In [1]:

```python
#import the necassary library
import nltk
nltk.download
```

Out[1]:

```
<bound method Downloader.download of <nltk.downloader.Downloader object at 0x1a
1efb1cf8>>
```

In [2]:

```python
#import the words
from nltk.corpus import brown
brown.categories()
```

Out[2]:

```
['adventure',
 'belles_lettres',
 'editorial',
 'fiction',
 'government',
 'hobbies',
 'humor',
 'learned',
 'lore',
 'mystery',
 'news',
 'religion',
 'reviews',
 'romance',
 'science_fiction']
```

In [3]:

```python
brown.words(categories='adventure')[:50]
```

Out[3]:

```
['Dan',
 'Morgan',
 'told',
 'himself',
 'he',
 'would',
 'forget',
 'Ann',
 'Turner',
 '.',
 'He',
 'was',
 'well',
 'rid',
 'of',
 'her',
 '.',
 'He',
 'certainly',
 "didn't",
 'want',
 'a',
 'wife',
 'who',
 'was',
 'fickle',
 'as',
 'Ann',
 '.',
 'If',
 'he',
 'had',
 'married',
 'her',
 ',',
 "he'd",
 'have',
 'been',
 'asking',
 'for',
 'trouble',
 '.',
 'But',
 'all',
 'of',
 'this',
 'was',
 'rationalization',
 '.',
 'Sometimes']
```

In [4]:

```python
#  IMPORT INAUGRAL CORPUS AND ACCESS DATA
# This includes every President's inaugral address

from nltk.corpus import inaugural
inaugural.fileids()
```

Out[4]:

```
['1789-Washington.txt',
 '1793-Washington.txt',
 '1797-Adams.txt',
 '1801-Jefferson.txt',
 '1805-Jefferson.txt',
 '1809-Madison.txt',
 '1813-Madison.txt',
 '1817-Monroe.txt',
 '1821-Monroe.txt',
 '1825-Adams.txt',
 '1829-Jackson.txt',
 '1833-Jackson.txt',
 '1837-VanBuren.txt',
 '1841-Harrison.txt',
 '1845-Polk.txt',
 '1849-Taylor.txt',
 '1853-Pierce.txt',
 '1857-Buchanan.txt',
 '1861-Lincoln.txt',
 '1865-Lincoln.txt',
 '1869-Grant.txt',
 '1873-Grant.txt',
 '1877-Hayes.txt',
 '1881-Garfield.txt',
 '1885-Cleveland.txt',
 '1889-Harrison.txt',
 '1893-Cleveland.txt',
 '1897-McKinley.txt',
 '1901-McKinley.txt',
 '1905-Roosevelt.txt',
 '1909-Taft.txt',
 '1913-Wilson.txt',
 '1917-Wilson.txt',
 '1921-Harding.txt',
 '1925-Coolidge.txt',
 '1929-Hoover.txt',
 '1933-Roosevelt.txt',
 '1937-Roosevelt.txt',
 '1941-Roosevelt.txt',
 '1945-Roosevelt.txt',
 '1949-Truman.txt',
 '1953-Eisenhower.txt',
 '1957-Eisenhower.txt',
 '1961-Kennedy.txt',
 '1965-Johnson.txt',
 '1969-Nixon.txt',
 '1973-Nixon.txt',
 '1977-Carter.txt',
 '1981-Reagan.txt',
 '1985-Reagan.txt',
 '1989-Bush.txt',
 '1993-Clinton.txt',
 '1997-Clinton.txt',
 '2001-Bush.txt',
 '2005-Bush.txt',
 '2009-Obama.txt',
 '2013-Obama.txt',
 '2017-Trump.txt']
```

In [5]:

```
inaugural.words(fileids='2009-Obama.txt')[:50]
```

Out[5]:

```
['My',
 'fellow',
 'citizens',
 ':',
 'I',
 'stand',
 'here',
 'today',
 'humbled',
 'by',
 'the',
 'task',
 'before',
 'us',
 ',',
 'grateful',
 'for',
 'the',
 'trust',
 'you',
 'have',
 'bestowed',
 ',',
 'mindful',
 'of',
 'the',
 'sacrifices',
 'borne',
 'by',
 'our',
 'ancestors',
 '.',
 'I',
 'thank',
 'President',
 'Bush',
 'for',
 'his',
 'service',
 'to',
 'our',
 'nation',
 ',',
 'as',
 'well',
 'as',
 'the',
 'generosity',
 'and',
 'cooperation']
```

In [6]:

```
inaugural.words(fileids='1861-Lincoln.txt')[:50]
```

Out[6]:

```
['Fellow',
 '-',
 'Citizens',
 'of',
 'the',
 'United',
 'States',
 ':',
 'In',
 'compliance',
 'with',
 'a',
 'custom',
 'as',
 'old',
 'as',
 'the',
 'Government',
 'itself',
 ',',
 'I',
 'appear',
 'before',
 'you',
 'to',
 'address',
 'you',
 'briefly',
 'and',
 'to',
 'take',
 'in',
 'your',
 'presence',
 'the',
 'oath',
 'prescribed',
 'by',
 'the',
 'Constitution',
 'of',
 'the',
 'United',
 'States',
 'to',
 'be',
 'taken',
 'by',
 'the',
 'President']
```

In [7]:

```python
# Frequency Distribution of words in a text
fd = nltk.FreqDist(inaugural.words(fileids='1861-Lincoln.txt'))
fd
```

Out[7]:

```
FreqDist({'the': 240, ',': 195, 'of': 146, 'to': 132, '.': 110, 'and': 101, 'be
': 76, 'in': 72, 'that': 57, 'a': 53, ...})
```

In [8]:

```python
from nltk.probability import ConditionalFreqDist
cfd = ConditionalFreqDist((len(word), word) for word in inaugural.words(fileids='1861-Linco
ln.txt'))
cfd
```

Out[8]:

```
<ConditionalFreqDist with 16 conditions>
```

In [ ]: