

## Gramática

El léxico es una forma de hablar de expresar pensamientos, emociones, acciones, .... En pocas palabras se tiene la comunicación entre individuos o grupo de individuos, para una correcta comunicación en el "léxico" se debe tener en cuenta la sintaxis, lo anterior se refiere a que la formación de las palabras este construidas de forma correcta, por ejemplo:

Una sintaxis incorrecta es : COMDORATAPU

Una sintaxis correcta es : COMPUTADORA

Se debe notar que ambas palabras tienen las mismas letras, pero solo una tiene la sintaxis correcta.

Cuando se tiene una sintaxis correcta también es necesario tomar en cuenta la semántica, en esta etapa es cuando las palabras que se van a combinar tengan una sintaxis correcta, es decir las palabras tengan sentido, por ejemplo:

Una semántica correcta es: COMPUTADORA ES UN DISPOSITIVO

Una semántica incorrecta es: UN ES DISPOSITIVO COMPUTADORA

Nótese que ambas sentencias tienen el mismo número y tipo de palabras, pero solo una tiene una semántica correcta, es decir la oración completa tiene un sentido lógico.

Todo lo mencionado anteriormente aplicará únicamente para el lenguaje español o castellano, si se estuviera hablando de una lengua diferente las reglas sintácticas y semánticas cambiaría, en incluso para algunos idiomas será necesario utilizar diferentes letras, por ejemplo para una sintaxis y semántica correcta será necesario conocer el **alfabeto** ruso { А Б В Г Д Е Ё Ж З И Й К Л М Н О П Р С Т У Ф Х Ц Ч Ш Щ Ъ Ы Ь Э Ю Я }.

Tomando en cuenta todo lo anterior se puede plantear que un lenguaje es un conjunto de palabras, y las palabras están formadas por un conjunto o colección finita de símbolos llamado alfabeto.

Matemáticamente un **alfabeto** es un conjunto no vacío y finito el cual es representado por  $\Sigma$ .

$$\Sigma = \{ a, b, c, d, \dots, z \}$$

Con los elementos del conjunto  $\Sigma$  se pueden formar un número muy grande de combinaciones con los símbolos, sin embargo, solo algunas combinaciones son correctas, por ejemplo:

computadora -> con significado

abbbccddaaz -> sin significado

Aquellas palabras que tiene un significado correcto dentro del lenguaje entrarán dentro de un diccionario, es decir el diccionario es un subconjunto de  $\Sigma$ . Al conjunto de palabras se le denota como:

$$\Sigma^*$$

El conjunto de palabras que entrarían dentro de  $\Sigma^*$  podrían ser { el, hombre, compró, balón, se, un }. Cuando se tiene definido el diccionario el siguiente paso será evaluar la sintaxis y la semántica correcta de la oración, por ejemplo “el hombre compró un balón” tiene una sintaxis y semántica correcta.

Los lenguajes se clasifican en lenguajes naturales y artificiales, los lenguajes naturales tienen la característica que se forman antes de que existan las reglas gramaticales, por su parte los lenguajes artificiales o formales primero se definen las reglas para que exista el lenguaje.

Un ejemplo de lenguaje formal son los lenguajes de programación, para poder hacer uso del lenguaje primero se deben definir las reglas por ejemplo el tipo de variables, constantes, funciones, librerías, etc ... es decir hay unas reglas formales que deben de ser cumplidas a un nivel sintáctico y semántico para poder compilar ( traducir ) un programa y este pueda ser ejecutado por la computadora.

Ejemplos:

$$\Sigma = \{ a, b, c, \dots z \} \quad L = \{ u \in \Sigma^* \mid u \text{ está en el RAE} \}$$

U es una palabra

L es un lenguaje finito

$$\Sigma = \{ a, b, c \} \quad L = \{ u \in \Sigma^* \mid u \text{ no contiene símbolo } c \}$$

ababababa  $\in L$

abababcba  $\notin L$

$$\Sigma = \{ 0, 1 \} \quad L = \{ u \in \Sigma^* \mid u \text{ es cualquier secuencia binaria impar} \}$$

0011  $\in L$

0010  $\notin L$

En programación : )

$$\Sigma = \{ \text{ASCII} \}$$

$$\text{C++} = \{ \text{Reglas símbolos en C++} \}$$

$$\text{Java} = \{ \text{Reglas símbolos en Java} \}$$

Operaciones con Lenguajes:

#### Inversa

$$A^R = \{ u^R \mid u \in A \}$$

#### Concatenación

$$A \bullet B = \{ uv \mid u \in A, v \in B \}$$

$AB \neq BA$  No es conmutativa

#### **Gramática formal**

Una gramática formal consta de un conjunto finito de símbolos terminales (las palabras de un lenguaje formal), un conjunto finito de símbolos no terminales, conjunto de reglas de producción con un lado izquierdo y otro derecho, y un conjunto inicial.

Para aplicar las reglas se sustituye la parte izquierda de la derecha. Una derivación es una secuencia de aplicación de las reglas.

Cada gramática define el lenguaje formal de todas las sentencias que están formadas exclusivamente por los símbolos terminales a los que se puede llegar mediante derivación a partir del símbolo inicial.

$G(V, T, S, P)$

$V$  : Conjunto de componentes – léxico (símbolos terminales)

$T$  : Conjunto terminal de elementos

$S$  : Símbolo inicial

$P$  : Conjunto de producciones, en general tienen

la forma:  $\underset{\text{izq}}{\mathbf{a}} \rightarrow \underset{\text{dere}}{\mathbf{B}}$

Ejemplo 1:

¿Qué lenguaje es generado por la gramática siguiente ?

$G(V, T, S, P)$

donde:

$V = \{ S, A, a, b \}$

$T = \{ a, b \}$

$S = S$

$P = \{ S \rightarrow aA, S \rightarrow b, A \rightarrow aa \}$

Para dar respuesta se deben de aplicar todas las producciones posibles, hasta encontrar los resultados finales

### Ejemplo 2

La palabra **cbab** ¿ pertenece o no pertenece? al lenguaje generado por la siguiente gramática:

**$G(V, T, S, P)$**

donde :

**$V = \{ a, b, c, A, B, C, S \}$**

**$T = \{ a, b, c \}$**

**$S = S$**

**$P = \{ S \rightarrow AB$**

**$A \rightarrow Ca$**

**$B \rightarrow Ba$**

**$B \rightarrow Cb$**

**$B \rightarrow b$**

**$C \rightarrow cb$**

**$C \rightarrow b \}$**