

# La potencia de WATT

La traducción terminológica auxiliada por Wikipedia

Octavio Augusto Sánchez Velázquez

`osanchezv@iingen.unam.mx`

Grupo de Ingeniería Lingüística

Facultad de Ingeniería

Facultad de Filosofía y Letras

Universidad Nacional Autónoma de México

Dirección Adjunta de Competitividad

INFOTEC

Coloquio de Lingüística Computacional, 2013

- 1 Motivación
  - Descripción del problema

- 1 Motivación
  - Descripción del problema
- 2 Wikipedia
  - Descripción de Wikipedia
  - Estructura
  - Vínculos entre lenguas

- 1 Motivación
  - Descripción del problema
- 2 Wikipedia
  - Descripción de Wikipedia
  - Estructura
  - Vínculos entre lenguas
- 3 WATT
  - Una propuesta de solución
  - Posibilidades de expansión

# Un mundo más pequeño

- Incremento de las telecomunicaciones

# Un mundo más pequeño

- Incremento de las telecomunicaciones
- Mayor contacto de idiomas

# Un mundo más pequeño

- Incremento de las telecomunicaciones
- Mayor contacto de idiomas
- Contacto de diferentes culturas

# Visiones de mundo en contacto

- La lengua codifica nuestra visión de mundo



# Visiones de mundo en contacto

- La lengua codifica nuestra visión de mundo
- Usamos términos para referirnos a conceptos específicos

# Visiones de mundo en contacto

- La lengua codifica nuestra visión de mundo
- Usamos términos para referirnos a conceptos específicos
- Entre más especializado el término, más específico el contexto

# Visiones de mundo en contacto

- La lengua codifica nuestra visión de mundo
- Usamos términos para referirnos a conceptos específicos
- Entre más especializado el término, más específico el contexto
- Un término puede o no tener equivalentes en otra lengua

# Visiones de mundo en contacto

- La lengua codifica nuestra visión de mundo
- Usamos términos para referirnos a conceptos específicos
- Entre más especializado el término, más específico el contexto
- Un término puede o no tener equivalentes en otra lengua
- Una nueva realidad conlleva la necesidad de nombrarla

# Comunicación entre dos no nativos

- Dos hablantes pueden tener como L2 una tercera

# Comunicación entre dos no nativos

- Dos hablantes pueden tener como L2 una tercera
  - Hablante 1: Habla español como lengua materna e inglés como segunda

# Comunicación entre dos no nativos

- Dos hablantes pueden tener como L2 una tercera
  - Hablante 1: Habla español como lengua materna e inglés como segunda
  - Hablante 2: Habla chino como lengua materna e inglés como segunda

# Comunicación entre dos no nativos

- Dos hablantes pueden tener como L2 una tercera
  - Hablante 1: Habla español como lengua materna e inglés como segunda
  - Hablante 2: Habla chino como lengua materna e inglés como segunda
- Algunas veces pueden existir dificultades en la comunicación



# Comunicación entre dos no nativos

- Dos hablantes pueden tener como L2 una tercera
  - Hablante 1: Habla español como lengua materna e inglés como segunda
  - Hablante 2: Habla chino como lengua materna e inglés como segunda
- Algunas veces pueden existir dificultades en la comunicación
- Se conoce el concepto pero se desconoce el término que lo denomina

# ¿Qué es Wikipedia?

- Enciclopedia web libre y gratuita

# ¿Qué es Wikipedia?

- Enciclopedia web libre y gratuita
- Editada de manera colaborativa

# ¿Qué es Wikipedia?

- Enciclopedia web libre y gratuita
- Editada de manera colaborativa
- Multilingüe

# ¿Qué es Wikipedia?

- Enciclopedia web libre y gratuita
- Editada de manera colaborativa
- Multilingüe
- Autorregulada

# Estructura

Guitarra - Wikipedia, la enciclopedia libre - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W Guitarra - Wikipedia, la enciclope...

es.wikipedia.org/wiki/Guitarra

Google

Artículo Discusión Leer Editar Editar fuente Ver historial Buscar

## Guitarra

*Para la planta del mismo nombre, véase *Dipsacus fulorum*.*

La **guitarra** es un instrumento musical de cuerda pulsada, compuesto de una caja de madera, un mástil sobre el que va adosado el diapasón o trastero —generalmente con un agujero acústico en el centro de la tapa (boca)—, y seis cuerdas. Sobre el diapasón van incrustados los trastes, que permiten las diferentes notas. Su nombre específico es **guitarra clásica**, **guitarra española**, **guitarra criolla** o **guitarra acústica**.

Es el instrumento más utilizado en géneros como *blues*, *rock*, *metal* y *flamenco*, y bastante frecuente en *cantautores*. También es utilizada en géneros tales como el *tango*, *rancheras* y *gruperas*, además del *folclore* de varios países.


Es parte fundamental de la orquesta de *pulso* y *púa* o *rondalla*, junto a la *bandurria* y el *laúd español*.

Instrumentos de la familia de la guitarra son el *requinto*, el *charango* y el *guitarrón*. Este último es de uso frecuente por los *mariaichis*.


**índice** [ocultar]

- 1 Partes de la guitarra
  - 1.1 Caja de resonancia
  - 1.2 Mástil
- 2 Historia
  - 2.1 Origen
  - 2.2 Edad Media
  - 2.3 Siglo XVI
  - 2.4 Siglo XVII
  - 2.5 Siglos XVII y XIX
  - 2.6 Siglo XX
- 3 Tipos de guitarra
  - 3.1 Guitarra acústica
  - 3.2 Guitarra de flamenco
  - 3.3 Guitarra italiana
  - 3.4 Guitarra eléctrica
  - 3.5 Guitarra electroacústica

**Guitarra**



**Tesitura**  
*(loco)*



**Características**

**Clasificación**    Cordófono (de cuerda pulsada)

**Instrumentos relacionados**    laúd, bandurria, guitarra eléctrica, bajo, ukelele, charango, guitarrón, cuatro, cavaquinho, timbal, bongo, djembe

Portada

Portal de la comunidad

Actualidad

Cambios recientes

Página nuevas

Página aleatoria

Ayuda

Donaciones

Notificar un error

▼ Imprimir/exportar

  Crear un libro

  Descargar como PDF

  Versión para imprimir

► Herramientas

► Otros proyectos

▼ En otros idiomas

Afrikaans

Alemannisch

АԥсӀӀ

Aragonés

Ænglisc

العربية

עברית

Астurianu

# Problemas en la estructura

- Existe la plantilla,

# Problemas en la estructura

- Existe la plantilla, pero no se usa siempre
- Las categorías son arbitrarias



# Problemas en la estructura

- Existe la plantilla, pero no se usa siempre
- Las categorías son arbitrarias
- No existen los mismos artículos en diferentes idiomas

# Problemas en la estructura

- Existe la plantilla, pero no se usa siempre
- Las categorías son arbitrarias
- No existen los mismos artículos en diferentes idiomas
- Casi ningún artículo existe en todos los idiomas

# Explotación de los vínculos de idioma

- Podemos seguir un vínculo a otro idioma

# Explotación de los vínculos de idioma

- Podemos seguir un vínculo a otro idioma
- El título en el nuevo idioma será el término

# Explotación de los vínculos de idioma

- Podemos seguir un vínculo a otro idioma
- El título en el nuevo idioma será el término
- Podemos conocer un término en una lengua desconocida

# Ejemplo

Гитара — Википедия - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W Гитара — Википедия

ru.wikipedia.org/wiki/Гитара

Google

Создать учётную запись Представиться системе

Статья Обсуждение Читать Править Текущая версия Править исходный текст История Поиск

**Википедия**  
Свободная энциклопедия

Заглавная страница  
Рубрикация  
Указатель А—Я  
Избранные статьи  
Случайная статья  
Текущие события

Участие  
Сообщить об ошибке  
Портал сообщества  
Форум  
Свежие правки  
Новые страницы  
Справка  
Пожертвования

Печать/экспорт

Инструменты

На других языках  
Afrikaans

ru.wikipedia.org/wiki/Заглавная\_страница\_близительные\_характеристики

Find: pause Previous Next Highlight all Match case

**Гитара**

Материал из Википедии — свободной энциклопедии [править | править исходный текст]


Текущая версия страницы пока не проверялась опытными участниками и может значительно отличаться от версии, проверенной 7 августа 2013; проверки требует 1 правка.

**Гита́ра** — струнный щипковый музыкальный инструмент, один из самых распространённых в мире. Применяется в качестве аккомпанирующего или *сольного* инструмента во многих музыкальных стилях и направлениях музыки, являясь основным инструментом в таких музыкальных стилях, как блюз, кантри, фламенко, рок-музыка, иногда джаз и др. Изобретённая в XX веке электрическая гитара оказала сильное влияние на массовую культуру. Исполнитель музыки на гитаре называется *гитарист*. Человек, изготавливающий и ремонтирующий гитары, называется *гитарный мастер* или *лютье*<sup>[1]</sup>.

**Содержание** [брать]

- История гитары
  - Происхождение
  - Происхождение названия
  - Испанская гитара
  - Русская гитара
  - Классическая гитара
  - Электргитара
- Устройство гитары
  - Основные части
  - Лады
  - Струны
  - Гитарный строй
  - Усиление звука

**Гитара**



# Ejemplo

Mecahuēhuētl - Huiquipedia: in yōlōxoxouhqu cēntlamablicāxmōtli - Mozilla Firefox

File Edit View History Bookmarks Tools Help

W Mecahuēhuētl - Huiquipedia, in y...

nah.wikipedia.org/wiki/Mecahuēhuētl

Tic̄h̄huāz cē cuentah Ximomachyōmaca/Ximocalaqu

Tlaxtlapalli  Tāmap̄shuaz

Xitēpalēhuia Huiquipedia tlatequitl̄h̄iztli nahuātl̄p̄an ic̄ q̄etzta media. ic̄ achi tlanōnōtzal̄iztli, onc̄ān yāuh.

## Mecahuēhuētl

**Mecahuēhuētl** in tlatzotzonalli. Nāhuatlah̄tōcopa, ṽōcā cē mecati auh huēhuētl̄ tlatzotzonalli. Mecahuēhuētl̄ thuan tlanexmecahuēhuētl̄ tlaqueh̄ilia ip̄an achi rock tlahcuil̄iztli

### Huēyi mecahuēhuētini

- Tomo Milićević
- Matthew Bellamy
- Jmi Hendrix
- Terry Kath
- Carlos Santana
- The Edge

### Nō xiquitta

- Cuīcayōtl

Achkāuh̄itlapalli  
Cāl̄ic̄uētl̄ tocalpōl  
Axcāncayōtl  
Yāncūc̄ tlapatl̄iztli  
Zāzōzānilli  
Tēpalēhuil̄iztli  
Tēyocat̄l̄iztli

In tlein motequit̄ilia  
In tlein q̄utzonhuia nic̄ān  
Tlapatl̄iztli  
Tzonhuil̄p̄an  
Nōnc̄uah̄q̄utz̄i  
āmati  
Tepōtlah̄cuil̄ōlli  
Moch̄ipa  
Tzonhuil̄iztli  
Información de la página  
Item en Wikidata  
Tlah̄toa in̄in  
Tlah̄cuil̄ōtech̄copa

Occ̄qūint̄in  
tlah̄tōcopa

Find:      Match case

# ¿Cómo obtengo un término en otro idioma?

- Busco en: “es.wikipedia.org”



# ¿Cómo obtengo un término en otro idioma?

- Busco en: “es.wikipedia.org”
- Debo conocer la clave de la lengua meta (en, ru, nah, ...)

# ¿Cómo obtengo un término en otro idioma?

- Busco en: “es.wikipedia.org”
- Debo conocer la clave de la lengua meta (en, ru, nah, ...)
- Busco el término en esa lengua

# ¿Cómo obtengo un término en otro idioma?

- Busco en: “es.wikipedia.org”
- Debo conocer la clave de la lengua meta (en, ru, nah, ...)
- Busco el término en esa lengua
- Lo imprimo en pantalla

# ¿Qué es WATT?

- Herramienta libre y gratuita

# ¿Qué es WATT?

- Herramienta libre y gratuita
- Desarrollada en python y en java

# ¿Qué es WATT?

- Herramienta libre y gratuita
- Desarrollada en python y en java
- Simulaba ser un humano

# ¿Qué es WATT?

- Herramienta libre y gratuita
- Desarrollada en python y en java
- Simulaba ser un humano
- Ahora usa la API (interfaz de programación de aplicaciones)

# Estado actual

- Genera la URI correcta



# Estado actual

- Genera la URI correcta
- Obtiene los títulos en diferentes idiomas

# Estado actual

- Genera la URI correcta
- Obtiene los títulos en diferentes idiomas
- Reconoce el idioma de entrada (falta evaluación)

# Estado actual

- Genera la URI correcta
- Obtiene los títulos en diferentes idiomas
- Reconoce el idioma de entrada (falta evaluación)
- Obtiene definiciones en otras lenguas

# Otras cosas que se podrían hacer

- Obtener la imagen de la infobox

# Otras cosas que se podrían hacer

- Obtener la imagen de la infobox
- Utilizar la transcripción fonética (cuando exista)

# Otras cosas que se podrían hacer

- Obtener la imagen de la infobox
- Utilizar la transcripción fonética (cuando exista)
- Hacerlo aplicación móvil (Android, iOS, Windows. . .)

# Otras cosas que se podrían hacer

- Obtener la imagen de la infobox
- Utilizar la transcripción fonética (cuando exista)
- Hacerlo aplicación móvil (Android, iOS, Windows...)
- Se aceptan sugerencias...

# Resumen

- El contacto de lenguas ocurre con mayor frecuencia



# Resumen

- El contacto de lenguas ocurre con mayor frecuencia
- Existen recursos libres que pueden explotarse

# Resumen

- El contacto de lenguas ocurre con mayor frecuencia
- Existen recursos libres que pueden explotarse
- La explotación automática de esos recursos puede facilitar el trabajo de los terminólogos

# ¿Dudas?

