

```
Problems Javadoc Declaration Console  
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)  
i go to school by bus. the bus is big . the school is also big. i like big bus and big school!  
i go to school by bus the bus is big the school is also big i like big bus and big school  
list1:[stav. at. vour. home. or. my, others]|
```

```
Problems Javadoc Declaration Console  
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)  
i go to school by bus. the bus is big . the school is also big. i like big bus and big school!  
i go to school by bus the bus is big the school is also big i like big bus and big school  
list1:[stav. at. vour. home. or. my, others]|
```

```
Problems Javadoc Declaration Console  
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)  
i go to school by bus. the bus is big . the school is also big. i like big bus and big school!  
i go to school by bus the bus is big the school is also big i like big bus and big school  
list1:[stav. at. vour. home. or. my, others]|
```

# PART1

## Problem: Simple Text Concordance

### Data Structure: List

Concordances are frequently used in linguistics, **when studying a text**.  
For example:

- comparing different usages of the same word.
- analysing keywords. analysing word frequencies.
- finding and analysing phrases and idioms.
- finding translations of subsentential elements, e.g. terminology, in bitexts and translation memories.
- creating indexes and word lists (also useful for publishing).

# Text Concordance

- This topic will focus on the function **of text analysis**.
- The text analysis **involve extraction process** to identify each words from a passage (input- keyboard/file).
- The **words** (cannot directly be displayed) need to be **hold** by a variable **for comparison purposes** (to prevent the same word displayed more than once).
- Since a passage **comprises of** several paragraphs and paragraph consists of **many words with unknown size**, we will be using LIST as a container to hold these data.

# Input for Text Concordance

- Passage usually consist of:
  - combination of words, sentences, and paragraphs.
  - *non-alphabetic characters* or symbols that is not a number or letter such as blank spaces, exclamation points (!), commas (, ), question marks (?), periods (.), underscores (\_), apostrophes ('), and at symbols (@).
  - blank line used to separate one paragraph to another.
- Extract all words and exclude
  - the *non-alphabetic characters* or symbols - [split](#) method
  - as well as the leading and trailing spaces - `trim` method

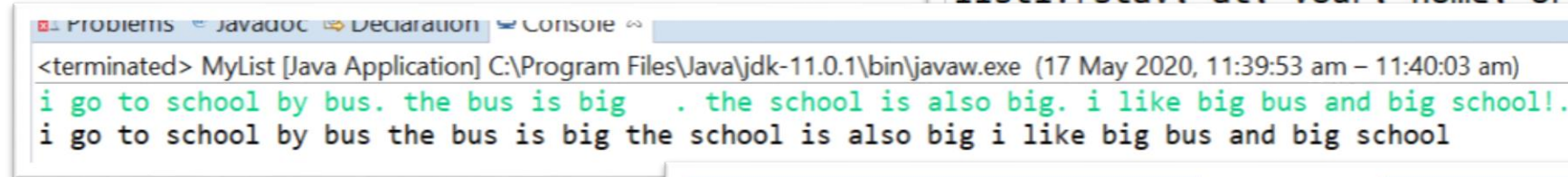
# Java Code: `split` and `trim` method

```
1  import java.util.*;
2
3  public class MyList {
4      public static void main(String args[]) {
5          Scanner in = new Scanner(System.in);
6          String passage = "Be safe..Stay home!. #Kita jaga kita." ;
7
8          String delims = "\\W+"; // split any non word
9          String [] words = passage.split(delims);
10         for (String str: words){
11             str = str.trim();
12             System.out.println(str);
13         }
14     }
}
```

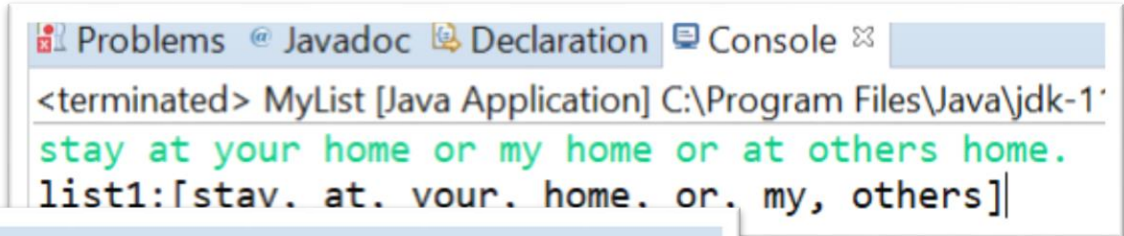
## Output:

```
Be
safe
Stay
home
Kita
jaga
kita
```

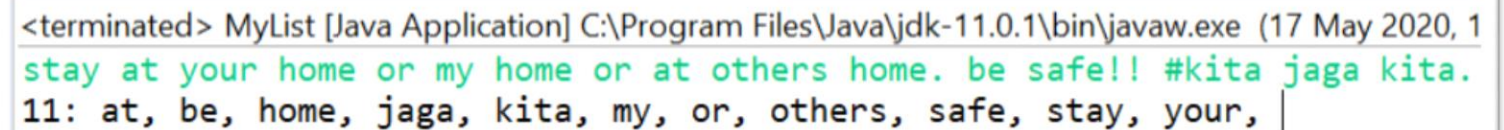
# Text Concordance:



```
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)
i go to school by bus. the bus is big . the school is also big. i like big bus and big school!.
i go to school by bus the bus is big the school is also big i like big bus and big school
```



```
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)
stay at your home or my home or at others home.
list1:[stav. at. vour. home. or. my, others]
```



```
<terminated> MyList [Java Application] C:\Program Files\Java\jdk-11.0.1\bin\javaw.exe (17 May 2020, 11:39:53 am – 11:40:03 am)
stay at your home or my home or at others home. be safe!! #kita jaga kita.
11: at, be, home, jaga, kita, my, or, others, safe, stay, your, |
```

1. Display all words from a passage
2. Display all unique words from a passage
3. Display all unique words in sorted order
4. Display all unique words with its frequencies in sorted order