# The General Problem of Describing Syntax

A Language, whether natural-like English, or artificial like C, Java, is a set of strings of characters from some alphabet. The strings of a language are called **Sentences or Statements**. So Syntax Rules of a language specify which strings of characters from the language's alphabet are in the language.

Formal descriptions of the Syntax of programming languages, for simplicity's sake, often donot include descriptions of the lowest-level Syntatic Units. These small units are called **Lexemes.**


Lexemes include its numeric literals, operators and special words, among others. We can think of programs as strings of lexemes rather than of characters.

Lexemes are partitioned into groups- for example. **The names of variables, methods, classes and so forth** in a programming language from a group called **identifiers.**

Each lexeme group is represented by a name, or token. **So, a token of a language is a category of its Lexemes.**

Consider the following Java Statement :-

```
index = 2 * count + 17;
```

| Lexemes | Tokens |
|---|---|
| index | identifier |
| = | equal_sign |
| 2 | int_literal |
| * | mult_op |
| Count | identifier |
| + | plus_op |
| 17 | int_literal |
| ; | Semicolon |

## 2 Distinct ways of defining a language :-

### A) • Language Recognizers :-

- A recognition device reads input strings of the language and decides whether the input strings belong to the language

**Syntax Analyzers :-** Determine whether the given program are syntatically correct.

### B) Language Generators :-

→ It generates sentences of a language.

→ people prefers certain forms of Generators over recognizers because they can more easily read and

Understand them.

By Contrast, the Syntax-Checking portion of a Compiler (a language recognizer) is not as useful as language description for a programmer because it can be used only in trial-and-error mode.

To determine Correct Syntax of a particular Statement using a Compiler, the programmer can only Submit Speculated Version and note whether the Compiler accepts it.

On the other hand, it is often possible to determine whether the Syntax of a particular Statement is Correct by Comparing it with the Structure of the generator