

# Projecte **TECNOPARLA**. Tecnologies de la parla en català

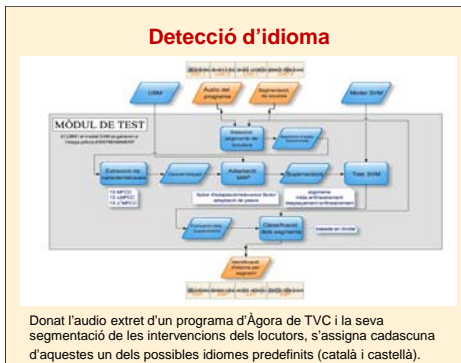
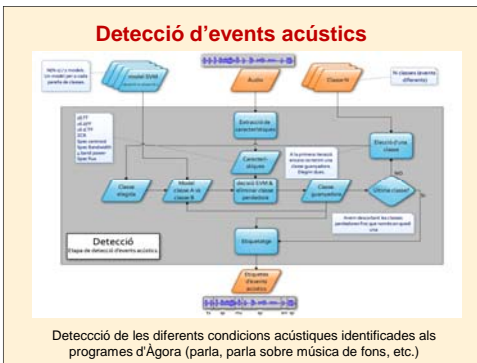
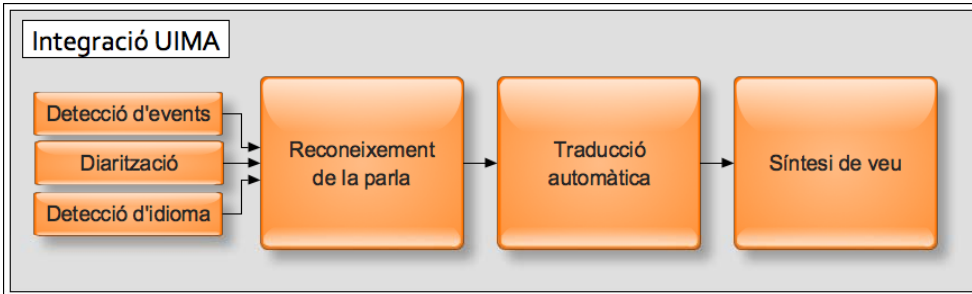
Contacte: J. Adrián Rodríguez Fonollosa, [adrian@gps.tsc.upc.edu](mailto:adrian@gps.tsc.upc.edu)

Grup de Tractament de la Parla – Centre de Tecnologies i Aplicacions del Llenguatge i la Parla (TALP)

Departament de Teoria del Senyal i Comunicacions – Universitat Politècnica de Catalunya

## Resum

En el marc de l'objectiu principal del projecte, el desenvolupament de tecnologies de la parla en català, s'ha desenvolupat un sistema automàtic de transcripció i traducció de notícies i debats televisius. Aquest sistema serveix com a demostrador de les possibilitats que ofereixen els sistemes desenvolupats de reconeixement de la parla, detecció d'events acústics, detecció d'idioma, diarització de locutors, traducció automàtica estadística i síntesi de veu. També ofereix una plataforma oberta i basada en estàndards internacionals d'integració de totes aquestes tecnologies.



La diarització és la tasca de reconèixer la identitat de les persones presents en una gravació. Moltes vegades per referència a aquest tipus de tecnologia es fa ús de la frase: "¿ qui parlo? ¿ qui interessa? Aquesta frase és molt explicativa... resumeix l'objectiu principal: segmentar en temps la intervenció de locutors i agrupar aquestes intervencions distingint les de locutors diferents. Si, tenim tota la realitat però a més la diarització pot recórrer-se en altres tecnologies... Com la segmentació d'events acústics per detectar silencis, sorolls, riuers, la detecció de les condicions acústiques o també del gènere del locutor. I tant! Amb aquesta informació el diaritzador pot adaptar-se millor a les diferents situacions que hi ha en una conversa. El riu: de qui parlo? ¿ M'engorja la la tauleta..."



### Audíio

- Programa *Àgora* emès a TVC
  - 34 debats altament moderats
  - gran variació de temes i personatges convidats
  - no inclou els anuncis publicitaris
- Temps total: 43h 25min
- Divisió: entrenament-desenvolupament-test (comuna per a tots els mòduls, d'acord a les seves necessitats)

### Bases de dades d'audíio i de text

### Text

- Corpus bilingüe CATALÀ-CASTELLÀ de *El Periòdic*

	CATALÀ	CASTELLÀ
Oracions		6.5 M
Paraules	139.0 M	139.2 M

### Sistema de traducció estadística Ngram-based

$\arg\max_{s'_i} P(s'_i | s_i)$  aproximació

$\arg\max_{s'_i} \sum_{m=1}^M \lambda_m h_m(s'_i, s_i)$  combinació de funcions característiques

Model de traducció

$$h_m(s'_i, s_i) = \prod_{j=1}^k p((s, t) | (s', t)_{-N+1}, \dots, (s, t)_{-1})$$

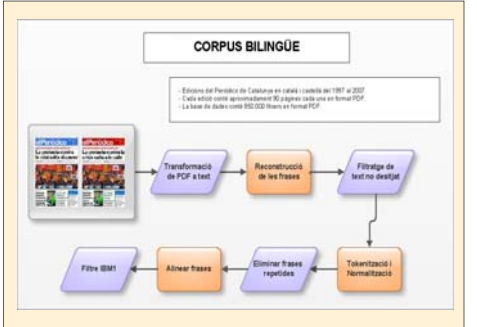
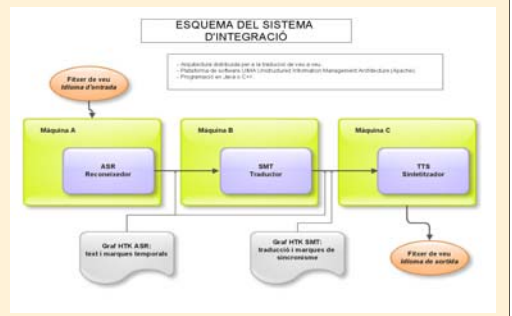
Model de llenguatge basat en Part-of-Speech

Bonificació de paraules

### Síntesi de la veu

Donat el text traduït i el seu alineament amb l'audio d'entrada es genera veu traduïda respectant cert sincronisme. S'està treballant en una nova funcionalitat per adaptar el timbre de la veu sintètica perquè s'assembli a la veu del parlant del programa font.

El sistema de síntesi de veu obté una classificació de 4 en una escala de 1 a 5. (La veu natural rep 4.6). La conversió de veu redueix la qualitat a 3.



### Demo

Original Language Catalan

entraria en aquesta casuística és a dir seria un cas en què combina es podria interpretar que hi ha una voluntat de pau de canviar la relació de poder mentre el tribunal constitucional

Catalunya entraria en esa casuística, es decir, sería un caso en el que combina podría a interpretar se que existe una voluntad de paz de cambiar la relación de poder mientras el tribunal constitucional.