

PRAC 7

```
sentence1 = "I will walk 500 miles and I would walk 500 more. Just to be the man who walks " + \
"a thousand miles to fall down at your door!"
sentence2 = "I played the play playfully as the players were playing in the play with playfulness"
```

```
from nltk import word_tokenize, sent_tokenize
print("Tokenized words:", word_tokenize(sentence1))
print("\nTokenized sentences:", sent_tokenize(sentence1))
```

```
from nltk import pos_tag
token = word_tokenize(sentence1) + word_tokenize(sentence2)
tagged = pos_tag(token)
print("Tagging Parts of Speech:", tagged)
```

```
from nltk.corpus import stopwords
stop_words = stopwords.words('english')
token = word_tokenize(sentence1)
cleaned_token = []
for word in token:
    if word not in stop_words:
        cleaned_token.append(word)
print('Unclean version:', token)
print("\nCleaned version:", cleaned_token)
```

```
from nltk.stem import PorterStemmer
stemmer = PorterStemmer()
token = word_tokenize(sentence2)
stemmed = [stemmer.stem(word) for word in token]
print(" ".join(stemmed))
```

```
from nltk.stem import WordNetLemmatizer
lemmatizer = WordNetLemmatizer()
token = word_tokenize(sentence2)
lemmatized_output = [lemmatizer.lemmatize(word) for word in token]
print(" ".join(lemmatized_output))
```

