposed', 'VBN')
('sweet', 'JJ')
('sweet', 'VB')
('t', 'JJ')

('t', 'NN')
('t', 'RB')
('t', 'VB')
('t', 'VBZ')
('ta', 'JJ')
('table', 'NN')
('tact', 'NN')
('take', 'VB')
('takes', 'VBZ')
('talk', 'NN')
('talking', 'VBG')
('tall', 'JJ')
('taste', 'VB')
('taxing', 'NN')
('teach', 'VB')
('tears', 'NNS')
('tell', 'VB')
('than', 'IN')
('that', 'DT')
('that', 'IN')
('that', 'WDT')
('the', 'DT')
('their', 'PRP\$')
('them', 'PRP')
('then', 'RB')
('there', 'EX')
('there', 'RB')
('these', 'DT')
('they', 'PRP')
('think', 'VB')
('think', 'VBP')
('this', 'DT')
('those', 'DT')
('thought', 'VBD')
('thoughts', 'NNS')
('three', 'CD')
('three-fold', 'JJ')
('through', 'IN')
('thus', 'RB')
('til', 'JJ')
('time', 'NN')
('times', 'NNS')
('tin-can', 'JJ')
('tips', 'NNS')
('to', 'TO')
('told', 'VBD')
('too', 'RB')
('took', 'VBD')

('tracing', 'VBG')
('trails', 'NNS')
('train', 'NN')
('trained', 'VBD')
('transit', 'NN')
('trashed', 'VBD')
('tree', 'NN')
('trees', 'NNS')
('tribute', 'NN')
('trudge', 'NN')
('truth', 'NN')
('tsunami', 'NN')
('turn', 'VBP')
('turtle', 'NN')
('two', 'CD')
('under', 'IN')
('unkempt', 'JJ')
('unstill', 'JJ')
('unweave', 'VBP')
('up', 'RP')
('upon', 'IN')
('urn', 'JJ')
('us', 'PRP')
('vain', 'NN')
('vanity', 'NN')
('ve', 'RB')
('view', 'NN')
('violent', 'JJ')
('vuitton', 'NN')
('wake', 'VBP')
('walk', 'NN')
('wander', 'VB')
('wander', 'VBP')
('want', 'VBP')
('was', 'VBD')
('watch', 'VBP')
('wave', 'VBP')
('way', 'NN')
('we', 'PRP')
('weeds', 'NNS')
('well', 'RB')
('were', 'VBD')
('wet', 'JJ')
('what', 'WP')
('when', 'WRB')
('where', 'WRB')
('which', 'WDT')
('while', 'IN')

```

('who', 'WP')
('why', 'WRB')
('wilder', 'NN')
('winds', 'NNS')
('winter', 'NN')
('wish', 'VBP')
('wispier', 'JJR')
('with', 'IN')
('wolf', 'NN')
('wolves', 'NNS')
('woman', 'NN')
('women', 'NNS')
('won', 'VBD')
('world', 'NN')
('wringing', 'VBG')
('y', 'JJ')
('ya', 'NN')
('year', 'NN')
('years', 'NNS')
('yellow', 'JJ')
('yesterday', 'NN')
('yet', 'RB')
('you', 'PRP')
('your', 'PRP$')
('yourself', 'PRP')
('zen', 'NN')
('à', 'JJ')
('‘', 'VBP')
('’', 'JJ')
('’', 'NN')
('’', 'NNP')
('’', 'VB')
('’', 'VBD')
('’', 'VBP')
('’', 'VBZ')

```

```

[16]: fdist = FreqDist(words)
      print("Most common words (with #):")
      for i in fdist.most_common(10): print(i)
      fdist.plot(10)
      plt.show()

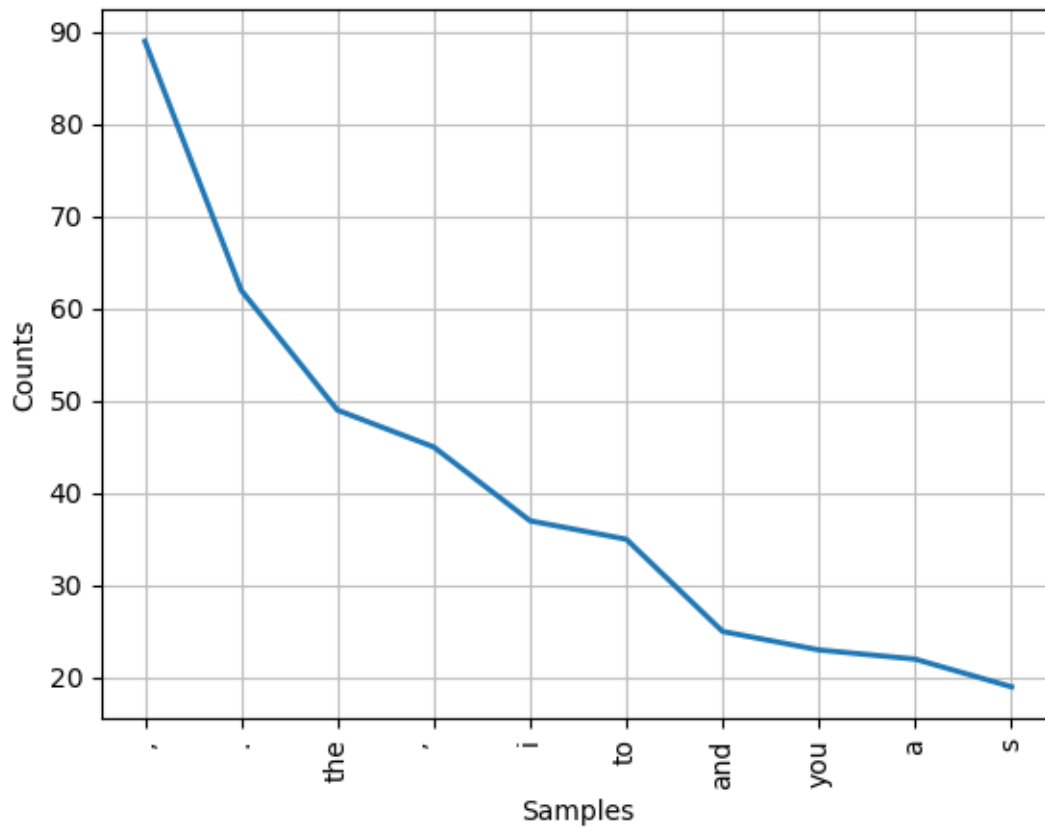
```

```

Most common words (with #):
(',', 89)
('.', 62)
('the', 49)
('’', 45)
('i', 37)

```

```
('to', 35)
('and', 25)
('you', 23)
('a', 22)
('s', 19)
```



```
[17]: words = word_tokenize(cleaned_text)
print(f"Total words in cleaned-up text: {len(words)}")
print("Parts of speech for each word (dups removed):")
for i in sorted(list(set(nltk.pos_tag(words))): print(i)
```

```
Total words in cleaned-up text: 642
Parts of speech for each word (dups removed):
('abyss', 'NN')
('aces', 'NNS')
('act', 'JJ')
('add', 'VBP')
('aim', 'NN')
('aims', 'JJ')
('air', 'NN')
('alarmed', 'VBD')
```


('alfred', 'VBD')
('alone', 'RB')
('alongside', 'IN')
('also', 'RB')
('although', 'IN')
('always', 'RB')
('amassed', 'VBN')
('amid', 'IN')
('anonymized', 'VBN')
('apartment', 'NN')
('arbored', 'VBD')
('ash', 'NN')
('ashtray', 'NN')
('ask', 'VB')
('asks', 'NNS')
('asphalt', 'NN')
('asphyxiation', 'NN')
('ate', 'NN')
('attack', 'NN')
('averts', 'NNS')
('avoid', 'JJ')
('awake', 'JJ')
('away', 'RB')
('awhile', 'IN')
('babel', 'NN')
('back', 'RB')
('back', 'RP')
('backburner', 'NN')
('behind', 'IN')
('behold', 'VBD')
('beholding', 'VBG')
('belong', 'IN')
('beneath', 'RB')
('betray', 'VBP')
('better', 'RB')
('beyond', 'IN')
('binge', 'NN')
('birds', 'NNS')
('bisect', 'NN')
('bit', 'NN')
('bit', 'RB')
('blame', 'NN')
('blamed', 'VBN')
('blank', 'JJ')
('blare', 'VBP')
('bleeding', 'VBG')
('blood', 'NN')
('blues', 'NNS')

('booze', 'NN')
('breath', 'NN')
('breathing', 'VBG')
('breeze', 'JJ')
('brien', 'NN')
('brim', 'VBP')
('bullseye', 'JJ')
('bullshitting', 'VBG')
('burn', 'VBP')
('burning', 'VBG')
('bury', 'VB')
('caked', 'VBD')
('came', 'VBD')
('canoe', 'NN')
('captive', 'JJ')
('care', 'NN')
('careening', 'VBG')
('carriage', 'VB')
('catch', 'JJ')
('chase', 'JJ')
('chi', 'NN')
('children', 'NNS')
('chirruping', 'VBG')
('christmas', 'NN')
('claims', 'NNS')
('clay', 'NN')
('cloudless', 'JJ')
('cold', 'JJ')
('colt', 'NN')
('come', 'JJ')
('come', 'VB')
('comes', 'VBZ')
('congealing', 'VBG')
('conked', 'VBN')
('conor', 'NN')
('contract', 'NN')
('convulsing', 'VBG')
('coping', 'VBG')
('corpse', 'VB')
('could', 'MD')
('count', 'VB')
('cover', 'NN')
('cracks', 'NNS')
('crashes', 'NNS')
('cries', 'NNS')
('crying', 'VBG')
('cypress', 'JJ')
('dad', 'NN')

('dammit', 'NN')
('damp', 'JJ')
('dance', 'NN')
('dare', 'NN')
('darned', 'VBD')
('daughters', 'NNS')
('day', 'NN')
('dead', 'JJ')
('deaths', 'NNS')
('delight', 'JJ')
('despair', 'NN')
('different', 'JJ')
('dirt', 'NN')
('discarded', 'VBD')
('disdain', 'NN')
('diving', 'VBG')
('door', 'NN')
('drags', 'VBZ')
('drink', 'VBP')
('driver', 'NN')
('driver', 'RB')
('dunno', 'NN')
('dusk', 'JJ')
('easier', 'JJR')
('ekphrastic', 'JJ')
('else', 'RB')
('empty', 'JJ')
('enough', 'JJ')
('enough', 'RB')
('epistolary', 'JJ')
('ever', 'RB')
('everything', 'NN')
('exercise', 'NN')
('eyes', 'NNS')
('face', 'NN')
('facedness', 'NN')
('fame', 'NN')
('family', 'NN')
('far', 'RB')
('fast', 'JJ')
('fatal', 'JJ')
('father', 'NN')
('father', 'RB')
('faux', 'JJ')
('feel', 'VB')
('feet', 'NNS')
('fields', 'NNS')
('find', 'IN')

('find', 'VBP')
('fire', 'VBP')
('firs', 'NNS')
('fitting', 'VBG')
('fixed', 'VBN')
('flash', 'JJ')
('flashbacks', 'NNS')
('fling', 'VBG')
('floor', 'NN')
('floorboards', 'NNS')
('fly', 'JJ')
('fold', 'JJ')
('footprints', 'NNS')
('forecast', 'JJ')
('forest', 'JJS')
('forever', 'RB')
('found', 'VBN')
('frame', 'NN')
('freely', 'RB')
('freer', 'JJ')
('friend', 'NN')
('friend', 'VBP')
('front', 'JJ')
('fuck', 'NN')
('gaze', 'VBP')
('gd', 'NN')
('ghost', 'NN')
('gibbous', 'JJ')
('giggling', 'VBG')
('give', 'JJ')
('glass', 'NN')
('go', 'VB')
('go', 'VBP')
('goddamn', 'JJ')
('gold', 'JJ')
('got', 'VBD')
('grasp', 'VBP')
('graves', 'NNS')
('green', 'JJ')
('greener', 'NN')
('grey', 'VBD')
('grief', 'NN')
('grime', 'NN')
('gross', 'JJ')
('ground', 'NN')
('guess', 'JJ')
('guide', 'NN')
('haiku', 'NN')

('hail', 'NN')
('hair', 'NN')
('hand', 'NN')
('hang', 'NN')
('hard', 'JJ')
('hate', 'NN')
('haunted', 'VBN')
('hay', 'NN')
('heart', 'NN')
('heated', 'VBN')
('heaven', 'NN')
('high', 'JJ')
('hold', 'VBP')
('home', 'NN')
('homes', 'NNS')
('honesty', 'JJ')
('hope', 'NN')
('hot', 'JJ')
('house', 'NN')
('howl', 'NN')
('hummingbirds', 'NNS')
('hushed', 'VBN')
('inky', 'JJ')
('insane', 'NN')
('jazz', 'NN')
('jersey', 'NN')
('juices', 'NNS')
('keep', 'VB')
('kin', 'JJ')
('kinda', 'NN')
('know', 'JJ')
('know', 'VBP')
('lambaste', 'NN')
('laughed', 'JJ')
('laughter', 'NN')
('lay', 'JJ')
('lay', 'NN')
('lead', 'VB')
('leave', 'JJR')
('leaves', 'VBZ')
('lesser', 'JJR')
('let', 'NN')
('let', 'VBP')
('life', 'NN')
('lift', 'JJ')
('lighter', 'JJ')
('lights', 'NNS')
('like', 'IN')

('limps', 'NNS')
('little', 'JJ')
('live', 'JJ')
('lone', 'NN')
('looking', 'VBG')
('lose', 'JJ')
('lost', 'VBD')
('lots', 'NNS')
('love', 'NN')
('loved', 'VBD')
('low', 'JJ')
('lumber', 'VBP')
('mailing', 'VBG')
('maimed', 'VBD')
('marred', 'VBD')
('masses', 'VBZ')
('match', 'NN')
('matchstick', 'VB')
('may', 'MD')
('maybe', 'RB')
('maze', 'NN')
('mechanism', 'NN')
('medical', 'JJ')
('meet', 'VB')
('melody', 'VBP')
('melting', 'VBG')
('merely', 'RB')
('michelangelo', 'JJ')
('midnight', 'NN')
('mind', 'NN')
('minds', 'NNS')
('mirror', 'NN')
('mirrors', 'NNS')
('missed', 'VBD')
('mistake', 'NN')
('mold', 'VBN')
('moment', 'NN')
('momentary', 'JJ')
('moon', 'NN')
('morning', 'NN')
('mountains', 'NNS')
('mourning', 'VBG')
('much', 'JJ')
('mud', 'JJ')
('must', 'MD')
('name', 'NN')
('name', 'VB')
('names', 'NNS')

('navigate', 'JJ')
('navigate', 'NN')
('near', 'IN')
('neck', 'NN')
('never', 'RB')
('new', 'JJ')
('news', 'NN')
('newscasts', 'NNS')
('night', 'NN')
('none', 'NN')
('nostrils', 'NNS')
('note', 'NN')
('nunathat', 'DT')
('o', 'JJ')
('offers', 'VBZ')
('offshore', 'RB')
('often', 'RB')
('old', 'JJ')
('one', 'CD')
('onwards', 'NNS')
('open', 'JJ')
('others', 'NNS')
('outlast', 'JJ')
('outside', 'JJ')
('pad', 'NN')
('paddles', 'NNS')
('pain', 'JJ')
('panoramic', 'JJ')
('parasite', 'NN')
('paris', 'JJ')
('passing', 'VBG')
('past', 'IN')
('past', 'JJ')
('path', 'NN')
('pause', 'NN')
('paved', 'VBD')
('peach', 'NN')
('peculiar', 'JJ')
('peerless', 'JJ')
('perched', 'VBN')
('pictures', 'NNS')
('pine', 'NN')
('pistol', 'NN')
('place', 'NN')
('play', 'NN')
('plum', 'NN')
('plume', 'JJ')
('pocket', 'JJ')

('poetry', 'NN')
('poets', 'NNS')
('portfolio', 'NN')
('posing', 'VBG')
('pray', 'VBP')
('prefer', 'VBP')
('pregamed', 'VBD')
('prithree', 'JJ')
('procession', 'NN')
('proud', 'JJ')
('proud', 'NN')
('pull', 'VB')
('punctuated', 'VBD')
('pursuit', 'NN')
('rage', 'NN')
('rain', 'VBP')
('ransacked', 'VBD')
('rap', 'VBP')
('reach', 'VB')
('reacts', 'NNS')
('real', 'JJ')
('really', 'RB')
('reasonable', 'JJ')
('receipt', 'NN')
('reflecting', 'VBG')
('refracts', 'NNS')
('regrets', 'NNS')
('rehash', 'NN')
('remain', 'VBP')
('remember', 'VBP')
('resent', 'NN')
('resist', 'NN')
('resolute', 'VB')
('respond', 'VBP')
('rests', 'NNS')
('retract', 'JJ')
('riches', 'NNS')
('ringing', 'VBG')
('rivulet', 'NN')
('road', 'NN')
('roads', 'NNS')
('roof', 'NN')
('rosy', 'JJ')
('rot', 'NN')
('roulette', 'NN')
('saints', 'NNS')
('sand', 'NN')
('saw', 'NN')

('scared', 'VBD')
('scraps', 'NNS')
('selves', 'NNS')
('seventeen', 'JJ')
('shaking', 'VBG')
('shards', 'NNS')
('shirt', 'NN')
('shook', 'NN')
('shot', 'NN')
('shovel', 'VBP')
('shoving', 'VBG')
('shrink', 'VBP')
('shrug', 'JJ')
('shunned', 'VBN')
('shut', 'VBD')
('sickle', 'NN')
('sidewalk', 'NN')
('sign', 'NN')
('sing', 'VBG')
('sink', 'NN')
('sinks', 'NNS')
('sirens', 'NNS')
('sit', 'NN')
('six', 'CD')
('skimming', 'VBG')
('sky', 'JJ')
('sky', 'NN')
('slamming', 'VBG')
('slang', 'JJ')
('slang', 'NN')
('slowly', 'RB')
('smacks', 'NNS')
('smashed', 'VBD')
('smelling', 'VBG')
('smiles', 'NNS')
('smiles', 'VBZ')
('smiling', 'VBG')
('sock', 'NN')
('somber', 'NNP')
('sometimes', 'RB')
('son', 'NN')
('sorry', 'NN')
('souls', 'NNS')
('speaks', 'JJ')
('speech', 'NN')
('spiderwebs', 'NN')
('spike', 'NN')
('squirrels', 'NNS')

('stand', 'NN')
('stand', 'VB')
('stands', 'VBZ')
('stapled', 'VBD')
('sticky', 'JJ')
('still', 'RB')
('stone', 'NN')
('straight', 'JJ')
('straighter', 'JJ')
('strained', 'VBD')
('streetlamps', 'NNS')
('streets', 'NNS')
('struck', 'JJ')
('stuck', 'VBN')
('sun', 'JJ')
('sun', 'NN')
('supplies', 'NNS')
('suppose', 'JJ')
('supposed', 'VBD')
('surrealism', 'VBP')
('sweet', 'JJ')
('sweet', 'NN')
('ta', 'JJ')
('table', 'NN')
('tact', 'JJ')
('take', 'VB')
('takes', 'VBZ')
('talk', 'NN')
('talking', 'VBG')
('tall', 'JJ')
('taste', 'VBP')
('taxing', 'VBG')
('teach', 'VB')
('tears', 'NNS')
('tell', 'VB')
('think', 'NN')
('thought', 'VBD')
('thoughts', 'NNS')
('three', 'CD')
('thus', 'RB')
('til', 'JJ')
('time', 'NN')
('times', 'NNS')
('tin', 'NN')
('tips', 'NNS')
('told', 'VBD')
('took', 'VBD')
('tracing', 'VBG')

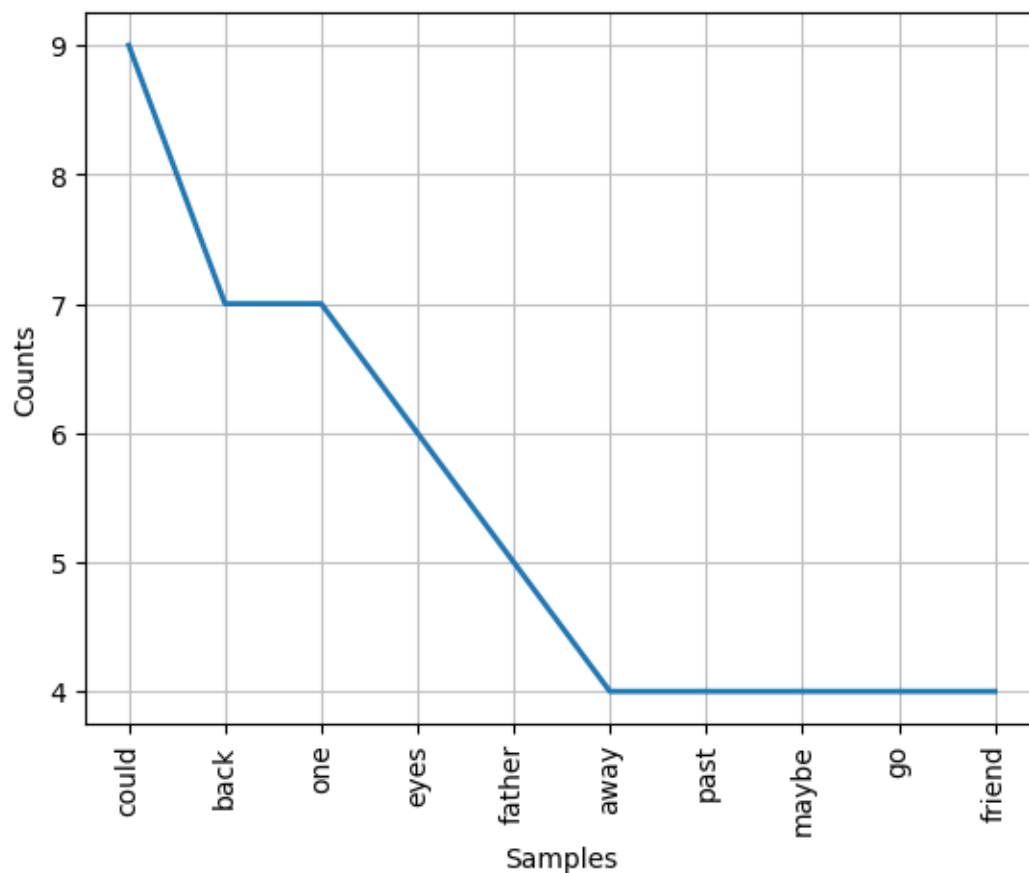
('trails', 'NNS')
('train', 'VBP')
('trained', 'VBD')
('transit', 'JJ')
('trashed', 'VBD')
('tree', 'JJ')
('trees', 'NNS')
('tribute', 'NN')
('trudge', 'NN')
('truth', 'NN')
('tsunami', 'NN')
('turn', 'NN')
('turn', 'VBP')
('turtle', 'NN')
('tvs', 'JJ')
('two', 'CD')
('unkempt', 'JJ')
('unstill', 'VB')
('unweave', 'JJ')
('upon', 'IN')
('urn', 'JJ')
('us', 'PRP')
('vain', 'JJ')
('vain', 'NN')
('vanity', 'NN')
('verse', 'NN')
('view', 'NN')
('violent', 'JJ')
('vuitton', 'NN')
('wake', 'VBP')
('walk', 'NN')
('wander', 'NN')
('wander', 'VB')
('wander', 'VBP')
('want', 'NN')
('want', 'VBP')
('watch', 'NN')
('wave', 'VBP')
('way', 'NN')
('weeds', 'NNS')
('well', 'RB')
('wet', 'JJ')
('wilder', 'NN')
('winds', 'NNS')
('winter', 'NN')
('wish', 'JJ')
('wispier', 'JJR')
('wolf', 'NN')

```
('wolves', 'NNS')
('woman', 'NN')
('women', 'NNS')
('world', 'NN')
('wringing', 'VBG')
('ya', 'VBP')
('year', 'NN')
('years', 'NNS')
('yellow', 'JJ')
('yesterday', 'NN')
('yet', 'RB')
('zen', 'NNP')
```

```
[18]: fdist = FreqDist(words)
      print("Most common words (with #):")
      for i in fdist.most_common(10): print(i)
      fdist.plot(10)
      plt.show()
```

Most common words (with #):

```
('could', 9)
('back', 7)
('one', 7)
('eyes', 6)
('father', 5)
('away', 4)
('past', 4)
('maybe', 4)
('go', 4)
('friend', 4)
```



```
[19]: print("""
Overall, this was pretty interesting data to look at. I was curious to see
↳which words were most common here and what that says about the poetry, but
↳also curious from a technical standpoint what happens with really long
↳'sentences' like in Conor's poetry.

It's especially interesting the things that spacy highlights - like "grey
↳forecast" as a person, for some reason, which is obviously not right. I
↳guess it doesn't really like very figurative/flowery language.

The WordClouds were also pretty fun to look at - I didn't realize "back" was
↳such a common word in general let alone in this poetry.

One downside I did notice - TF-IDF seems very awkward for very varying sentence
↳lengths. It really liked the longest sentences, presumably since they share
↳a lot of similarities etc. with the other ones, but that's something to
↳consider when using TF-IDF with future datasets.
""")
```

Overall, this was pretty interesting data to look at. I was curious to see which words were most common here and what that says about the poetry, but also curious from a technical standpoint what happens with really long 'sentences' like in Conor's poetry.

It's especially interesting the things that spacy highlights - like "grey forecast" as a person, for some reason, which is obviously not right. I guess it doesn't really like very figurative/flowery language.

The WordClouds were also pretty fun to look at - I didn't realize "back" was such a common word in general let alone in this poetry.

One downside I did notice - TF-IDF seems very awkward for very varying sentence lengths. It really liked the longest sentences, presumably since they share a lot of similarities etc. with the other ones, but that's something to consider when using TF-IDF with future datasets.

[]: