

Ciclo “Las aplicaciones de la fonética”

Departamento de Filología Española, Universidad Autónoma de Madrid
24 de abril de 2007

Las tecnologías del habla

Joaquim Llisterri

Grup de Fonètica, Departament de Filologia
Espanyola, Universitat Autònoma de Barcelona

Joaquim.Llisterri@uab.cat

<http://liceu.uab.cat/~joaquim>



Universitat
Autònoma
de Barcelona

Joaquim Llisterri
Grup de Fonètica, Departament de Filologia Espanyola

Ciclo “Las aplicaciones de la fonética”

Departamento de Filología Española, Universidad Autónoma de Madrid
24 de abril de 2007

Materiales en Internet

[http://liceu.uab.cat/
~joaquim/
speech_technology/
UAM_07/UAM_07.html](http://liceu.uab.cat/~joaquim/speech_technology/UAM_07/UAM_07.html)



Universitat
Autònoma
de Barcelona

Joaquim Llisterri
Grup de Fonètica, Departament de Filologia Espanyola

Las tecnologías del habla

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ La comunicación mediatizada por ordenador



Universitat
Autònoma
de Barcelona

Joaquim Llisterri
Grup de Fonètica, Departament de Filologia Espanyola

Las tecnologías del habla

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ La comunicación mediatizada por ordenador

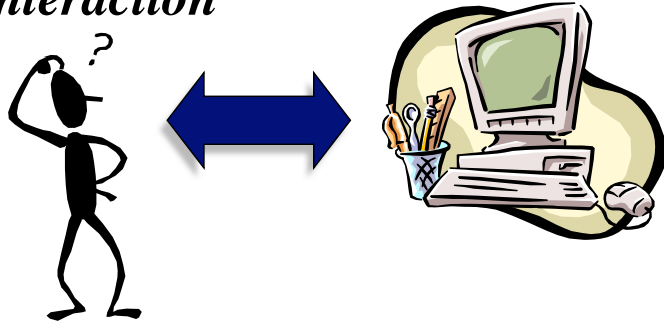


Universitat
Autònoma
de Barcelona

Joaquim Llisterri
Grup de Fonètica, Departament de Filologia Espanyola

Objetivos de un sistema de diálogo

HCI, *Human - Computer Interaction*



Las tecnologías del habla

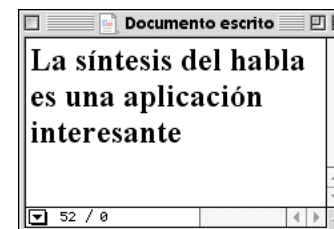
- **Objetivo**
 - **Facilitar la comunicación oral entre personas y sistemas informáticos**

Tecnologías del habla

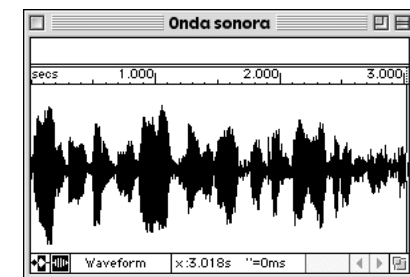
- **Salida vocal**
 - ▶ **Síntesis del habla**
- **Entrada vocal**
 - ▶ **Reconocimiento del habla**
- **Interacción vocal**
 - ▶ **Sistemas de diálogo**

Las tecnologías del habla La síntesis del habla

Del texto ...

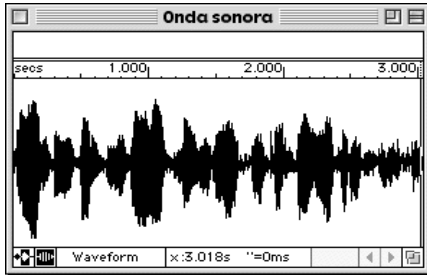


... al habla 



Las tecnologías del habla El reconocimiento del habla

Del habla ... 



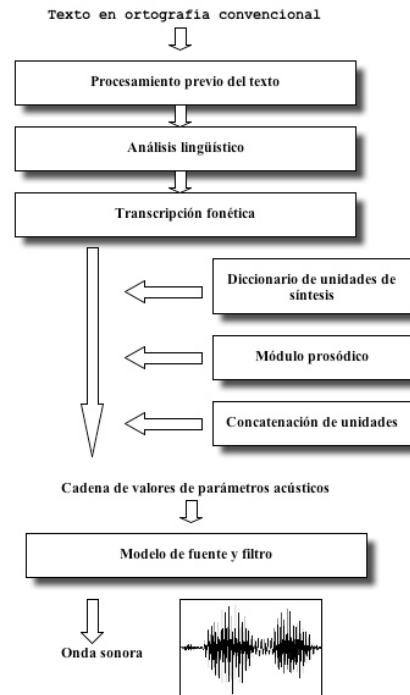
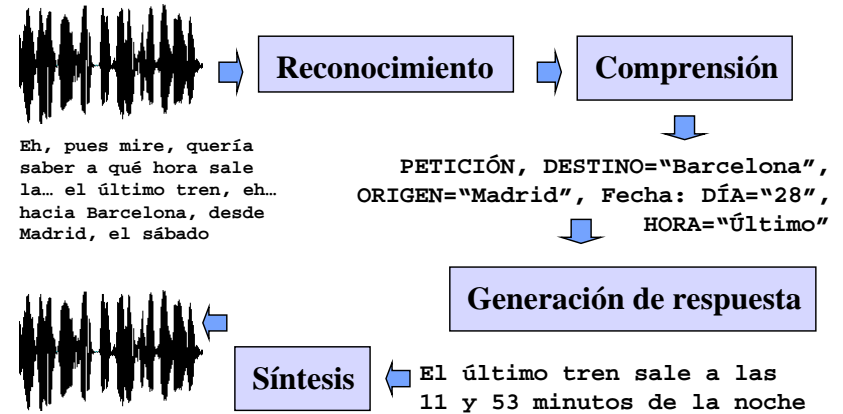
... al texto

Documento escrito
**La síntesis del habla
es una aplicación
interesante**

Las tecnologías del habla en español

- ✓ Las tecnologías del habla
- ✓ **La conversión de texto en habla**
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ La comunicación mediatizada por ordenador

Las tecnologías del habla Los sistemas de diálogo



Módulos de tratamiento lingüístico de un conversor de texto en habla

Unidades de síntesis

Difonema

- Segmento acústico que incluye la transición entre dos fonemas consecutivos, formado por la parte estacionaria del primero, la transición del primero al segundo y la parte estacionaria del segundo

/kasa/ - /k/ + /ka/ + /as/ + /sa/ + /a/

Síntesis basada en corpus

Selección de unidades

- El corpus de síntesis constituye el inventario de unidades
- Se selecciona la cadena fonética más larga que corresponde al texto de entrada del sistema de síntesis
- Se minimiza el número de concatenaciones y se busca la adecuación prosódica
- Diseño y evaluación de algoritmos de selección

Síntesis basada en corpus

- *Corpus-based speech synthesis*
- Sistemas de síntesis basados en la selección de unidades (*unit selection*) de longitud variable (*non-uniform*)
- Las unidades de síntesis están recogidas en un corpus y se concatenan una vez realizada la selección

Síntesis basada en corpus

Diseño del corpus

- Definición del inventario de alófonos y de sus combinaciones
- Cobertura de los alófonos y de sus combinaciones (difonemas)
- Cobertura de los fenómenos prosódicos

LLISTERRI, J.- MACHUCA, M.- MADRIGAL, N.- MANCINI, F.- MASSIMINO, P.- de la MOTA, C.- RIERA, M.- RÍOS, A. (2004)
 "Aspectos lingüísticos en el diseño de un convertor de texto en habla en castellano y en catalán: El sistema LoquendoTTS®", in *VI Congreso de Lingüística General*. Universidade de Santiago de Compostela, Santiago, 3-7 de mayo de 2004. Universidade de Santiago de Compostela, Facultade de Filoloxía, Área de Lingüística Xeral, pp. 521-522.
http://liceu.uab.es/~joaquim/publicacions/Llisterrí_et_al_04_Conversor_Texto_Habla_Castellano_Catalan_Loquendo.pdf

- **Corpus inicial de 3 millones de palabras extraídas de textos periodísticos, literarios, culturales, científicos, didácticos, etc.**
- **Selección automática de secuencias que se acerquen al equilibrio fonético teniendo en cuenta el contexto de aparición de los alófonos**
- **Edición manual de 200 frases enunciativas e interrogativas para ajustar el equilibrio y la estructura sintáctica**



Universitat
Autònoma
de Barcelona

Joaquim Llisterrí
Grup de Fonètica, Departament de Filologia Espanyola

LLISTERRI, J.- MACHUCA, M.- MADRIGAL, N.- MANCINI, F.- MASSIMINO, P.- de la MOTA, C.- RIERA, M.- RÍOS, A. (2004)
 "Aspectos lingüísticos en el diseño de un convertor de texto en habla en castellano y en catalán: El sistema LoquendoTTS®", in *VI Congreso de Lingüística General*. Universidade de Santiago de Compostela, Santiago, 3-7 de mayo de 2004. Universidade de Santiago de Compostela, Facultade de Filoloxía, Área de Lingüística Xeral, pp. 521-522.
http://liceu.uab.es/~joaquim/publicacions/Llisterrí_et_al_04_Conversor_Texto_Habla_Castellano_Catalan_Loquendo.pdf

- **Selección del locutor en función de la variedad lingüística, la elocución y la plasticidad: locutores profesionales**
- **Control lingüístico de la grabación de las 200 frases**
- **Correspondencia entre la realización del locutor y la transcripción fonética generada por el sistema**



Universitat
Autònoma
de Barcelona

Joaquim Llisterrí
Grup de Fonètica, Departament de Filologia Espanyola

LLISTERRI, J.- MACHUCA, M.- MADRIGAL, N.- MANCINI, F.- MASSIMINO, P.- de la MOTA, C.- RIERA, M.- RÍOS, A. (2004)
 "Aspectos lingüísticos en el diseño de un convertor de texto en habla en castellano y en catalán: El sistema LoquendoTTS®", in *VI Congreso de Lingüística General*. Universidade de Santiago de Compostela, Santiago, 3-7 de mayo de 2004. Universidade de Santiago de Compostela, Facultade de Filoloxía, Área de Lingüística Xeral, pp. 521-522.
http://liceu.uab.es/~joaquim/publicacions/Llisterrí_et_al_04_Conversor_Texto_Habla_Castellano_Catalan_Loquendo.pdf

- **Base de datos acústica de unos 200 Mb codificada a 16 kHz en PCM (*Pulse Code Modulation*)**
- **Reducción del tamaño mediante técnicas de codificación de la señal**



Universitat
Autònoma
de Barcelona

Joaquim Llisterrí
Grup de Fonètica, Departament de Filologia Espanyola

**GTH, Grupo de Tecnologías del Habla,
Universidad Politécnica de Madrid**
<http://www-gth.die.upm.es/research/synthesis/synth-form-concat-e.html/>

Introduzca su texto y recuerde que debe estar escrito en castellano ya que, en caso contrario, no obtendrá el resultado deseado. Puede utilizar cualquier símbolo del alfabeto castellano, siempre que utilice la codificación iso-8859 (vocales acentuadas (á,é,í,ó,ú), ñ, tanto mayúscula como minúscula, signos de puntuación especiales (¿,¡), etc.).

Este trabajo se centra en las tecnologías del habla para la lengua española tal como se han desarrollado en los últimos años en España, aunque también se hace referencia a algunos productos creados fuera de nuestras fronteras.

Say it!



Conversión de texto en habla por concatenación



Universitat
Autònoma
de Barcelona

Joaquim Llisterrí
Grup de Fonètica, Departament de Filologia Espanyola



Universitat
Autònoma
de Barcelona

Joaquim Llisterrí
Grup de Fonètica, Departament de Filologia Espanyola

**GTP, Grup de Tractament de la Parla,
Universitat Politècnica de Catalunya**
<http://gps-tsc.upc.es/veu/soft/demos/tts.php3>

1. Select a configuration

METHOD	LANGUAGE	SPEAKER	SPEECH RATE
Basic Diphone	Spanish	<input type="radio"/> female <input type="radio"/> male	Medium
	Catalan	<input type="radio"/> female <input type="radio"/> male	
Unit-Selection	Catalan	<input type="radio"/> female	AUDIO FORMAT
	Spanish	<input checked="" type="radio"/> female	

2. Type your text here OR copy one of the examples below. Then press READ NOW to listen to the result.

Este trabajo se centra en las tecnologías del habla para la lengua española tal como se han desarrollado en los últimos años en España, aunque también se hace referencia a algunos productos creados fuera de nuestras fronteras.

CLEAR READ NOW



Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

GTS, Grupo de Tratamiento de la Señal, Universidade de Vigo
<http://www.gts.tsc.uvigo.es/cotovia/>

Conversión de texto en habla basada en corpus



Conversión de texto en habla basada en difonemas



Synthesis type: Diphone-based | Language: Spanish | Voice: Sonia (female) | 3D effects: None

Speed*: 50% 75% 100% 150% 200%
Pitch*

Este trabajo se centra en las tecnologías del habla para la lengua española tal como se han desarrollado en los últimos años en España, aunque también se hace referencia a algunos productos creados fuera de nuestras fronteras.

Listen



Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Telefónica I+D, Madrid

Conversión de texto en habla basada en concatenación de unidades (1996)



Conversión de texto en habla basada en corpus (2003)



Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Verbio Technologies SL
http://www.verbio.com/webverbio2/html/demos_ttsonline.php



La demostración de síntesis on-line le permite evaluar la calidad de los locutores Verbio a través de cualquier texto que desee.

Simplemente siga los siguientes pasos:

- Escoja idioma y locutor.
- Escriba el texto a sintetizar y pulse 'Leer'.
- Podrá escuchar el resultado en un archivo de sonido que se descargará en su equipo.
- Tenga en cuenta que el proceso de descarga puede tardar en función del tamaño del texto y la velocidad de la conexión.

Idioma: Español | Locutor: Laura

Borrar Leer



Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola



Verbio Technologies SL

http://www.verbio.com/webverbio2/html/demos_news.php

Siga los siguientes pasos:

- Elija el periódico que desee consultar: El Universal (Español Mexicano)
- Escoja una sección del periódico: Minuto x Minuto
- Seleccione un locutor Verbio: Celia
- Pulse el botón de cualquier titular para ampliar la noticia. Podrá escuchar la noticia entera en un archivo de sonido que se descargará en su equipo. También puede consultar el texto completo de la noticia haciendo clic sobre el enlace puntos suspensivos (...)

- Deja incendio en club nocturno de Moscú 10 muertos (...)
- Evacúan a pasajeros de avión israelí tras detectar humo (...)
- Dudan de Tom Cruise para realizar filme sobre atentado contra Hitler (...)
- Llega Carrillo a 400 dobles en victoria de Tigres 5-1 Campeche (...)



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Actor, Loquendo

<http://actor.loquendo.com/actordemo/>



Cat, f Cat, m

ACTOR 5 DEMO

Enter a message:

Este ejemplo es el resultado del desarrollo de una voz española, llevada a cabo con una nueva técnica de síntesis del habla. Como puede verse, la calidad es buena. ¿No les parece que sí? Por cierto, el sistema también existe en otras lenguas.



Es, m, f

EsMx, f

EsCh, f

EsAr, m



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Loquendo Gilded TTS

<http://actor.loquendo.com/actordemo/>

Buenos días a todos! Soy Carmen, la voz sintética femenina de Loquendo. _Throat Tengo que contar que ultimamente me he vuelto más expresiva y sentimental. Me río si soy feliz, _Laugh_03 silbo si me parece el caso, _Whistle_01 y muchas otras cosas más. Adiós! un besote! _Smack_02



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

AT&T Labs Natural Voices

<http://www.naturalvoices.att.com/demos/>

Recorded Text-to-Speech Demo (Latin American Spanish, Female)

Step 1 SELECT ONE Output Format

wav (PC) aiff (mac) au basic

NOTE: Current selection is preset for your PC.

Step 2 Phrases

SELECT ONE of the phrases below and click SUBMIT to hear the Text-to-Speech conversion:

- "El tour de Francia de ciclismo empezó el sábado pasado, con la victoria de un corredor de un equipo francés" ("The cycling Tour de France began last Saturday with a victory for a rider from a French team")
- "Argentina vence 3-0 a Ghana y es campeón mundial juvenil" ("Argentina beats Ghana 3-0 and are junior world champions")
- "Los astrónomos han descubierto un nuevo cuerpo espacial en el Sistema Solar, más allá de la órbita de Neptuno" ("Astronomers have discovered a new planet-like object in the Solar System, further out than the orbit of Neptune")

SUBMIT

EspMx, f

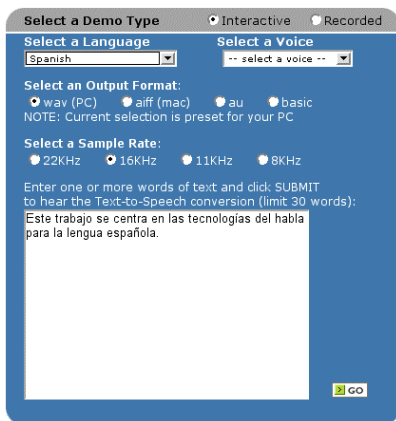


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

AT&T Labs Natural Voices

<http://www.naturalvoices.att.com/demos/>



EspMx, f

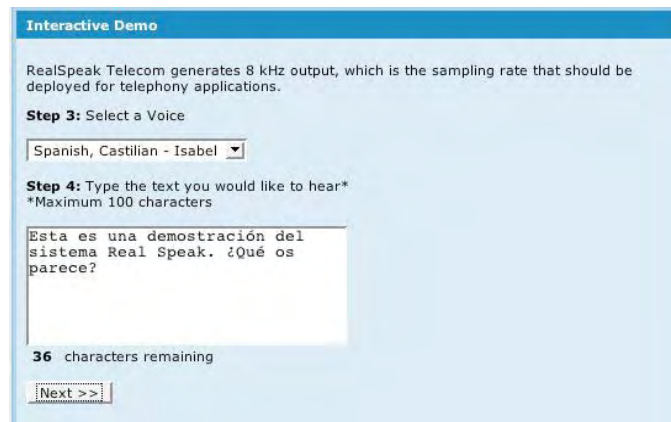


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

RealSpeak, Nuance

<http://www.nuance.com/realspeak/demo/default.asp>



Eng, f



EsMx, f



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Acapela Group

<http://demo.acapela-group.com/>

Acapela HQ TTS Interactive Demo

Select the language and voice :

Spanish Maria (Female)

Options :

Output file format : Wav Standard

Voice speed : Normal

Enter your text :

Default text Clear text

Este ejemplo muestra el resultado de la conversión de texto en habla en español. ¿Les parece interesante? ¡Yo creo que sí!

129 characters remaining

Say it!

Like what you hear ? Want more information ? Contact us ! >>>

En, f



Es, f



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Las tecnologías del habla

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ **El reconocimiento automático del habla**
- ✓ Los sistemas de diálogo
- ✓ La comunicación mediatizada por ordenador

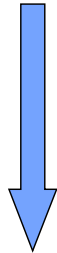


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Tipo de enunciado

- dificultad



+ dificultad

- Reconocimiento de palabras aisladas
- Reconocimiento de palabras conectadas
- Reconocimiento de habla continua



Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Número de locutores

- dificultad



+ dificultad

- Sistemas dependientes del locutor
- Sistemas independientes del locutor

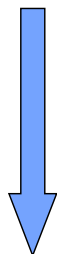


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Tamaño del vocabulario

- dificultad



+ dificultad

- Vocabulario restringido
- Vocabulario sin restricciones

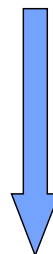


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Entorno

- dificultad



+ dificultad

- Entorno silencioso
- Entorno con ruido predecible
- Entorno con ruido aleatorio

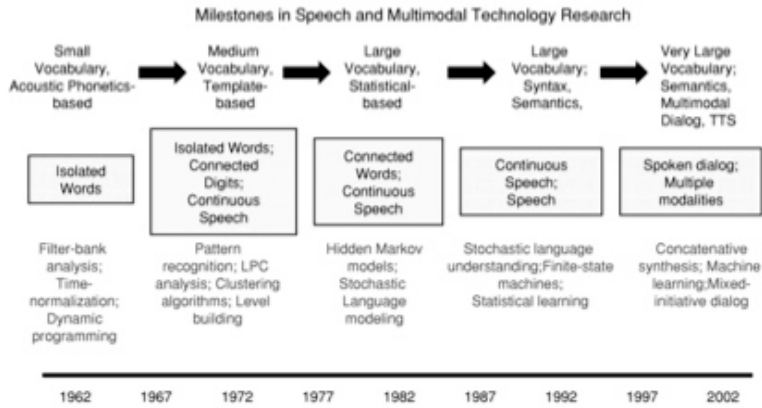


Universitat Autònoma de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Desarrollo histórico del reconocimiento del habla

JUANG, B.-H. - RABINER, L. R. (2006) "Automatic speech recognition: History", in BROWN, K. (Ed.) *Encyclopedia of Language & Linguistics*. Amsterdam: Elsevier. pp. 806-819.



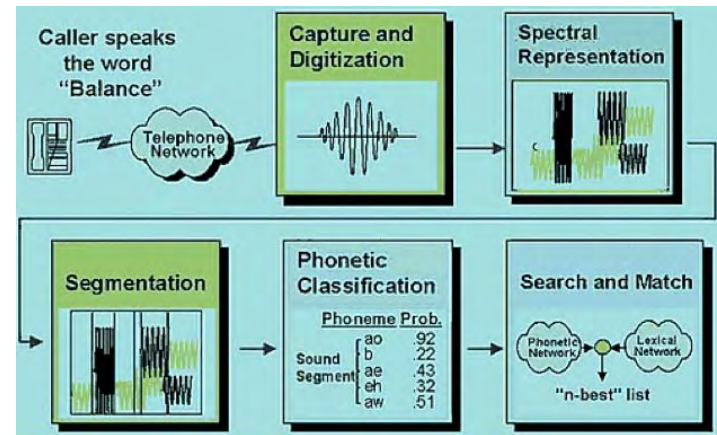
Universitat Autònoma de Barcelona

Joaquim Llisteri

Grup de Fonètica, Departament de Filologia Espanyola

Speech-Enabled Interactive Voice Response Systems. Web ProForum Tutorials, International Engineering Consortium.

http://www.iec.org/online/tutorials/speech_enabled/index.html

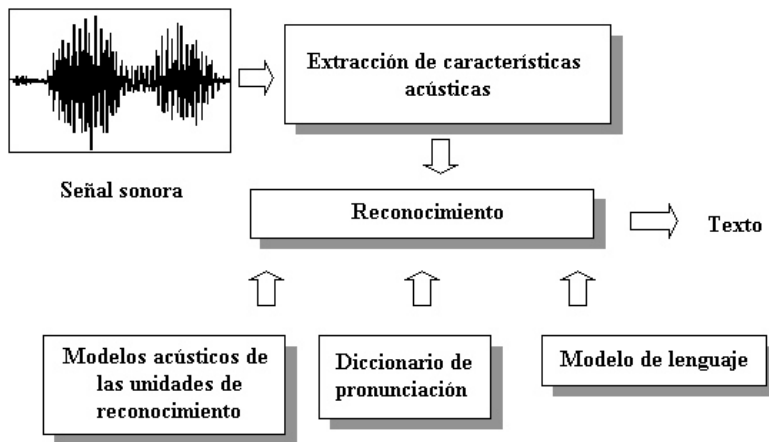


Universitat Autònoma de Barcelona

Joaquim Llisteri

Grup de Fonètica, Departament de Filologia Espanyola

Esquema básico de un sistema de reconocimiento automático del habla

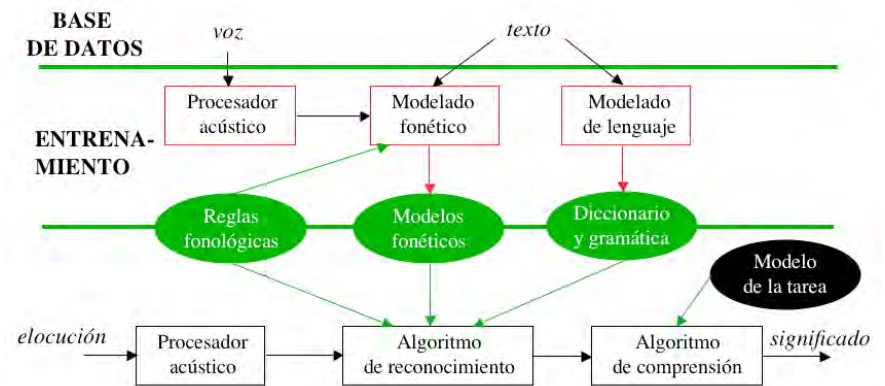


Universitat Autònoma de Barcelona

Joaquim Llisteri

Grup de Fonètica, Departament de Filologia Espanyola

Modelo de reconocimiento de habla continua Mariño (1999)



Universitat Autònoma de Barcelona

Joaquim Llisteri

Grup de Fonètica, Departament de Filologia Espanyola

Reconocimiento de palabras aisladas

- **Control del sistema operativo de un ordenador**



<http://www.youtube.com/watch?v=SKm-g7hUSag>



Universitat
Autònoma
de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Windows Vista

Primera presentación pública

<http://www.youtube.com/watch?v=G-dYiTc88Pc>



Universitat
Autònoma
de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Dictado automático

Dragon Naturally Speaking

- **Edición estándar y profesional**
- **Vocabulario legal**
- **Vocabulario médico**
- **Versión en español**



<http://www.nuance.com/naturallyspeaking/>
<http://spain.nuance.com/naturallyspeaking/>

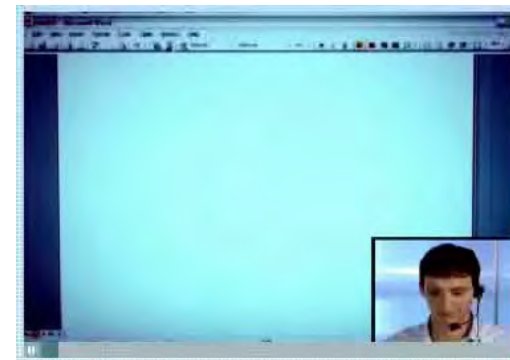


Universitat
Autònoma
de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Dragon Naturally Speaking

<http://www.nuance.com/talk/>

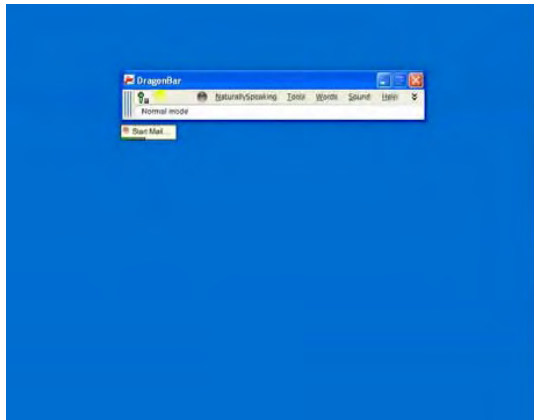


Universitat
Autònoma
de Barcelona

Joaquim Llisteri
Grup de Fonètica, Departament de Filologia Espanyola

Gestión y dictado de correos electrónicos Dragon Naturally Speaking

<http://www.nuance.com/naturallyspeaking/demos/>



Nuance Mobile Dictation “The amazing race”

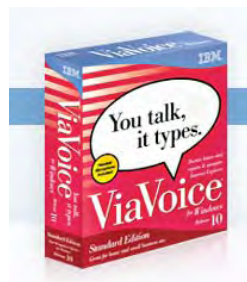


<http://www.youtube.com/watch?v=-L4Jk6GDud0>

Dictado automático

IBM Via Voice

- 300.000 palabras
- Vocabulario legal
- Vocabulario médico
- Primeras versiones en español



<http://www-306.ibm.com/software/voice/viavoice/>

<http://www.nuance.com/viavoice/>

Audio Information Retrieval Recuperación de información sonora

- TRANSCRIGAL, Grupo de Tecnologías de la Señal, Universidade de Vigo

<http://www.webapps.gts.tsc.uvigo.es/Transcrigal/>

- Transcripción automática y recuperación de contenido en telediarios en castellano y en gallego

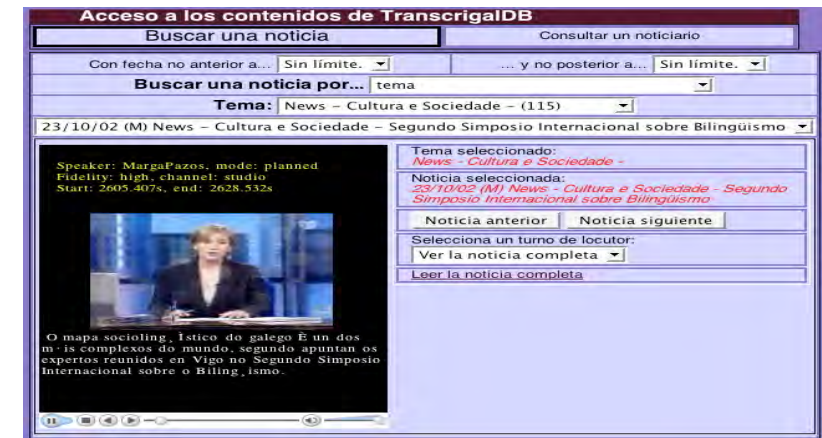
TRANSCRIGAL

<http://www.webapps.gts.tsc.uvigo.es/Transcrigal/>



TRANSCRIGAL

<http://www.webapps.gts.tsc.uvigo.es/Transcrigal/>



Audio Information Retrieval

Recuperación de información sonora

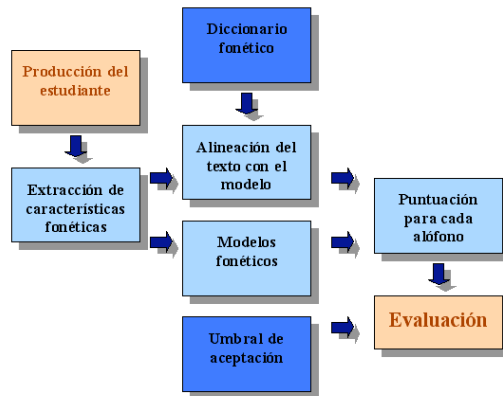
- Nuance Dragon Audio Mining
<http://www.nuance.com/audiomining/>
- Producto comercial para la búsqueda de información mediante palabras clave o frases en bases de datos sonoras
- Versión en español

Nuance Dragon Audio Mining

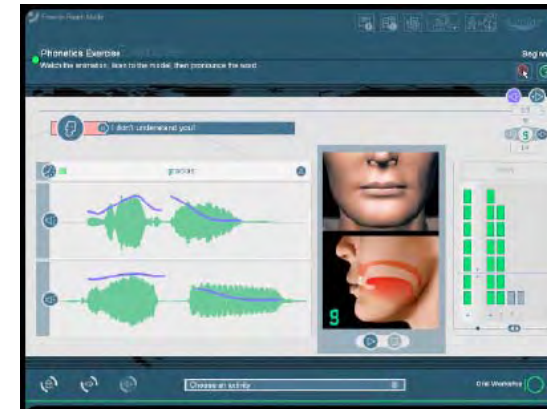
<http://www.nuance.com/audiomining/>



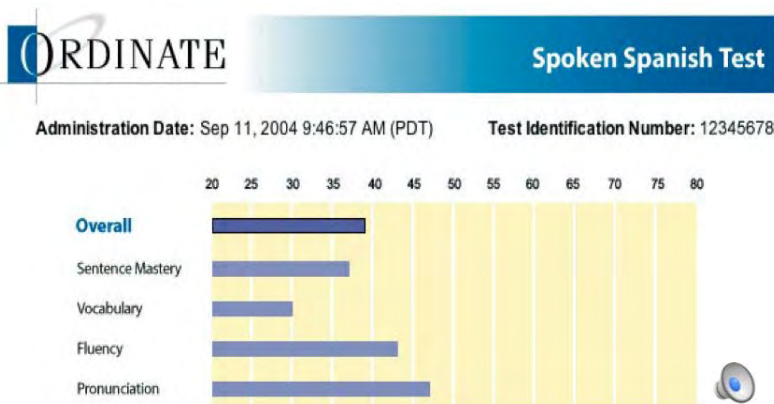
Reconocimiento automático del habla y enseñanza de la pronunciación asistida por ordenador



LAFFORD, B.A. (2004) "Review of *Tell Me More Spanish*",
Language Learning and Technology 8, 3: 21-34.
<http://llt.msu.edu/vol8num3/pdf/review1.pdf>



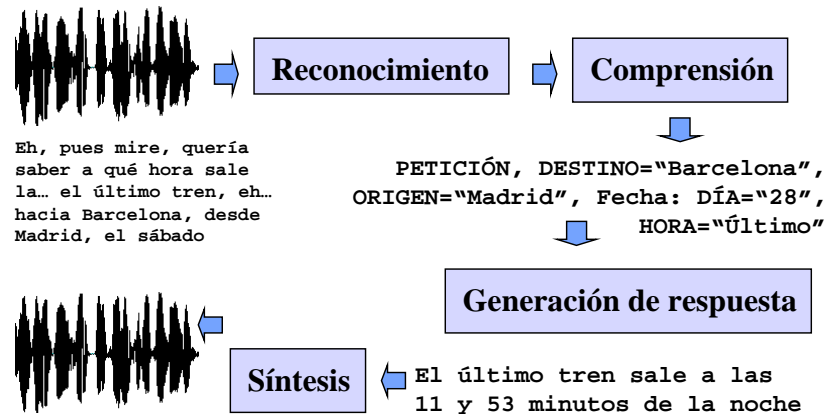
Ordinate Corporation, Menlo Park, California
<http://www.ordinate.com/>



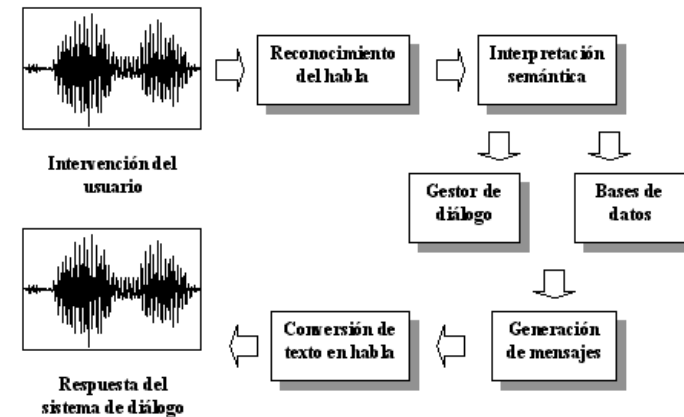
Las tecnologías del habla en español

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ La situación de las tecnologías del habla en España

Los sistemas de diálogo



Módulos de un sistema de diálogo



Módulos de un sistema de diálogo

Módulo de reconocimiento del habla

- Convierte la señal sonora de entrada - continua- en una representación escrita - discreta - de las palabras reconocidas

► Tecnologías del habla

- Reconocimiento automático del habla (ASR, *Automatic Speech Recognition*)

Módulos de un sistema de diálogo

Módulo de interpretación semántica

- Determina el “significado” (contenido) de la secuencia de palabras reconocida
- **Procesamiento del lenguaje natural**
 - Comprensión del lenguaje (NLU, *Natural Language Understanding*)

Módulos de un sistema de diálogo

Comprensión del contenido

◆ Comprensión del lenguaje natural (NLU)

Eh, pues mire, quería saber a qué hora sale la... el último tren, eh... hacia Barcelona, desde Madrid, el sábado



PETICIÓN,
DESTINO="Barcelona",
ORIGEN="Madrid", FECHA:
DÍA="28", HORA="Último"

Módulos de un sistema de diálogo

Módulo de gestión del diálogo

- Gestiona la interacción entre la persona y el sistema para la realización de la tarea deseada
- Módulo central de control en un sistema de diálogo

Módulos de un sistema de diálogo

Módulo de generación de lenguaje

- Crea un enunciado bien formado a partir de la representación interna proporcionada por el sistema
- ◆ Procesamiento del lenguaje natural
 - Generación del lenguaje (NLG, *Natural Language Generation*)

Módulos de un sistema de diálogo

Creación de respuestas

◆ Generación de lenguaje natural (NLG)

PETICIÓN, DESTINO="Barcelona",
ORIGEN="Madrid", Fecha: DÍA="28",
HORA="Último"



El último tren sale a las 11 y 53 minutos de la noche

Módulos de un sistema de diálogo


Módulo de conversión de texto en habla

- Transforma en una señal sonora continua- el texto creado por el módulo de generación de respuestas -discreto-

➤ Tecnologías del habla

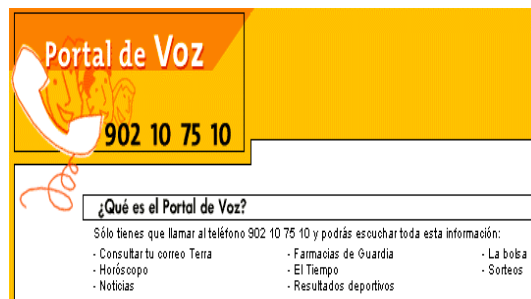
- Conversión de texto en habla (TTS, *Tex-to-Speech Synthesis*)

BASURDE (1998-2001)
Sistema de diálogo oral en dominios restringidos
<http://gps-tsc.upc.es/veu/basurde/Home.htm>

- Reconocimiento
 - “Para conseguir una tasa de reconocimiento aceptable es deseable que la frase se pronuncie evitando ruidos y a una velocidad de locución normal”
- Comprensión
 - Información telefónica de horarios y precios de trenes regionales y de grandes líneas
- Salida vocal 

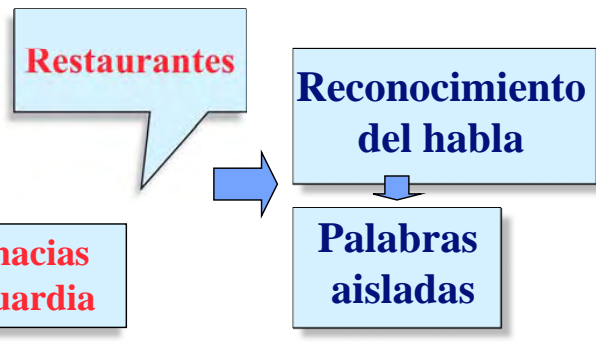
Acceso al ordenador mediante el teléfono

• Portales de voz




Portal de voz

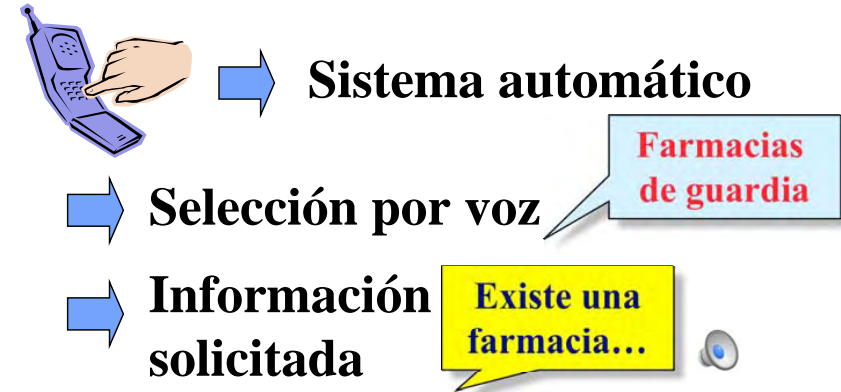
Selección por voz



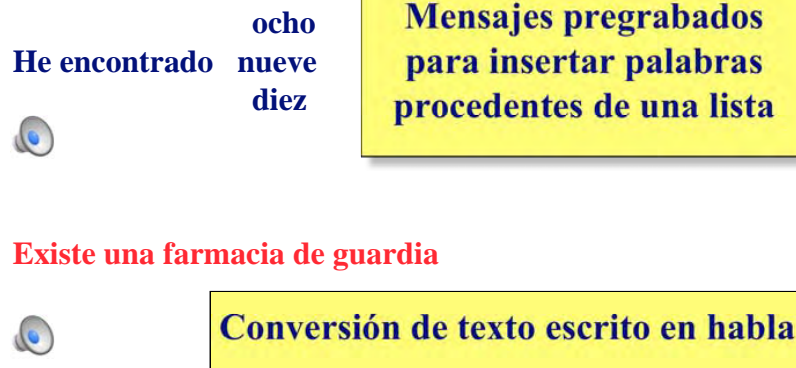
Portal de voz



Portal de voz



Portal de voz



Venta de entradas

CineEntradas

Ydilo Advanced Voice Solutions



<http://www.ydilo.com/esp/caseStudies.html>

Banca Telefónica

BpT, Banca por Teléfono Natural Natural Vox

- **STI:** Bienvenido a Banca Telefónica, buenas tardes.
- **Usu.:** Buenas tardes.
- **STI:** Por favor, diga o teclee los dígitos de su DNI.
- **Usu.:** 2-2-3-1-9-4-8-7.
- **STI:** Ahora los cuatro dígitos de su clave de acceso.
- **Usu.:** 2-4-4-7.
- **STI:** De acuerdo, ¿qué operación desea?
- **Usu.:** Quería comprar unas acciones.
- **STI:** Un momento, por favor, ¿de los contratos de los que es titular?
- **Usu.:** Si, si.

http://www.natvox.es/demos_bpt.html



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

Gestión del correo electrónico

TelCorreo Universidade de Vigo Universidade de Compostela



Sistema de diálogo para acceso telefónico a servicios telemáticos

<http://www.gts.tsc.uvigo.es/telcorreo/>



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

Cita previa

http://www.verbio.com/webverbio/html/demos_basic.php



demos basicas

Pruebe nuestra [tecnología de reconocimiento del habla](#) y sorpréndase con esta demo básica en distintos idiomas.

Esta demostración le permitirá comprobar la fiabilidad del sistema de reconocimiento, mediante un ejemplo sencillo de "formulario tipo" que deberá rellenar mediante la voz.

Simplemente siga los siguientes pasos:

- Llame al **93 4097304**
- Elija el idioma y el tipo de formulario a rellenar (formulario **personal** o **cita previa**)
- El sistema le asignará un **número de sesión** para poder comprobar on-line el resultado de la demostración (si introduce inmediatamente después de oírlo el código de sesión asignado en esta misma página podrá ver la demo en tiempo real)
- Vaya introduciendo los datos que se le solicitan siguiendo el orden del formulario.

Nota: Los datos proporcionados, como el NIF, el número de teléfono, o el código postal, deben ser válidos en el territorio español y los podrá introducir de forma natural con agrupaciones de cifras y no sólo dígito a dígito.



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

Las tecnologías del habla

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ **La comunicación mediatizada por ordenador**

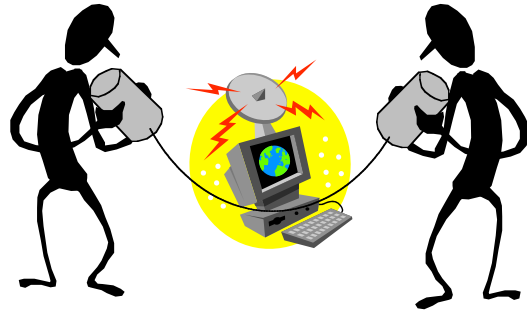


Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

La comunicación persona-persona mediatizada por una máquina

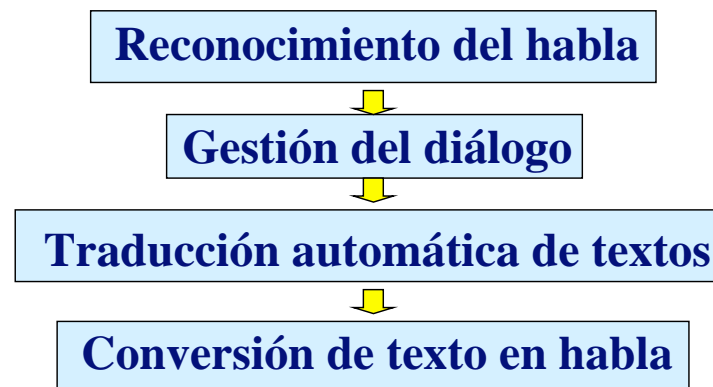
CMM, *Computer-Mediated Communication*



La comunicación persona-persona mediatizada por una máquina

- Traducción entre modalidades
 - Escrito - Escrito
 - ▶ Traducción automática de textos
(MT, *Machine Translation*)
 - Oral - Oral
 - ▶ Traducción automática del habla
(SLT, *Spoken Language Translation*)

Traducción automática del habla *Spoken Language Translation*



Traducción automática del habla

<http://www.is.cs.cmu.edu/papers/speech/1996/COMPUTER-alex.ps.gz>

“...sí sí el viernes diecinueve puedo sí porque sabes me voy de viaje d hoy la verdad así es que este mes es muy viajero me voy el día seis de viaje y estoy hasta el doce así que el día diecinueve me viene muy bien francamente...”
yes yes on friday nineteen can yes because know I go me of trip D today the truth such is that this month is very traveler I go me the day six of trip and I am until the twelve as soon as the day nineteen comes me very well outspokenly

Verbmobil

<http://verbmobil.dfki.de/overview-us.html>

- Sistema móvil de traducción de diálogos espontáneos (1993-2000)
- Dominios restringidos: citas, preparación de viajes, reservas de hotel
- Lenguas: alemán, inglés y japonés



WAHLSTER, W. (2000) "Mobile Speech-to-Speech Translation of Spontaneous Dialogs: An Overview of the Final Verbmobil System", in WAHLSTER, W. (Ed.) *Verbmobil: Foundations of Speech-to-Speech Translation*. Heidelberg - New York: Springer Verlag (Artificial Intelligence). <http://verbmobil.dfki.de/ww.html>

Verbmobil

<http://verbmobil.dfki.de/overview-us.html>



La comunicación persona-persona mediatizada por una máquina

- Traducción entre modalidades
 - Escrito - Oral
 - ↳ Conversión de texto en habla (TTS, *Text-to-Speech Synthesis*)
 - Oral - Escrito
 - ↳ Reconocimiento automático del habla (ASR, *Automatic Speech Recognition*)

La comunicación persona-persona mediatizada por una máquina

- Traducció entre modalitats
 - Lengua oral - Movimientos labiales
 - ▶ Reconocimiento del habla
 - ▶ Generación de movimientos labiales a partir del reconocimiento



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

Teleface

Department of Speech, Music and Hearing, KTH, 2002
<http://www.speech.kth.se/teleface/>



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

La comunicación persona-persona mediatizada por una máquina

- Traducción entre modalidades
 - Lengua oral - Lengua de señas
 - ▶ Reconocimiento del habla o de la imagen
 - ▶ Traducción de/a lengua de señas
 - ▶ Conversión de texto en habla o de texto a señas



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

eSign, Essential Sign Language Information on Government Networks

<http://www.sign-lang.uni-hamburg.de/esign/>



Generación automática de traducción a lengua de señas para programas de televisión



Universitat
Autònoma
de Barcelona

Joaquim Llisterra
Grup de Fonètica, Departament de Filologia Espanyola

iCommunicator

<http://www.myicomunicator.com/>



- ▶ Reconocimiento del habla
- ▶ Conversión de texto en habla
- ▶ Traducción de lengua oral a lengua de señas

La comunicación persona-persona mediatizada por una máquina

- ▶ Traducción entre modalidades
- Acceso a la información en otras lenguas
- Acceso a la información mediante la modalidad más adecuada a las necesidades específicas de cada usuario

Otras aplicaciones de las tecnologías del habla...



<http://www.youtube.com/watch?v=YhWmzTjScfA>

Las tecnologías del habla

- ✓ Las tecnologías del habla
- ✓ La conversión de texto en habla
- ✓ El reconocimiento automático del habla
- ✓ Los sistemas de diálogo
- ✓ La comunicación mediatizada por ordenador

Ciclo “Las aplicaciones de la fonética”
Departamento de Filología Española, Universidad Autónoma de Madrid
24 de abril de 2007

Materiales en Internet

[http://liceu.uab.cat/
~joaquim/
speech_technology/
UAM_07/UAM_07.html](http://liceu.uab.cat/~joaquim/speech_technology/UAM_07/UAM_07.html)