

# Evaluación del Modulo 2

Laura Rodríguez Navas  
rodrigueznava@posgrado.uimp.es

Enero 2020

1. Explique cuáles son los dos grandes paradigmas del PLN e indique sus diferencias. Trate de ampliar la información que hay en la lección.

Los dos grandes paradigmas del PLN son la Lingüística generativa y la Tecnología de la Lengua.

La **Lingüística generativa**, también conocida como teoría racionalista y Chomskiana, describe las estructuras del lenguaje humano definidas en el cerebro (Lenguaje-I) a partir de las derivaciones de dicho lenguaje, las cuales se encuentran impresas en los textos.

La **Tecnología de la Lengua**, también conocida como Ingeniería Lingüística o PLN estadístico, modela e identifica los parámetros del lenguaje mediante el análisis y procesamiento exhaustivo del Lenguaje-E, a partir de la reproducción física de este.

La primera diferencia es que la Lingüística generativa fundamenta que el lenguaje es un mecanismo tan complejo que no puede ser adquirido por los sentidos. Contrariamente, la Tecnología de la Lengua fundamenta que la mente humana tiene la capacidad innata de establecer asociaciones, reconocer patrones y de generalizar ocurrencias de eventos que son percibidos a través de los sentidos. Por lo tanto, el lenguaje puede ser adquirido por los sentidos.

La segunda diferencia es que la Lingüística generativa se refiere al Lenguaje-I ("lengua interna" o interiorizada) que contrasta con el uso del Lenguaje-E ("lengua exterior") de la Tecnología de la Lengua. Técnicamente el Lenguaje-I se refiere a la representación mental o conocimiento lingüístico inconsciente que un hablante tiene de su lengua y por tanto es un objeto mental. En cambio, el Lenguaje-E abarca los aspectos de lengua relacionados con su uso social, los hábitos socio-lingüísticos y aspectos externos del uso de la lengua en comunidades humanas.

2. ¿Cuáles son los niveles de análisis o procesamiento de un sistema típico de PLN? No se limite a indicar los nombres, defínalos y explique las relaciones existentes entre ellos.

Los niveles de análisis o procesamiento de un sistema típico de PLN son: Tokenización y Segmentación, Análisis Léxico, Análisis Sintáctico, Análisis Semántico y Análisis Pragmático.

El proceso de **Tokenización y Segmentación** consiste en, dado un fragmento de texto, identificar palabras y tokens, así como el conjunto de oraciones en el que se organiza el texto

## El Análisis Léxico

El análisis léxico requiere haber identificado las unidades mínimas que constituyen un mensaje, así como su organización en oraciones. Por tanto primero se debe realizar el proceso de Tokenización y Segmentación.

El proceso de **Análisis Sintáctico** se analiza la estructura de un mensaje y si se adecua a la gramática de una lengua.

Antes del análisis sintáctico se requiere identificar las unidades mínimas del lenguaje: las palabras y tokens, así como toda la información relativa a ellas. Por tanto primero se debe realizar el análisis léxico.

El proceso de **Análisis Semántico** se identifica el significado de las palabras y de las figuras lingüísticas presentes en un mensaje.

El proceso de **Análisis Pragmático** se trata de determinar el significado del discursos subyacente en el mensaje.

3. Indique cuáles son las dependencias que se deben tener en cuenta a la hora de desarrollar un tokenizador. No se limite a nombrar las dependencias, desarrolle su respuesta.

La **Codificación de caracteres**

El **Idioma**

El **Corpus**

La **Aplicación**

4. Explique las diferencias existentes entre un analizador y un generador léxico. Se valorará positivamente la descripción de ejemplos de analizadores y generadores léxicos.
5. La lección se ha centrado en los analizadores sintácticos basados en Gramáticas Libres de Contexto, por lo que ¿podría describir las operaciones fundamentales de un analizador LR(0)?
6. Defina qué es la identificación de roles semánticos, así como qué es FrameNet. Recuerde que se valorarán positivamente la presentación de ejemplos.