



Universidad de Colima

Facultad de Ingeniería Mecánica y Eléctrica

Ingeniería en Computación Inteligente

Procesamiento del lenguaje natural

Oswaldo carrillo Zepeda

Dolores Fernanda Castillo De Dios

6ºD

Componentes del Procesamiento de Lenguaje Natural

Colima, col, miércoles 18 de mayo de 2022

## ¿Qué es?

El Procesamiento del Lenguaje Natural es el campo de conocimiento de la Inteligencia Artificial que se ocupa de la investigación la manera de comunicar las máquinas con las personas mediante el uso de lenguas naturales, como el español, el inglés o el chino.

Virtualmente, cualquier lengua humana puede ser tratada por los ordenadores. Lógicamente, limitaciones de interés económico o práctico hace que solo las lenguas más habladas o utilizadas en el mundo digital tengan aplicaciones en uso.

Pensemos en cuántas lenguas hablan Siri (20) o Google Assistant (8). El inglés, español, alemán, francés, portugués, chino, árabe y japonés (no necesariamente en este orden) son las que cuentan con más aplicaciones que las entienden. Google Translate es la que más lenguas trata, superando el centenar... pero hay entre 5000 y 7000 lenguas en el mundo.

## Componentes

- **Análisis morfológico o léxico:** Consiste en el análisis interno de las palabras que forman oraciones para extraer lemas, rasgos flexivos, unidades léxicas compuestas. Es esencial para la información básica: categoría sintáctica y significado léxico.
- **Análisis sintáctico:** Consiste en el análisis de la estructura de las oraciones de acuerdo con el modelo gramatical empleado (lógico o estadístico).
- **Análisis semántico:** Proporciona la interpretación de las oraciones, una vez eliminadas las ambigüedades morfosintácticas.
- **Análisis pragmático:** Incorpora el análisis del contexto de uso a la interpretación final. Aquí se incluye el tratamiento del lenguaje figurado (metáfora e ironía) como el conocimiento del mundo específico necesario para entender un texto especializado.

## Análisis morfológico o léxico

Su funcionalidad se apoya en identificar la interacción que está establecido en medio de las unidades mínimas que conforman un vocablo, como podría ser el reconocimiento de sufijos o prefijos.

El léxico es el grupo de información acerca de cada palabra que el sistema usa para el procesamiento.

Comúnmente el léxico solamente tiene la raíz de los vocablos con maneras regulares, siendo el analizador morfológico el que se ocupa de establecer si el género, número o flexión que conforman lo demás del vocablo son adecuados.

### Análisis sintáctico

El resultado de este proceso se apoya en producir la composición que corresponde a las categorías sintácticas formadas por todas las unidades léxicas que aparecen en la frase.

### Análisis semántico

En muchas aplicaciones del PLN las metas del estudio apuntan hacia el procesamiento del sentido.

No obstante, las técnicas de representación del sentido no han obtenido los resultados deseados, y varias preguntas continúan sin descubrir resoluciones satisfactorias.

Realizando uso de esta concepción modular es viable diferenciar entre sentido sin dependencia y sentido dependiente del entorno.

El primero, tratado por la semántica, se refiere al sentido que los vocablos poseen por sí mismas sin tener en cuenta el sentido adquirido de acuerdo con el uso en una cierta situación. Por otro lado, el elemento relevante de una sentencia asociado a las situaciones en que ésta se da, es estudiado por la pragmática y conocido como sentido dependiente del entorno.

### Análisis pragmático

Incorpora información adicional al estudio del sentido de la sentencia en funcionalidad del entorno donde surge. Hablamos de uno de los niveles de estudio más complicados, el fin del cual es integrar al estudio semántico la aportación significativa que tienen la posibilidad de hacer los competidores, la evolución del discurso o información presupuesta.