



Facultad de Ingeniería Mecánica y Eléctrica
Ingeniería en Computación Inteligente

Materia: Optativa I (Procesamiento de lenguaje natural).

Profesor: Oswaldo Carrillo Zepeda
Alumno: Jazmín Azucena González Peredia

6^{to} semestre.
Grupo B.

Niveles de PLN.

25 de mayo de 2022.

Niveles de PLN.

Comprender el lenguaje natural exige:

- Extraer el significado individual de las palabras
- Comprender las relaciones entre las palabras
- Referir el significado literal al contexto de actuación del sistema

Todo esto se alcanza a través de un análisis de los componentes del lenguaje a *diferentes niveles*.

Niveles de descripción del lenguaje:

Nivel fonético	
Nivel fonológico	
Nivel léxico	<i>Léxicos computacionales</i>
Nivel morfológico	<i>Analizadores morfológicos</i>
Nivel sintáctico	<i>Analizadores sintácticos</i>
Nivel semántico	<i>Redes semánticas (WordNet) Corpus</i>
Nivel pragmático	
Nivel ilocutivo	

Nivel de análisis: Fonética.

Estudia la acústica (cómo se articulan) y fisiología (cómo funcionan los órganos –lengua, paladar, labios, nariz, etc.) al producir los sonidos.

- Los sonidos son materiales y audibles.

saco, zapato, cepa

grafías diferentes → mismo fonema /s/.

Nivel de análisis: Fonológico.

- Es el tratamiento de los sonidos para detectar unidades de expresión (palabras).
- Parte de la lingüística que estudia los fonemas o descripciones teóricas de los sonidos vocálicos y consonánticos que forman una lengua.
- Se aplica en el procesamiento del lenguaje oral.
- Investiga qué diferencias fónicas están relacionadas con diferencias de significado, qué relaciones establece un fonema con los demás dentro de un paradigma

cala /k/

pala /p/

mala /m/

sala /s/

bala /b/

-ala (el fonema omitido determina la diferencia)

- Un fonema es la representación abstracta de un sonido que pertenece a la lengua, al sistema.
- No hay correspondencia biunívoca entre fonema y grafía (letras).

/b/ → {b, v}

/y/ → {ll, y (inicial de sílaba)}

/k/ → {k, qu, c (ca,co,cu)}

/s/ → {s, z, c (ce,ci)}

Fonética y fonología.

- Sistemas de análisis y síntesis del habla.
- Reconocimiento de voz.
- Generación de voz.
- Estos desarrollos se dirigen hacia el interés de diseñar programas para el dictado y reconocimiento automático.
- No presentan aún desarrollos significativos puesto que no se logra modelar flexiones y variaciones en entonaciones particulares en las lenguas.

Nivel de análisis: Léxico.

El concepto de léxico encierra varios significados, todos ligados al mundo de lingüística. Léxico es el vocabulario de un idioma o de una región, el diccionario de una lengua o el caudal de modismos y voces de un autor.

- Distingue entre palabras ortográficas y palabras gramaticales.
- Obtiene información léxica de diccionarios, ontologías..

Léxicos computacionales.

- Diccionarios de uso común en soporte electrónico
- Volcado automático a una estructura de base de datos
- Vocabulario controlado
- Granularidad uniforme
- Aportan diferentes tipos de información léxica

Nivel de análisis: Morfológico.

La morfología estudia la estructura de las palabras y su relación con las categorías del lenguaje.

- El objetivo del análisis morfológico automático es llevar a cabo una clasificación morfológica de las palabras.
- Por ejemplo, el análisis de la palabra gatos resulta en **gato+Sustantivo+Genero+Plural**,

que nos indica que se trata de un sustantivo plural con género masculino y que su forma normalizada (lema) es gato.

Analizadores morfológicos:

- Análisis y síntesis automático morfológico (sufijos, prefijos, raíces y flexiones)
- Morfotáctica: reglas de combinación de morfemas
- Variaciones fonológicas: cambios al combinar morfemas
- Hace falta la construcción de diccionarios robustos
- Aunque ha habido grandes desarrollos, se requiere una estandarización efectiva del lenguaje para la marcación

Nivel de análisis: Sintáctico.

- Incluye todo lo relacionado con el significado de las palabras.
- Se ocupa también de la clasificación de las palabras según su origen,
- Su uso, sus posibles transformaciones...
- Reconoce, extrae y representa sintácticamente válidas (o inválidas):

Es especialmente difícil automatizar en español donde no hay un orden fijo de las palabras que componen una oración.

Ejemplo:

- ¿Estuviste en el parque el sábado?
- ¿El sábado estuviste en la parque?
- ¿Tu estabas en el parque el sábado?

Nivel de análisis: Semántico.

- Estudia los cambios de significación de palabras/grupos de palabras y el papel que desempeñan las categorías y sus combinaciones en el significado de la frase.
- Dependiendo de la categoría sintáctica se puede hacer referencia a diferentes tipos de la realidad.
- Construcción de redes semánticas donde se presenten conceptos y relaciones entre ellos.

Redes semánticas.

- Una red semántica es una manera de representar el conocimiento basada en las relaciones que establecen los conceptos representados en las redes lexicosemánticas se representan los conceptos lexicalizados.
- EuroWordNet es la red lexicosemántica más extendida, de manera que se ha constituido como un estándar de representación castellano, catalán e inglés.
- Se utiliza para:
 - ◆ El etiquetado automático de corpus
 - ◆ Léxico computacional para obtener relaciones entre palabras
 - ◆ Recuperación de información

Nivel de análisis: Pragmático.

Interpretación en un contexto determinado (incorpora referencias implícitas)

- ◆ “Le dio un libro”
- ◆ “No les gustó”
- Se encarga de como el contexto (situación) influye en la interpretación del significado.
- Se preocupa como utilizamos el lenguaje para obtener diversos objetivos en el mundo que nos rodea. Utilizamos el lenguaje para informar, prometer, pedir o preguntar entre otras funciones.

Nivel de análisis: Ilocutivo.

- Es la idea o el concepto de la frase, es decir, aquello que se dice.
- Es la intención o finalidad concreta del acto de habla.
- Tiene problemas de asignación de intenciones:

“Los platos están sucios” (por tanto, ¡lávalos!)

Bibliografía.

Calderón Nepamuceno, D. M. (2019, septiembre). *Tema 1.3: Niveles de descripción del Lenguaje*.

Uaemex.mx. https://ri.uaemex.mx/bitstream/handle/20.500.11799/108125/secme-28839_1.pdf?sequence=1&isAllowed=y