

In [7]:

```
import nltk
#nltk.download('averaged_perceptron_tagger')
texts = ["""The horse (Equus ferus caballus) is one of two extant subspecies of Equus ferus. It is an odd-toed ungulate mammal belonging to the taxonomic family Equidae. The horse has evolved over the past 45 to 55 million years from a small multi-toed creature, Eohippus, into the large, single-toed animal of today. Humans began domesticating horses around 4000 BC, and their domestication is believed to have been widespread by 3000 BC. Horses in the subspecies caballus are domesticated, although some domesticated populations live in the wild as feral horses. These feral populations are not true wild horses, as this term is used to describe horses that have never been domesticated, such as the endangered Przewalski's horse, a separate subspecies, and the only remaining true wild horse. There is an extensive, specialized vocabulary used to describe equine-related concepts, covering everything from anatomy to life stages, size, colors, markings, breeds, locomotion, and behavior.""]
```

Horses' anatomy enables them to make use of speed to escape predators and they have a well-developed sense of balance and a strong fight-or-flight response. Related to this need to flee from predators in the wild is an unusual trait: horses are able to sleep both standing up and lying down, with younger horses tending to sleep significantly more than adults. Female horses, called mares, carry their young for approximately 11 months, and a young horse, called a foal, can stand and run shortly following birth. Most domesticated horses begin training under saddle or in harness between the ages of two and four. They reach full adult development by age five, and have an average lifespan of between 25 and 30 years.""]

In [8]:

```
for text in texts:
    sentences = nltk.sent_tokenize(text)
    for sentence in sentences:
        words = nltk.word_tokenize(sentence)
        tagged_words = nltk.pos_tag(words)
        print(tagged_words)
```

[('The', 'DT'), ('horse', 'NN'), ('(', '('), ('Equus', 'NNP'), ('ferus', 'RB'), ('caballus', 'NN'), (')', ')'), ('is', 'VBZ'), ('one', 'CD'), ('of', 'IN'), ('two', 'CD'), ('extant', 'JJ'), ('subspecies', 'NNS'), ('of', 'IN'), ('Equus', 'NNP'), ('ferus', 'NN'), ('.', '.')]
[('It', 'PRP'), ('is', 'VBZ'), ('an', 'DT'), ('odd-toed', 'JJ'), ('ungulate', 'NN'), ('mammal', 'NN'), ('belonging', 'VBG'), ('to', 'TO'), ('the', 'DT'), ('taxonomic', 'JJ'), ('family', 'NN'), ('Equidae', 'NNP'), ('.', '.')]
[('The', 'DT'), ('horse', 'NN'), ('has', 'VBZ'), ('evolved', 'VBN'), ('over', 'IN'), ('the', 'DT'), ('past', 'JJ'), ('45', 'CD'), ('to', 'TO'), ('55', 'CD'), ('million', 'CD'), ('years', 'NNS'), ('from', 'IN'), ('a', 'DT'), ('small', 'JJ'), ('multi-toed', 'JJ'), ('creature', 'NN'), (',', ','), ('Eohippus', 'NNP'), (',', ','), ('into', 'IN'), ('the', 'DT'), ('large', 'JJ'), (',', ','), ('single-toed', 'JJ'), ('animal', 'NN'), ('of', 'IN'), ('today', 'NN'), ('.', '.')]
[('Humans', 'NNS'), ('began', 'VBD'), ('domesticating', 'VBG'), ('horses', 'NNS'), ('around', 'IN'), ('4000', 'CD'), ('BC', 'NNP'), (',', ','), ('and', 'CC'), ('their', 'PRP\$'), ('domestication', 'NN'), ('is', 'VBZ'), ('believed', 'VBN'), ('to', 'TO'), ('have', 'VB'), ('been', 'VBN'), ('widespread', 'VBN'), ('by', 'IN'), ('3000', 'CD'), ('BC', 'NNP'), ('.', '.')]
[('Horses', 'NNS'), ('in', 'IN'), ('the', 'DT'), ('subspecies', 'NNS'), ('caballus', 'CC'), ('are', 'VBP'), ('domesticated', 'VBN'), (',', ','), ('although', 'IN'), ('some', 'DT'), ('domesticated', 'JJ'), ('populations', 'NNS'), ('live', 'VBP'), ('in', 'IN'), ('the', 'DT'), ('wild', 'JJ'), ('as', 'IN'), ('feral', 'JJ'), ('horses', 'NNS'), ('.', '.')]
[('These', 'DT'), ('feral', 'JJ'), ('populations', 'NNS'), ('are', 'VBP'), ('not', 'RB'), ('true', 'JJ'), ('wild', 'JJ'), ('horses', 'NNS'), (',', ','), ('as', 'IN'), ('this', 'DT'), ('term', 'NN'), ('is', 'VBZ'), ('used', 'VBN'), ('to', 'TO'), ('describe', 'VB'), ('horses', 'NNS'), ('that', 'WDT'), ('have', 'VBP'), ('never', 'RB'), ('been', 'VBN'), ('domesticated', 'VBN'), (',', ','), ('such', 'JJ'), ('as', 'IN'), ('the', 'DT'), ('endangered', 'VBN'), ('Przewalski', 'NNP'), ('"', 'POS'), ('horse', 'NN'), (',', ','), ('a', 'DT'), ('separate', 'JJ'), ('subspecies', 'NNS'), (',', ','), ('and', 'CC'), ('the', 'DT'), ('only', 'JJ'), ('remaining', 'VBG'), ('true', 'JJ'), ('wild', 'JJ'), ('horse', 'NN'), ('.', '.')]
[('There', 'EX'), ('is', 'VBZ'), ('an', 'DT'), ('extensive', 'JJ'), (',', ','), ('specialized', 'JJ'), ('vocabulary', 'NN'), ('used', 'VBN'), ('to', 'TO'), ('describe', 'VB'), ('equine-related', 'JJ'), ('concepts', 'NNS'), (',', ','), ('covering', 'VBG'), ('everything', 'NN'), ('from', 'IN'), ('anatomy', 'NN'), ('to', 'TO'), ('life', 'NN'), ('stages', 'NNS'), (',', ','), ('size', 'NN'), (',', ','), ('colors', 'NNS'), (',', ','), ('markings', 'NNS'), (',', ','), ('breeds', 'NNS'), (',', ','), ('locomotion', 'NN'), (',', ','), ('and', 'CC'), ('behavior', 'NN'), ('.', '.')]
[('Horses', 'NNS'), ('"', 'POS'), ('anatomy', 'NN'), ('enables', 'VBZ'), ('them', 'PRP'), ('to', 'TO'), ('make', 'VB'), ('use', 'NN'), ('of', 'IN'), ('speed', 'NN'), ('to', 'TO'), ('escape', 'VB'), ('predators', 'NNS'), ('and', 'CC'), ('they', 'PRP'), ('have', 'VBP'), ('a', 'DT'), ('well-developed', 'JJ'), ('sense', 'NN'), ('of', 'IN'), ('balance', 'NN'), ('and', 'CC'), ('a', 'DT'), ('strong', 'JJ'), ('fight-or-flight', 'JJ'), ('response', 'NN'), ('.', '.')]
[('Related', 'VBN'), ('to', 'TO'), ('this', 'DT'), ('need', 'NN'), ('to', 'TO'), ('flee', 'VB'), ('from', 'IN'), ('predators', 'NNS'), ('in', 'IN'), ('the', 'DT'), ('wild', 'NN'), ('is', 'VBZ'), ('an', 'DT'), ('unusual', 'JJ'), ('trait', 'NN'), (':', ':'), ('horses', 'NNS'), ('are', 'VBP'), ('able', 'JJ'), ('to', 'TO'), ('sleep', 'VB'), ('both', 'DT'), ('standing', 'VBG'), ('up', 'RP'), ('and', 'CC'), ('lying', 'VBG'), ('down', 'RP'), (',', ','), ('with', 'IN'), ('younger', 'JJR'), ('horses', 'NNS'), ('tending', 'VBG'), ('to', 'TO'), ('sleep', 'VB'), ('significantly', 'RB'), ('more', 'JJR'), ('than', 'IN'), ('adults', 'NNS'), ('.', '.')]
[('Female', 'JJ'), ('horses', 'NNS'), (',', ','), ('called', 'VBD'), ('mares', 'NNS'), (',', ','), ('carry', 'VBP'), ('their', 'PRP\$'), ('young', 'J

```
J'), ('for', 'IN'), ('approximately', 'RB'), ('11', 'CD'), ('months', 'NN
S'), ('', ' '), ('and', 'CC'), ('a', 'DT'), ('young', 'JJ'), ('horse', 'N
N'), ('', ' '), ('called', 'VBD'), ('a', 'DT'), ('foal', 'NN'), ('', ' '),
','), ('can', 'MD'), ('stand', 'VB'), ('and', 'CC'), ('run', 'VB'), ('shor
tly', 'RB'), ('following', 'VBG'), ('birth', 'NN'), ('.', '.')]
[('Most', 'JJ'), ('domesticated', 'JJ'), ('horses', 'NNS'), ('begin', 'VB
P'), ('training', 'VBG'), ('under', 'IN'), ('saddle', 'NN'), ('or', 'CC'),
('in', 'IN'), ('harness', 'NN'), ('between', 'IN'), ('the', 'DT'), ('age
s', 'NNS'), ('of', 'IN'), ('two', 'CD'), ('and', 'CC'), ('four', 'CD'),
('.', '.')]
[('They', 'PRP'), ('reach', 'VBP'), ('full', 'JJ'), ('adult', 'NN'), ('dev
elopment', 'NN'), ('by', 'IN'), ('age', 'NN'), ('five', 'CD'), ('', ' '),
('and', 'CC'), ('have', 'VBP'), ('an', 'DT'), ('average', 'JJ'), ('lifespa
n', 'NN'), ('of', 'IN'), ('between', 'IN'), ('25', 'CD'), ('and', 'CC'),
('30', 'CD'), ('years', 'NNS'), ('.', '.')]

```