

```
import nltk
from nltk.tokenize import word_tokenize
from collections import Counter

# Download tokenizer if not already done
nltk.download('punkt_tab')

# Input text
text = "This is a test. This test is only a test."

# Tokenize the text into words
words = word_tokenize(text.lower()) # convert to lowercase to catch case duplicates

# Filter out non-alphabetic tokens (e.g., punctuation)
words = [word for word in words if word.isalpha()]

# Count word frequencies
word_counts = Counter(words)
print(word_counts)

# Find duplicate words
duplicates = {word: count for word, count in word_counts.items() if count > 1}

# Display result
print("Duplicate Words Found:")
for word, count in duplicates.items():
    print(f"{word}: {count} times")
```

```
→ [nltk_data] Downloading package punkt_tab to /root/nltk_data...
[nltk_data]  Unzipping tokenizers/punkt_tab.zip.
Counter({'test': 3, 'this': 2, 'is': 2, 'a': 2, 'only': 1})
Duplicate Words Found:
this: 2 times
is: 2 times
a: 2 times
test: 3 times
```