



## Treball final de grau

**Estudi:** Grau en Enginyeria Biomèdica

**Títol:** Sistema de control per veu per a la manipulació d'un instrument quirúrgic auxiliar en cirurgia mínimament invasiva

**Document:** Resum

**Alumne:** Gerard Torras Canela

**Tutor:** Xavier Cufí Solé

**Departament:** Arquitectura i Tecnologia de Computadors

**Àrea de coneixement:** Arquitectura i Tecnologia de Computadors

**Convocatòria (mes/any):** Juny 2025

## 1. INTRODUCCIÓ

La cirurgia mínimament invasiva (CMI) ha transformat el camp de la medicina quirúrgica gràcies a la reducció del trauma al pacient, la disminució del dolor postoperatori i una recuperació més ràpida. No obstant això, aquesta tècnica requereix una gran precisió i estabilitat en la manipulació dels instruments, especialment de la càmera laparoscòpica, que sol ser controlada per un assistent. Aquesta dependència pot generar dificultats d'ergonomia, retards en el procediment i falta de fluïdesa durant la intervenció.

Amb l'avenç dels sistemes robòtics quirúrgics i la integració de tecnologies emergents com el reconeixement de veu, s'obren noves possibilitats per millorar l'autonomia i l'eficiència del cirurgià. En aquest context, sorgeix la necessitat de desenvolupar un sistema que permeti controlar de manera directa i intuïtiva la càmera sense requerir assistents addicionals, utilitzant només ordres verbals en llengua natural.

Aquest treball s'emmarca en aquest repte tecnològic i assistencial, proposant una solució experimental basada en tecnologia accessible i modular, adaptada a entorns de simulació i amb projecció cap a aplicacions mèdiques reals.

## 2. OBJECTIU DEL PROJECTE

L'objectiu principal d'aquest treball és desenvolupar un sistema robòtic capaç de controlar una càmera endoscòpica mitjançant ordres de veu, orientat a la seva aplicació en cirurgia mínimament invasiva (CMI). El sistema ha de permetre que el mateix cirurgià pugui ajustar la posició i orientació de la càmera de forma directa i verbal, sense necessitat d'interacció física ni assistència externa, millorant així l'autonomia i eficiència de l'acte quirúrgic.

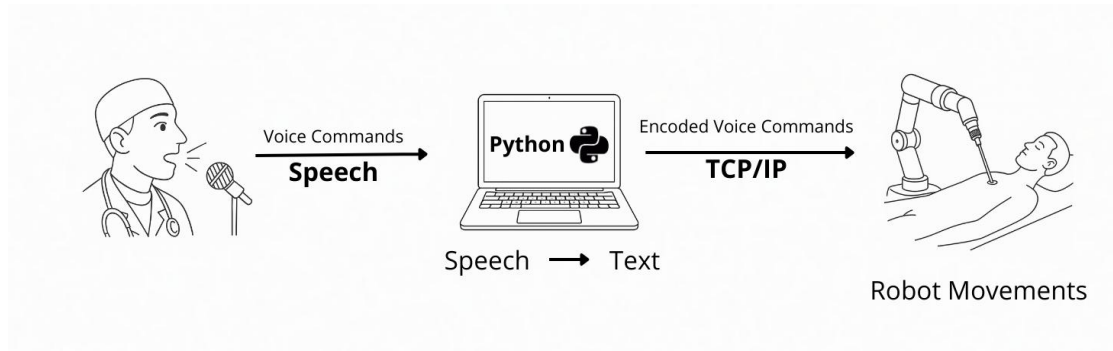
Aquest control vocal es materialitza a través de la integració de diferents components:

- Un sistema de reconeixement de veu en català, desenvolupat en Python, que interpreta comandes verbals predefinides.
- Una connexió de xarxa TCP/IP que comunica el mòdul de veu amb el robot industrial.
- Un braç robòtic Stäubli TX60, programat en llenguatge VAL3, que simula el moviment de la càmera laparoscòpica sobre un punt fixat (el trocar).

A nivell funcional, el robot ha de poder realitzar moviments de:

- Translació vertical (endinsar o retirar la càmera).
- Pivotatge lateral (rotació al voltant del trocar mantenint-lo com a centre de moviment).

L'arquitectura completa del sistema es pot observar a la Figura 1, que mostra el flux funcional des de la comanda vocal fins al moviment físic executat pel robot.



**Figura 1.** Esquema funcional del sistema de control per veu per a la manipulació d'un robot quirúrgic auxiliar.

(Font: adaptada de Figura 2.3 de la memòria)

Aquest enfocament busca oferir una alternativa intuïtiva, assequible i extensible als sistemes quirúrgics comercials, especialment pensada per a entorns de simulació, entrenament o recerca, amb la possibilitat d'adaptar-se en un futur a condicions d'ús clínic real.

### 3. DESENVOLUPAMENT TÈCNIC

El sistema desenvolupat es basa en la integració de tecnologies de control robòtic, reconeixement de veu i comunicació per xarxa, aplicades a un entorn de simulació quirúrgica. El projecte s'ha estructurat en diversos mòduls interconnectats, cadascun amb una funció específica dins del flux global del sistema.

#### 3.1. Arquitectura general

El sistema està compost per tres blocs principals:

- Entrada vocal: l'usuari dicta comandes en català, com ara "entra", "surt", "esquerra" o "dreta".
- Processament en Python: un script captura la veu via micròfon, l'envia a l'API de Google Speech-to-Text i interpreta la comanda mitjançant un diccionari personalitzat.
- Control robòtic via socket: la comanda codificada (ex. 0, 1, 2, 3) es transmet mitjançant TCP/IP al robot Stäubli TX60, que l'interpreta i executa el moviment corresponent en temps real.

Aquesta arquitectura modular permet una resposta ràpida i flexible, així com una fàcil adaptació a futures ampliacions del sistema.

### 3.2. Reconeixement de veu

El mòdul de reconeixement de veu s'ha implementat en Python 3 mitjançant la llibreria *speech\_recognition*. El sistema funciona en bucle continu, capturant ordres i comparant-les amb un conjunt tancat de comandes conegudes. Les ordres reconegudes es converteixen a valors numèrics i es transmeten de forma immediata al robot.

Per garantir la simplicitat i robustesa del sistema, s'ha optat per un conjunt reduït de comandes, tot i que el sistema és escalable.

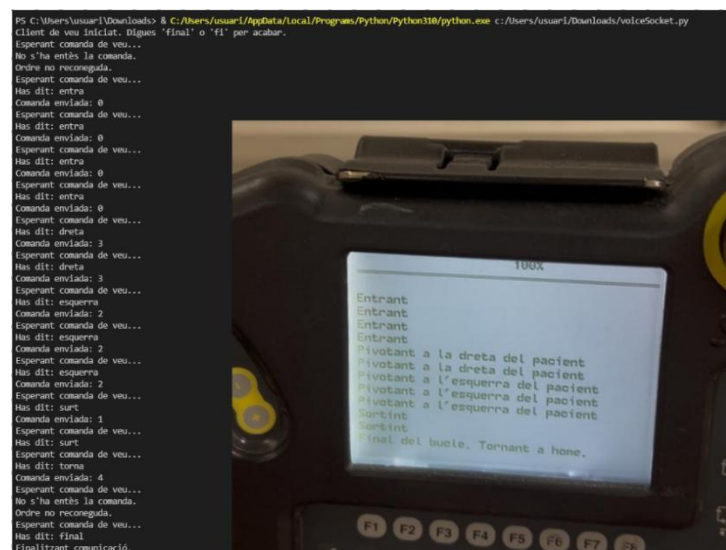
### 3.3. Comunicació per sockets TCP/IP

La comunicació entre el client Python i el robot s'ha implementat mitjançant sockets TCP/IP, seguint un esquema client-servidor:

- El robot actua com a servidor, esperant comandes entrants.
- L'ordinador amb el mòdul de veu actua com a client, enviant les comandes quan es detecta una ordre vàlida.

Aquesta comunicació és estable, bidireccional i síncrona, i ha permès un control fluït del robot en temps real durant les proves.

El funcionament pràctic de la comunicació entre el mòdul de veu i el robot es mostra a la Figura 2, amb un exemple real del sistema interpretant una comanda i executant-la en temps real.



**Figura 2.** Comunicació entre el sistema de reconeixement de veu i el robot Stäubli en temps real.

(Font: Figura 5.1 de la memòria)

### 3.4. Control del robot amb VAL3

El robot industrial Stäubli TX60, amb 6 graus de llibertat, ha estat programat mitjançant el llenguatge VAL3, utilitzant l'entorn SRS19. El control s'ha implementat seguint els principis següents:

- Definició de punts clau: posicions inicials (*jHome*, *jWork*), punts d'inserció (trocar) i trajectòries.
- Moviments definits:
  - Translació: desplaçaments cap a dins o fora de la cavitat.
  - Pivotatge: rotació sobre l'eix Y, mantenint el punt de pivot fix (implementació de Remote Center of Motion).
- Actualització dinàmica de la tool i correcció geomètrica per assegurar que l'eina simula correctament el moviment d'una càmera quirúrgica sobre un trocar.

El sistema admet tant control remot (per veu) com control local (per botons físics), per fer proves o validacions sense necessitat del mòdul de veu.

## 4. RESULTATS

Un cop desenvolupat i integrat el sistema complet, es van dur a terme diverses proves de validació funcional per comprovar el correcte funcionament del reconeixement de veu, la comunicació entre mòduls i l'execució dels moviments del robot.

Durant les proves, es va constatar que:

- El sistema de reconeixement de veu en català permet controlar els moviments del robot mitjançant ordres com “entra”, “surt”, “dreta” i “esquerra”, sense necessitat d'un assistent físic.
- El mòdul Python interpreta correctament les ordres i les transmet per sockets TCP/IP al robot industrial Stäubli TX60, mantenint una comunicació estable.
- El robot executa moviments de translació i pivotatge amb fluïdesa, simulant el comportament d'una càmera laparoscòpica que manté fix el punt d'inserció (trocar), respectant el principi de Remote Center of Motion.
- El sistema també disposa d'un control manual alternatiu mitjançant botons físics, que permet validar el moviment del robot sense dependència del mòdul de veu.
- Es va verificar la coherència entre l'ordre vocal donada i la resposta mecànica del robot, així com la facilitat de canvi entre modes de control.

Aquestes proves s'han realitzat en un entorn de simulació, però els resultats obtinguts demostren que el sistema és funcional, fiable i coherent amb els objectius del projecte, obrint la porta a futures millores o adaptacions en entorns més avançats.

## 5. CONCLUSIONS

El projecte ha assolit amb èxit l'objectiu principal de desenvolupar un sistema de control per veu per a la manipulació d'un robot quirúrgic auxiliar, validant-ne tant el reconeixement vocal en català com la resposta precisa del robot davant de les comandes.

Els resultats obtinguts evidencien que la solució és eficaç, modular i escalable, i que pot ser integrada en entorns de simulació o formació quirúrgica, amb la possibilitat de ser adaptada a contextos clínics reals en fases posteriors.

D'altra banda, s'han identificat diverses possibilitats de millora i ampliació del sistema, com ara:

- Integració d'un sistema de reconeixement de veu offline per independitzar-se de serveis externs.
- Increment del nombre d'ordres per permetre moviments més complexos.
- Incorporació d'una interfície gràfica per fer el sistema més accessible en entorns formatius.
- Validació en entorns clínics reals o amb usuaris (cirurgians) per avaluar l'ergonomia i l'acceptació.

En conjunt, el treball representa una aportació pràctica i innovadora dins l'àmbit de la robòtica biomèdica, amb aplicació directa a la millora de la cirurgia assistida per veu.