

# 1.B. Lenguajes de programación

## 4. Compilación.

### 4.1. Análisis léxico.

Wikipedia define el [analizador léxico](#) como:

Un analizador léxico **es la primera fase de un compilador**, consiste **en un programa que recibe como entrada el código fuente** (secuencia de caracteres, normalmente generada en un editor de textos) **y produce una salida compuesta de tokens** (componentes léxicos) **o símbolos**.

**Los tokens** sirven **para una posterior etapa del proceso de traducción**, siendo **la entrada para el analizador sintáctico**.

La especificación de un lenguaje de programación a menudo incluye **un conjunto de reglas que definen el léxico**. Estas reglas consisten **comúnmente en expresiones regulares**, que indican **el conjunto de posibles secuencias de caracteres que definen un token o lexema**.

En algunos lenguajes de programación **es necesario establecer patrones para caracteres especiales** (como el espacio en blanco) **que la gramática pueda reconocer sin que constituya un token en sí**.

**Por ejemplo**, en un lenguaje cuyos únicos lexemas sean: uno, dos, tres y cuatro.

- El texto uno uno tres cuatro, estaría formado por cuatro tokens: "uno" ; "uno" ; "tres" ; "cuatro"
- El texto uno una tres cuatro, produciría un error a nivel léxico, ya que al iniciar la búsqueda de un token tras el primer uno, no encuentra un lexema reconocible.

**Mientras existan errores léxicos no es posible completar la compilación del programa.**