

PROCESAMIENTO DEL LENGUAJE NATURAL

DR. JOSE EDUARDO FERRER CRUZ



SESIÓN DE CIERRE

AGENDA

- RESUMEN DEL CURSO
- EXAMEN FINAL

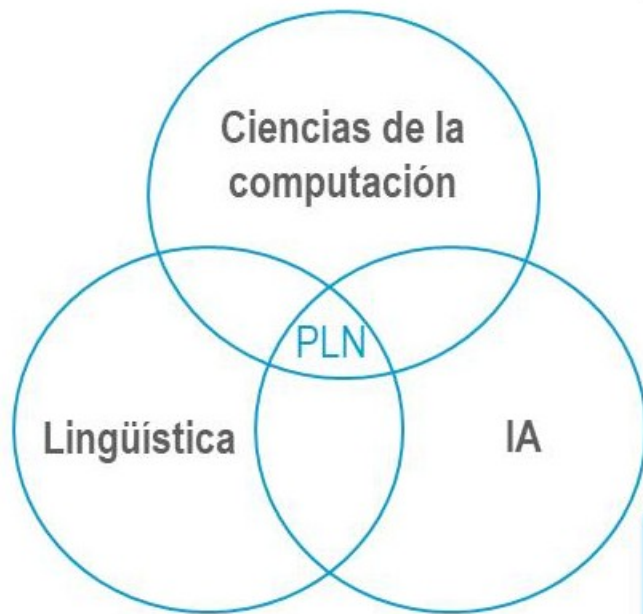
CURSO

1. INTRODUCCIÓN
2. ANÁLISIS MORFOLOGICO
3. POS TAGGING
4. GRAMATICAS PARA EL ANALISIS SINTACTICO
5. ANALISIS SINTACTICO
6. SEMANTICA
7. ANALISIS SEMANTICO
8. SEMANTICA LEXICA
9. RECURSOS PARA EL PROCESAMIENTO DE LENGUAJE NATURAL
10. AGENTES CONVERSACIONALES

Procesamiento de lenguaje natural

«Campo de la inteligencia artificial que tiene como objetivo que las máquinas sean capaces de realizar tareas que involucren el lenguaje humano».

«Conjunto de herramientas, algoritmos, modelos... que permiten procesar un texto en lenguaje natural en un conjunto de datos estructurados para ser procesados por un algoritmo».

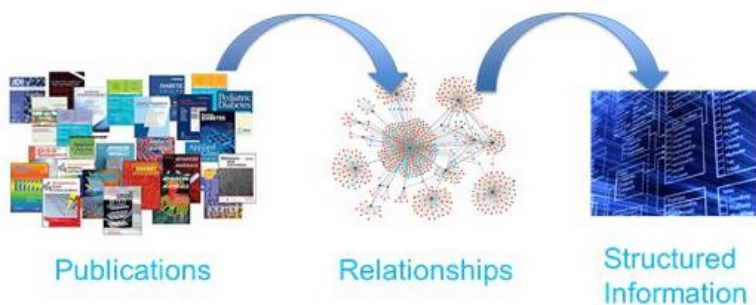


Objetivos

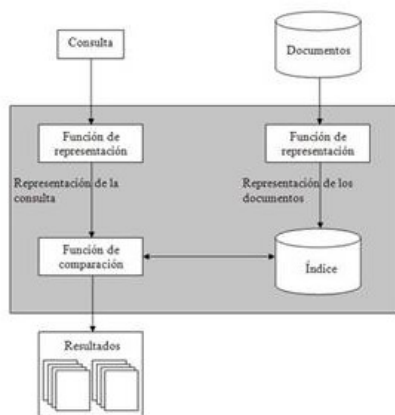
- ▶ Asimilar el concepto de PLN.
- ▶ Conocer las áreas de conocimiento implicadas.

Aplicaciones del PLN

Extracción de información



Recuperación de información



Análisis de sentimientos



Sistemas de diálogo o agentes conversacionales



Traducción automática



Objetivos

- ▶ ¿Dónde se aplica?
- ▶ ¿Qué tipo de resultados se pueden obtener?

Fuentes:

<https://www.questionpro.com/blog/es/herramienta-de-analisis-de-sentimientos/>

<https://www.pngwing.com/es/free-png-pbise>

Conocimiento del lenguaje

Fonética y fonología

Conocimiento sobre los sonidos lingüísticos.

Morfología

Conocimiento de los componentes significativos de las palabras.

Sintaxis

Conocimiento de las relaciones estructurales entre palabras.

Semántica

Conocimiento del significado.

Pragmática

Conocimiento de la relación del significado con los objetivos y las intenciones.

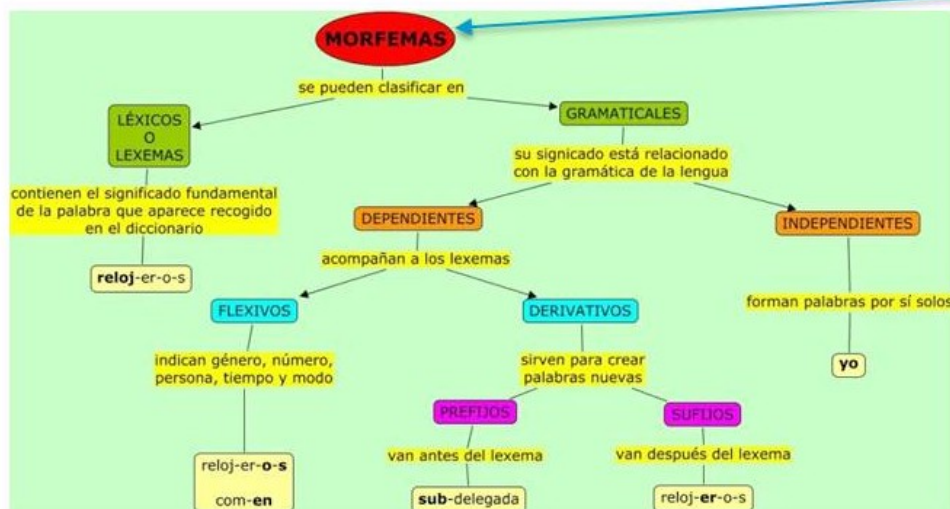
Discurso

Conocimiento sobre unidades lingüísticas más grandes que un solo enunciado.

Morfología: Morfemas

Parte de la gramática que estudia la estructura de las palabras y de sus elementos constitutivos (Drae, s.f.).

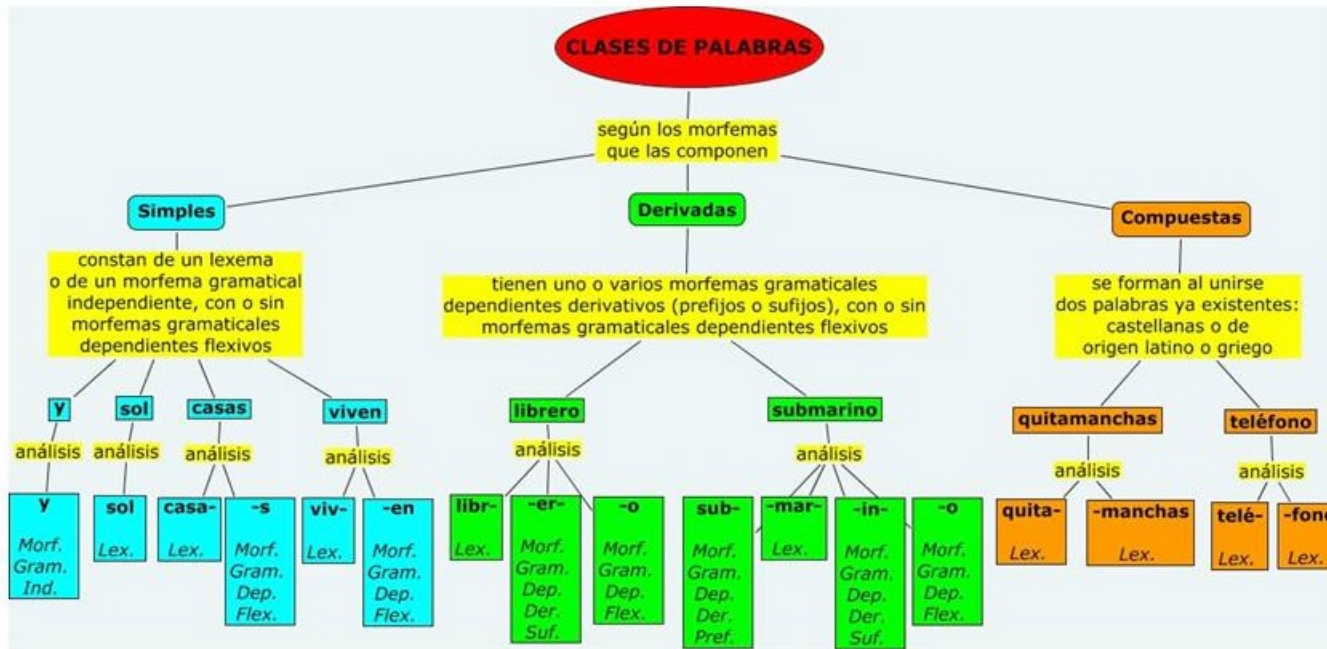
Unidades mínimas con significado que forman las palabras.



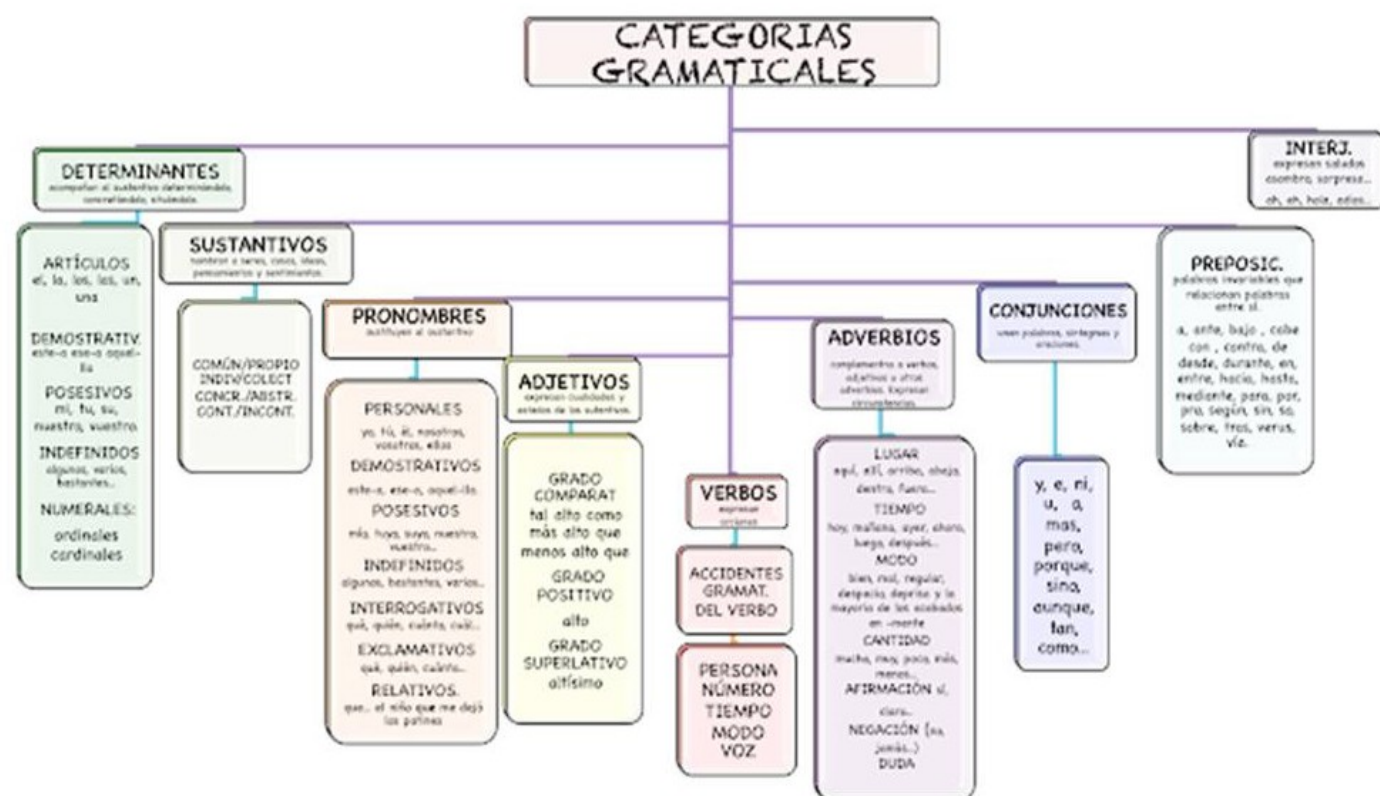
Objetivos

- Conocer de qué elementos están formadas las palabras.
- Conocer cómo se modifica el significado de una palabra según los morfemas que la componen.

Tipos de palabras y elementos del análisis morfológico

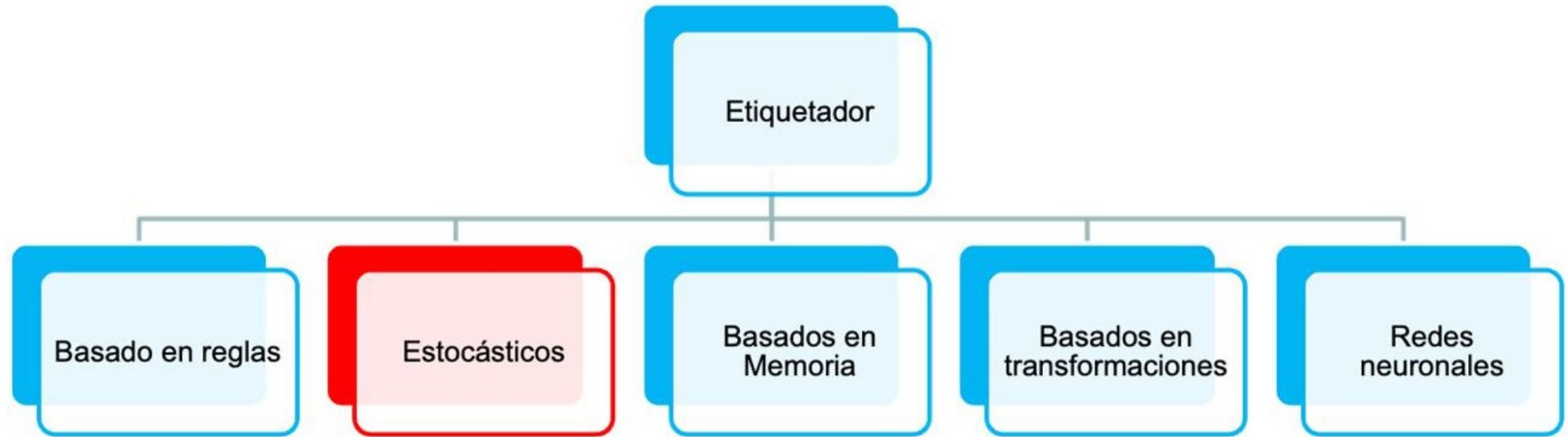


Categorías morfosintácticas o gramaticales



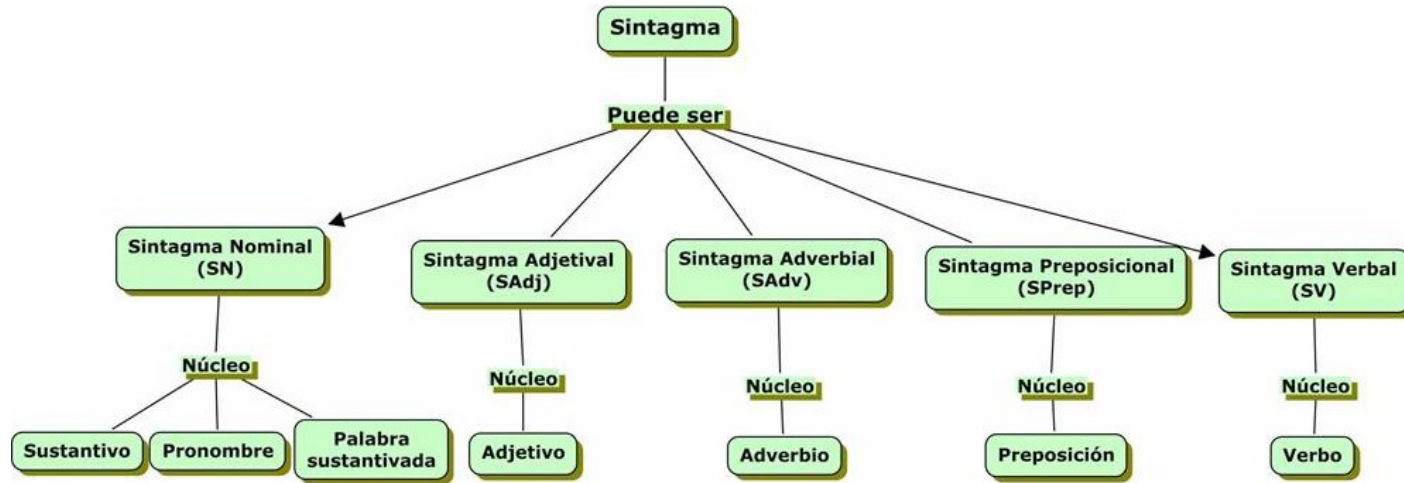
«Tipos de palabras de acuerdo con la función que desempeñan en la oración».

Tipos de etiquetadores morfosintácticos

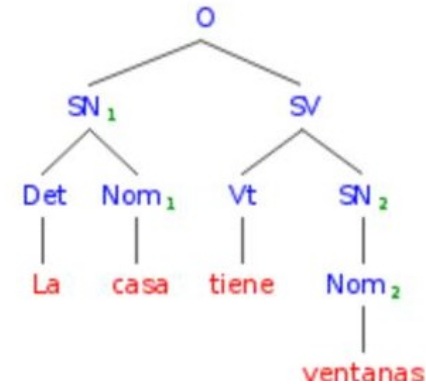


Sintaxis, constituyentes y árboles sintácticos

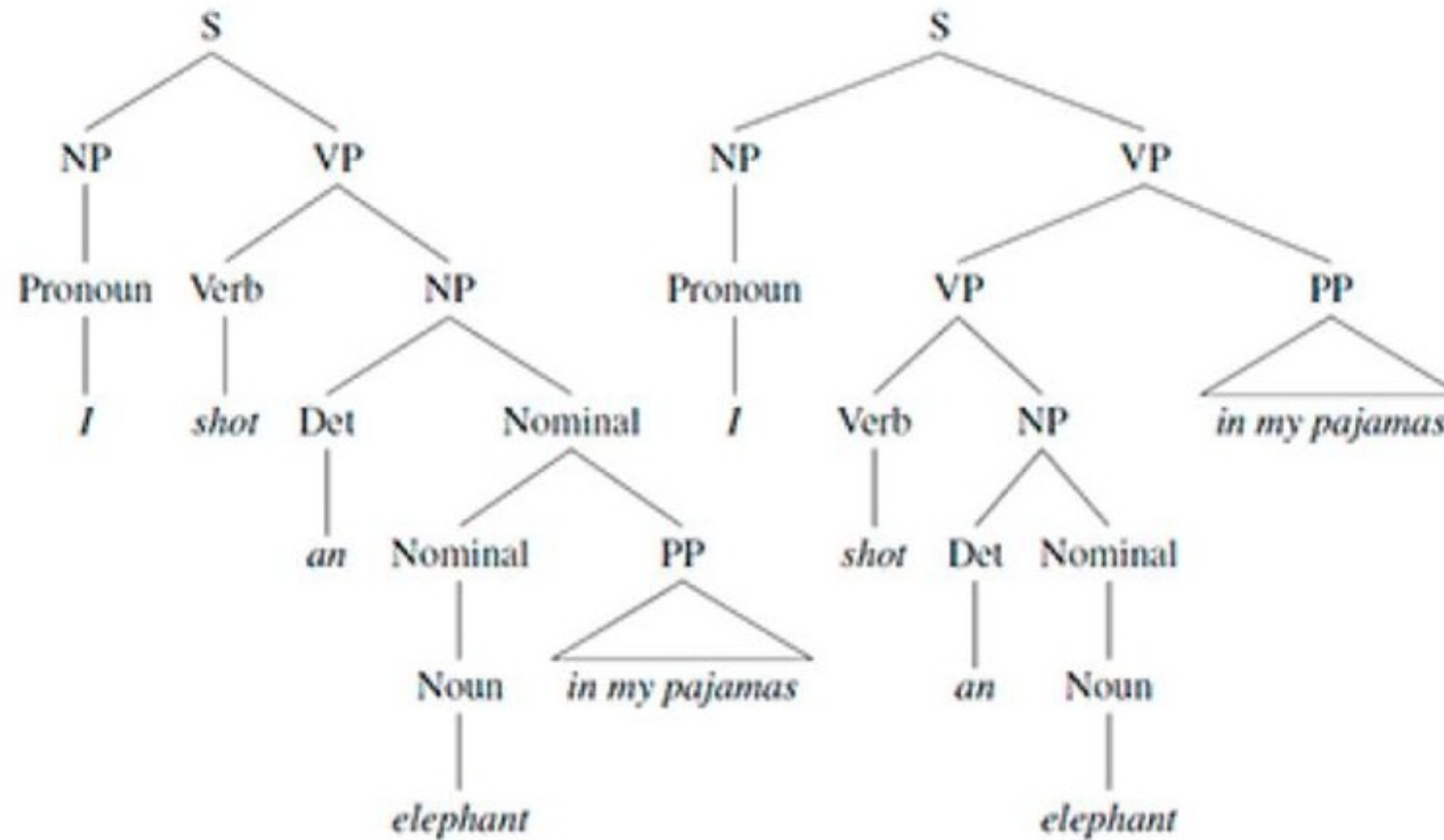
«Parte de la gramática que estudia el modo en que se combinan las palabras y los grupos que estas forman para expresar significados, así como las relaciones que se establecen entre todas esas unidades.» (RAE, s.f.).



Sintagma: grupo o conjunto de palabras que tienen una función determinada.

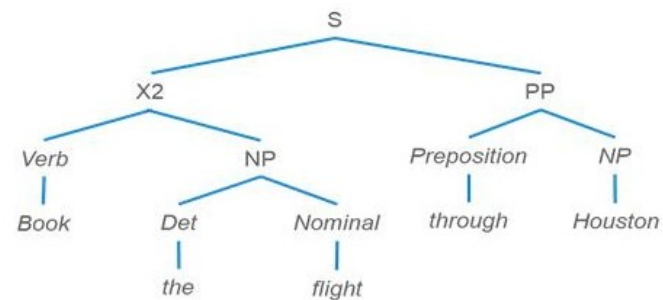
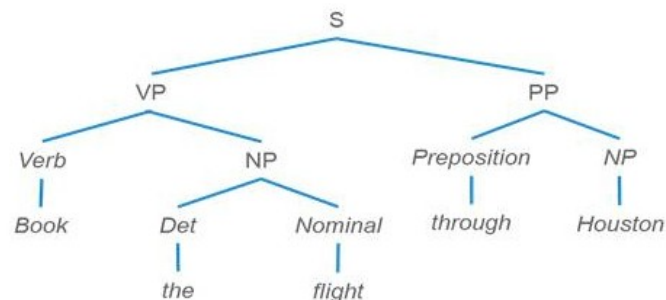
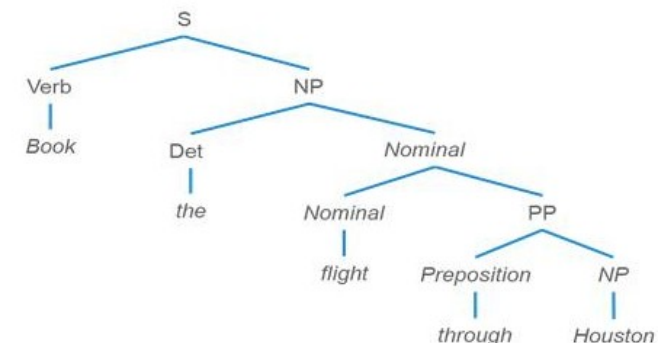


Ambigüedad en el análisis sintáctico



Métodos para análisis sintáctico basados en programación dinámica

Algoritmo de Cocke-Kasami-Younger (CKY)



Semántica y sus tipos

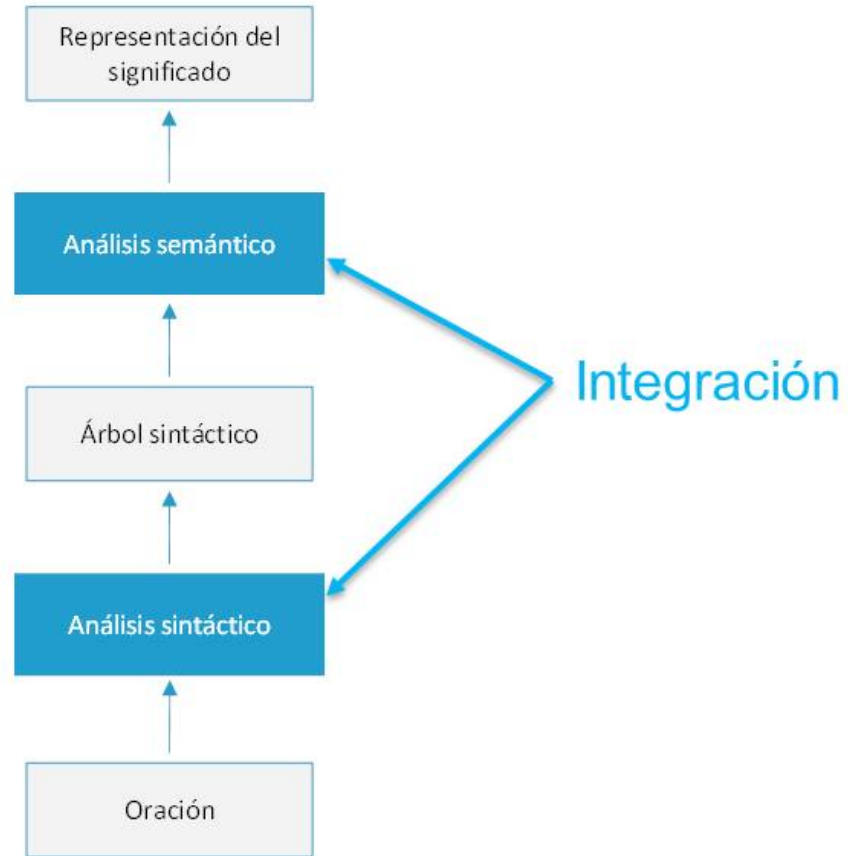
Semántica

«Disciplina que estudia el significado de las unidades lingüísticas y de sus combinaciones» (RAE, s. f.)

Semántica léxica

«Estudia el significado de las palabras, así como las diversas relaciones de sentido que se establecen entre ellas» (RAE, s. f.)

Análisis semántico dirigido por la sintaxis



Objetivos

- Conocer las fases que se realizan en un análisis semántico dirigido por la sintaxis.

Significado de una palabra y sus sentidos

Lema

Se considera como la raíz de la palabra, representa el significado y es la forma gramatical que se usa en los diccionarios.

- ▶ Homonimia → Ambigüedad léxica.
- ▶ Homografía → Conversión del texto al habla.
- ▶ Homofonía → Errores ortográficos.

1. Definir el sentido de una palabra.
2. Agrupar los sentidos de las palabras.

Diccionario



Base de datos léxica

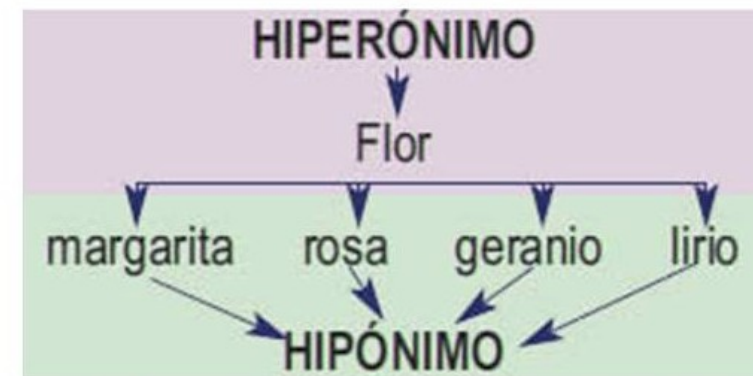


PRINCETON UNIVERSITY

WordNet

A Lexical Database for English

Relaciones entre las palabras



Desambiguación del sentido de las palabras

Algoritmos basados en aprendizaje supervisado

Entrenan un modelo clasificador que permita desambiguar las palabras

Algoritmos basados en conocimiento

Realizan un entrenamiento indirecto con aprendizaje supervisado débil

Algoritmos basados en aprendizaje semi supervisado

Aprenden a través de *bootstrapping*

Algoritmos basados en aprendizaje no supervisado

Inducción del sentido de las palabras

Recursos lingüísticos

love noun

 Save Word

\ Tav 🔊 \

Essential Meaning of *love*


- 1** : a feeling of strong or constant affection for a person
// motherly/maternal *love*
// fatherly/paternal *love*
See More Examples
- 2** : attraction that includes sexual desire : the strong affection felt by people who have a romantic relationship
// a declaration of *love*
// He was just a lonely man looking for *love*.
See More Examples
- 3** : a person you love in a romantic way
// a lost *love*
// You never forget your **first love**. [=the first person you loved in a romantic way]

Diccionario

Repertorio donde se recogen, según un orden determinado, las palabras de una lengua acompañadas de su definición o explicación con información sobre diferentes sentidos.

Thesaurus

love verb

 Save Word

Synonyms & Antonyms of *love* (Entry 1 of 2)

- 1** to hold dear
// patriots who *loved* their country well enough to die for it

Synonyms for *love*

appreciate, cherish, prize, treasure, value

Words Related to *love*

delight (in), dig, enjoy, fancy, groove (on), like, relish, revel (in)
admire, apprise, esteem, regard, respect, revere, reverence, venerate
enshrine, memorialize
adore, caress, dote (on), idolize, worship

Phrases Synonymous with *love*

set store by (or set store on)

Near Antonyms for *love*

undervalue
abhor, abominate, despise, detest, execrate, hate, loathe
disdain, high-hat, scorn, scout, slight, sniff (at), snub
bad-mouth, belittle, cry down, decry, deprecate, depreciate, disparage, kiss off, minimize, put down, write off
abandon, forget, neglect

Tesauro

Lista de significados de las palabras y las relaciones entre estos significados.


Lexicón

El vocabulario de una persona, lengua o rama del conocimiento.

Recursos lingüísticos

Corpus

colección de textos representativos de una lengua que se utilizan para el análisis lingüístico

 GOLDENAGE.CERVANTESVIRTUAL.COM

[¿ SOBRE EL CORPUS](#) [ENCUESTA](#)

CORPUS DE
SONETOS
SIGLO DE ORO

Autor

Introduce un nombre de autor (ej:Lope de Vega).

Título

Búsqueda por palabras en el título de la obra

Poema

Búsqueda por palabras en el título del poema

Verso

Búsqueda por palabras en el verso

Métrica

Introduce una métrica siguiendo la documentación del corpus (ej:+++++---) o con números separados por comas marcando la secuencia de sílabas tónicas (ej:1,2,4,6,10).

Filtrar por tipo de estrofa Todos ▾

Ordenado por Relevancia ▾

Número de resultados 10 ▾

INSTRUMENTOS
INTERPRETACIÓN
DE LA LENGUA

utilir

Herramientas y bibliotecas para PLN en Python

Natural Language Toolkit

- Developers Guide
- Integration Testing
- Adding a Corpus
- Book Development
- Dependency Parsing
- Machine Translation
- Semantic Role Labeling
- Semantic Parsing
- Sentiment Analysis
- Topic Modeling
- Twitter Processing

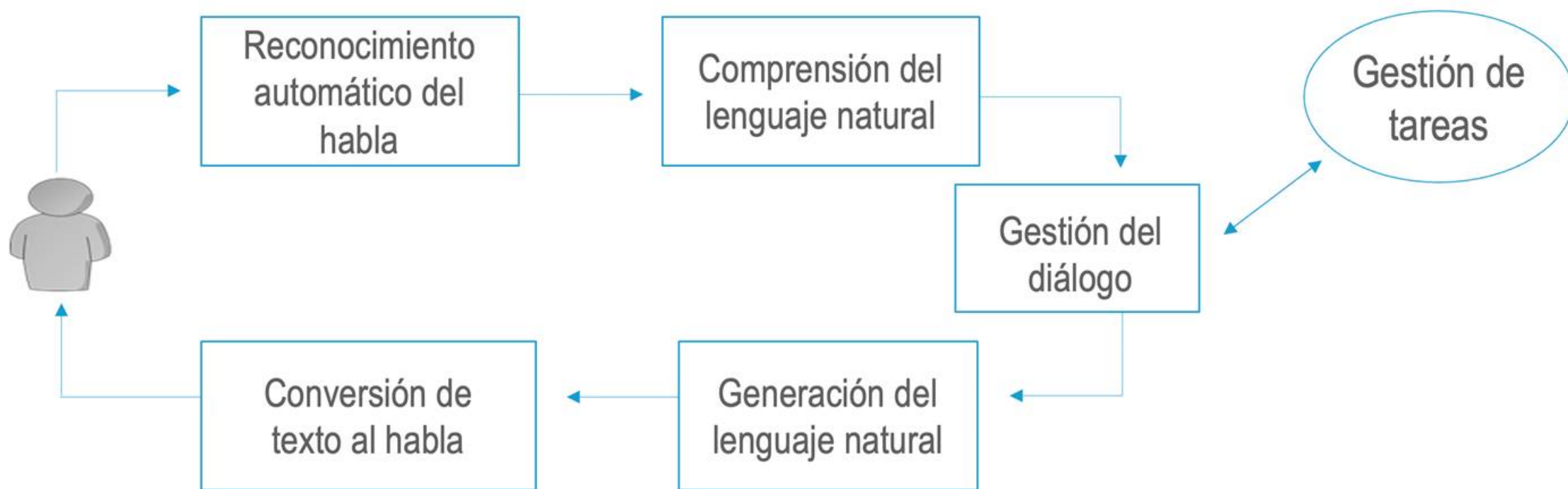
Fuente: <https://github.com/nltk/nltk/wiki>



Conversaciones



Estructura de un agente conversacional



Chatbots



Med Tech Company

Discover how this Med Tech company is leveraging conversational AI to accelerate its digital transformation strategy.

[VIEW CASE STUDY](#)

Swiss Telco

This leading Swiss telco uses Teneo to deliver conversational solutions 24/7 that are easy to deploy, are integrable with other components and adapt to Switzerland's multilingual environment.

[VIEW CASE STUDY](#)

Swedish Insurer

One of Sweden's biggest mutual insurance companies enabled its internal Center of Excellence, to investigate enterprise-grade Conversational AI solutions as a long-term opportunity to build even better customer experience and service, and firmly position them as an "insuretech company".

[VIEW CASE STUDY](#)

CPaaS Integration

This global financial organization used the Teneo Platform to build a conversational application that would allow their CRM teams to submit reports on the phone, by conversing with a Virtual Assistant (VA) in a very natural way.

[VIEW CASE STUDY](#)

RPA Integration

With easy integration into a wide range of back-end technologies such as Robotic Process Automation (RPA), Teneo is enhancing the customer experience and increasing business productivity.

[VIEW CASE STUDY](#)

Circle K

In what has anecdotally been described as one of the quickest projects in Circle K, virtual customer agent Kay is already reducing call volumes, while delivering 24/7 customer support, having been developed in only a few months with Teneo from Artificial Solutions.

[VIEW CASE STUDY](#)

EXAMEN FINAL

- 30 de septiembre
- Consta de 10 preguntas
 - 9 de opción múltiple / falso-verdadero
 - 1 pregunta abierta

EXAMEN FINAL

- ▶ 30 de septiembre
- ▶ Consta de 10 preguntas
 - 9 de opción múltiple / falso-verdadero
 - 1 pregunta abierta

Es el morfema que conforma la raíz o parte invariante de la palabra y proporciona el significado principal de la palabra.

A.Sintagma

B.Lexicón

C.Lexema

EXAMEN FINAL

- ▶ 30 de septiembre
- ▶ Consta de 10 preguntas
 - 9 de opción múltiple / falso-verdadero
 - 1 pregunta abierta

La superdiagonal de la matriz de análisis sintáctico contiene las categorías gramaticales para cada una de las palabras de la oración.

a. VERDADERO

b. FALSO

EXAMEN FINAL

- ▶ 30 de septiembre
- ▶ Consta de 10 preguntas
 - 9 de opción múltiple / falso-verdadero
 - 1 pregunta abierta

El algoritmo basado en _____ utiliza grandes bases de datos en lugar de corpus anotados y se basa en el número de palabras comunes en las definiciones de sentido, así como de grafos o redes semánticas

A.aprendizaje supervisado

B.conocimiento

C.aprendizaje semi supervisado

EXAMEN FINAL

- 30 de septiembre
- Consta de 10 preguntas
 - 9 de opción múltiple / falso-verdadero
 - 1 pregunta abierta

¿Qué es el NLTK?

Es un conjunto de bibliotecas y programas para el procesamiento del lenguaje natural (PLN) simbólico y estadísticos para el lenguaje de programación Python.



DR. JOSE EDUARDO FERRER CRUZ

www.unir.net