# Voice Control of *da Vinci®* Surgical System

Lindsey Dean and H. Shawn Xu
Mentor: Anton Deguet

# Outline

1. Background
2. Motivation and Significance
3. Project Goal
4. Technical Approach
5. Deliverables
6. Tasks and Dependencies
7. Timeline and Milestones
8. Management Plan

# Background

# What is the *da Vinci*®?

- A robotic teleoperated surgical system that enables surgeons "to perform complex and delicate operations through a few tiny incisions with increased vision, precision, dexterity and control."
- A new generation of surgery
- Developed by Intuitive Surgical (NASDAQ: ISRG) in 1999
- $1-2.3M not including any attachments or maintenance plan

# *da Vinci*® Surgery

- Surgeon seated at HD console
- *EndoWrist* instruments
  - Precise
  - Tremor-reduced
  - Natural
- Ports: MI incisions

# Advantages

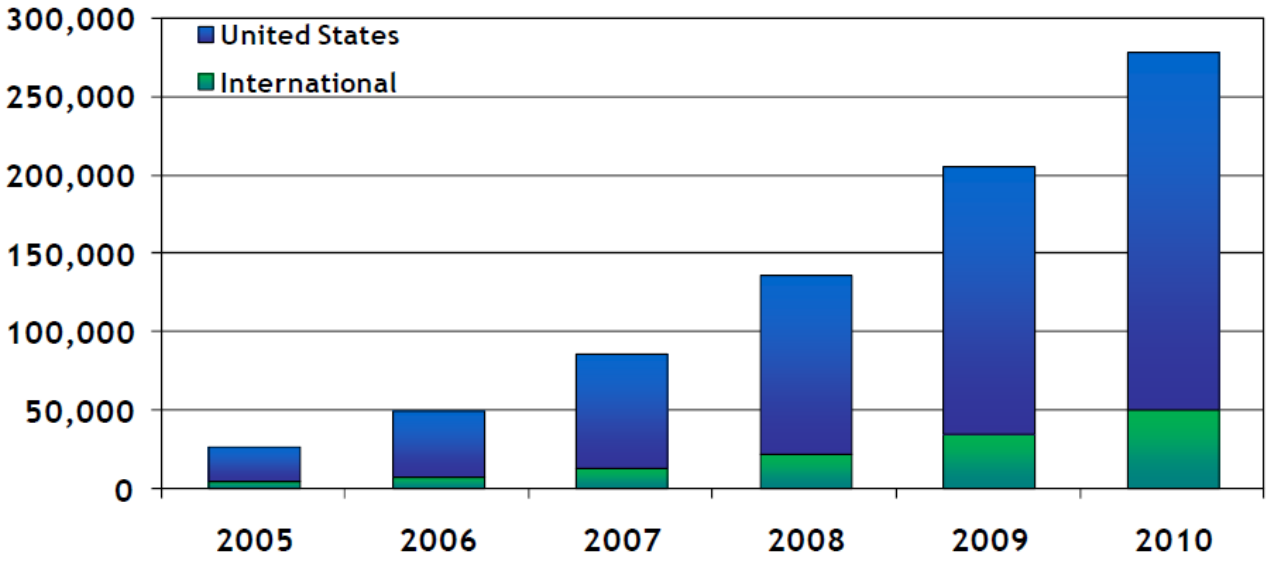
- Patient Value
  - Efficacy/Invasiveness
- Surgeon Value
  - Repeatable & Teachable
  - Reliable
- Hospital Value
  - Economic Benefit

# Uses

- Urology
  - **Prostatectomy**, nephrectomy, cystectomy, and pyeloplasty
- Gynecology
  - **Hysterectomy**, sacral colpopexy, myomectomy, and endometrial resection
- General Surgery
  - Colorectal procedures
- Cardiothoracic
  - Mitral valve repair, revascularization
- Head & Neck
  - Transoral procedures

# Usage Statistics

- 1,752 systems installed worldwide
- 278,000 procedures performed in 2010
  - Over 40% average yearly growth since 2005



# Motivation and Significance

# Problems

- Too many features to interact with
- A surgeon only has
  - One head
  - Two hands
  - Two feet
- Complex gestures
- Stop-start procedures
- Leads to inefficiency

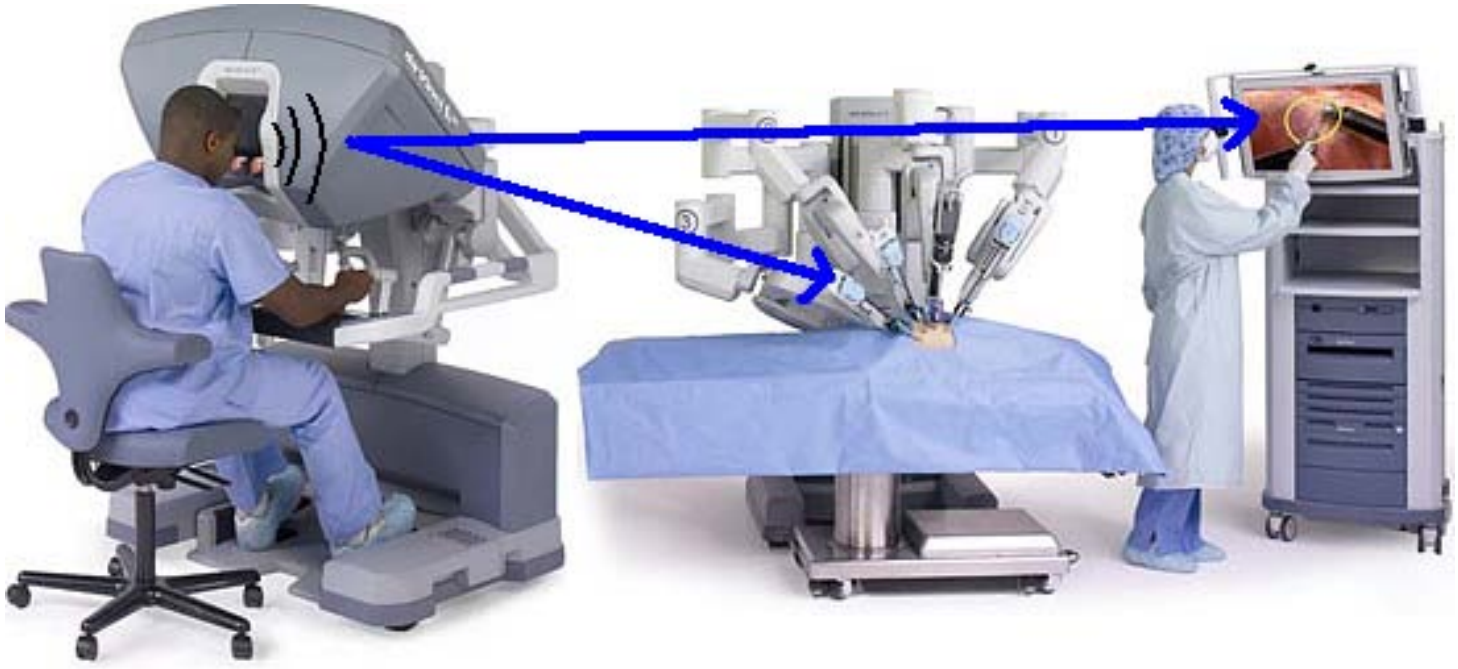
# An Analogous Example:

# Current Solutions

- Dual console interaction (Intuitive)
- Onscreen interactive digital 3DUI (CISST)
  - Allows surgeon
    - to overlay images on his viewing screen as he is performing surgery
    - to mark locations
    - to perform basic tasks such as measuring distance between two relevant points
- But…
  - Still need to pause surgery

# Proposed Solution

- Voice Control



# Project Goal

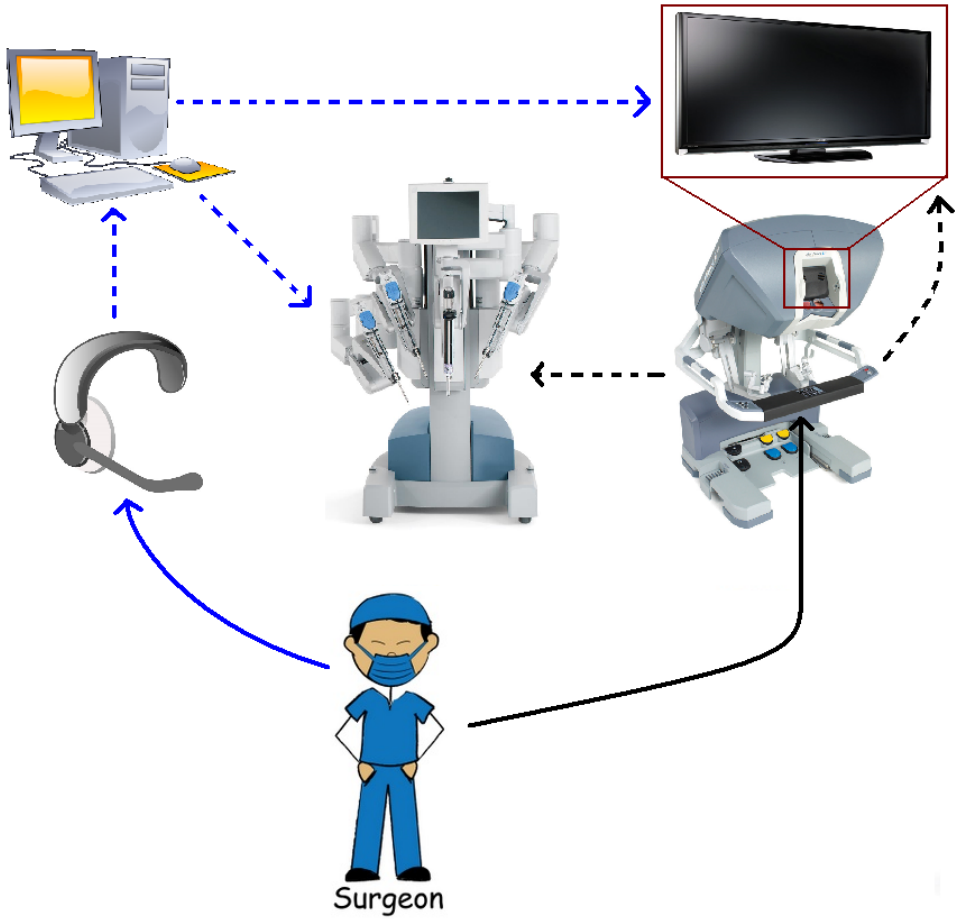
- Develop a way for the surgeon to interact with the surgical tools, camera, display, etc. using his/her voice

# Technical Approach

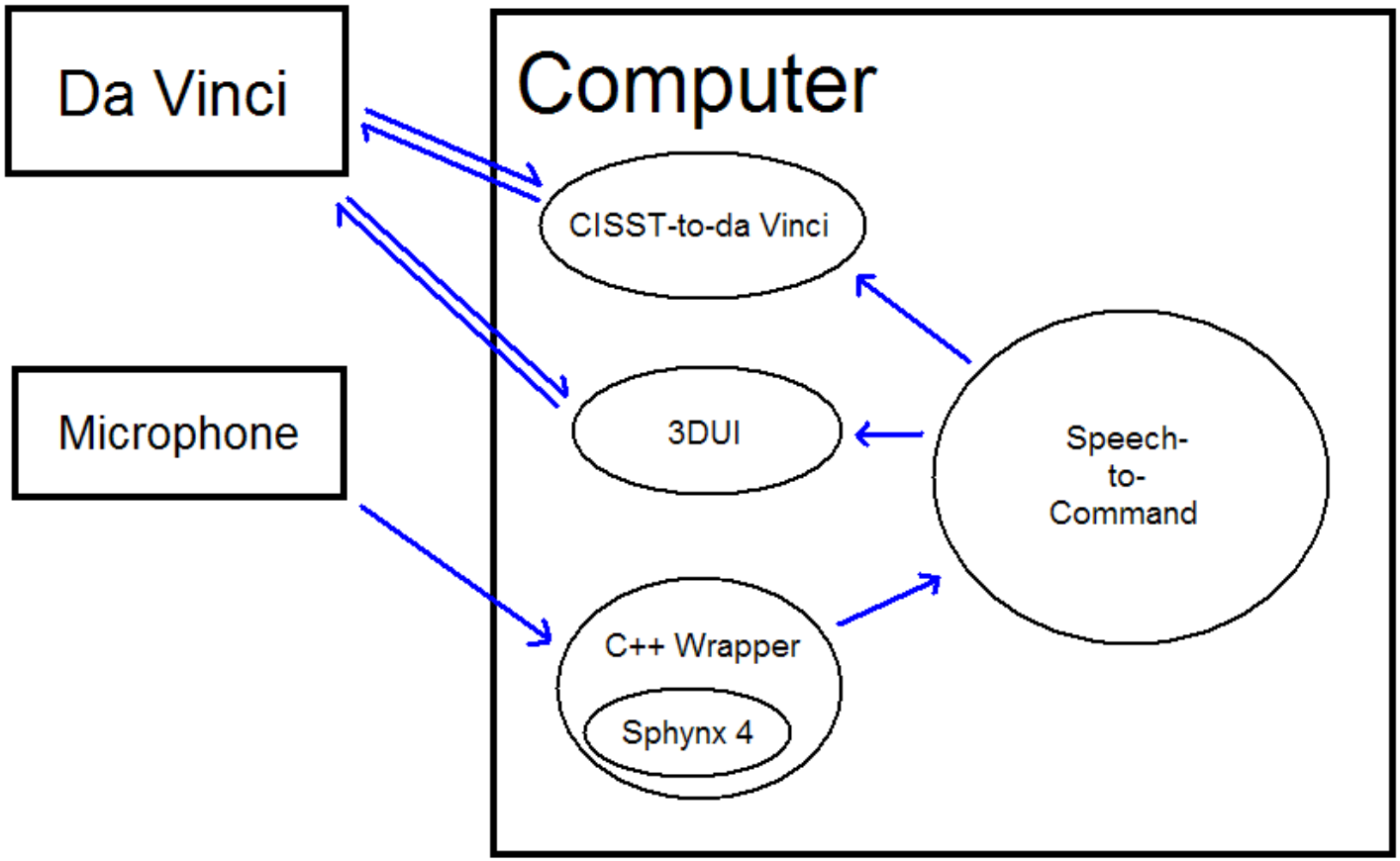
# What We Will Use

- Hardware
  - Microphone
  - *Da Vinci*® System
  - PC
- Software
  - CISST Libraries for interacting with *da Vinci*® Surgical System
  - C++ wrapper of Sphynx 4 JAVA speech recognition system
  - 3DUI: an interactive digital UI that surgeons can see and use during surgery
  - A text-to-speech package

# System Overview



# Software Architecture

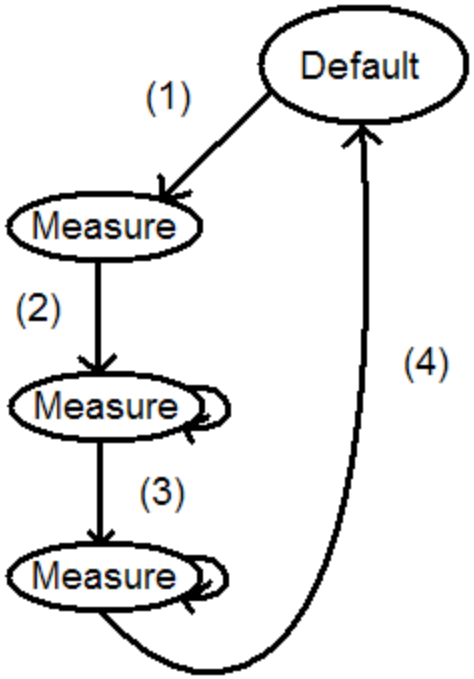


# State-Based Approach

- Currently, we are favoring a state-based implementation for speech-to-command
- Program keeps track of current "state" of system
- User can switch between different states using voice commands
- System state determines which voice commands are authorized at any time
- Universally accepted commands?

# Example

1. User says “measure left.” System switches to state ‘Measure Left’
2. User says “begin.” Because state is ‘Measure’, system begins measuring.
3. User moves tool controlled by left hand to new location and says “stop.” System displays distance moved.
4. User says “exit.” System exits to default state.



# Deliverables

# Deliverables

- Minimum
  - Well-documented program that adds singular functionality
  - A video demonstration of voice control
- Expected
  - Add multi-state functionality
  - Additional demonstration(s) that show different functions voice can perform on Da Vinci
- Maximum
  - Fully-functioning library of states and commands that can be easily expanded upon

# Tasks and Dependencies

# Exploratory Phase

- Experiment with voice recognition in other applications
  - What works well? What doesn't?
- Experiment with *da Vinci*® system and 3DUI
  - Determine what is most appropriate for voice integration
  - Determine what makes sense as "states"
- Goal:
  - Formulate lists of states and commands
  - Find ideas for video demonstrations

# Design Phase

- Familiarize ourselves with CISST libraries, specifically control of *da Vinci*® robot and 3DUI
- Familiarize ourselves with Sphynx 4 and C++ wrapper
- Flesh out software architecture
- Goal:
  - A detailed design of how the software will work, including classes, methods, key variables, etc.

# Implementation Phase

## 2 Concurrent Paths

1. Build speech-to-command program
2. Film video demonstrations

## Our Plan

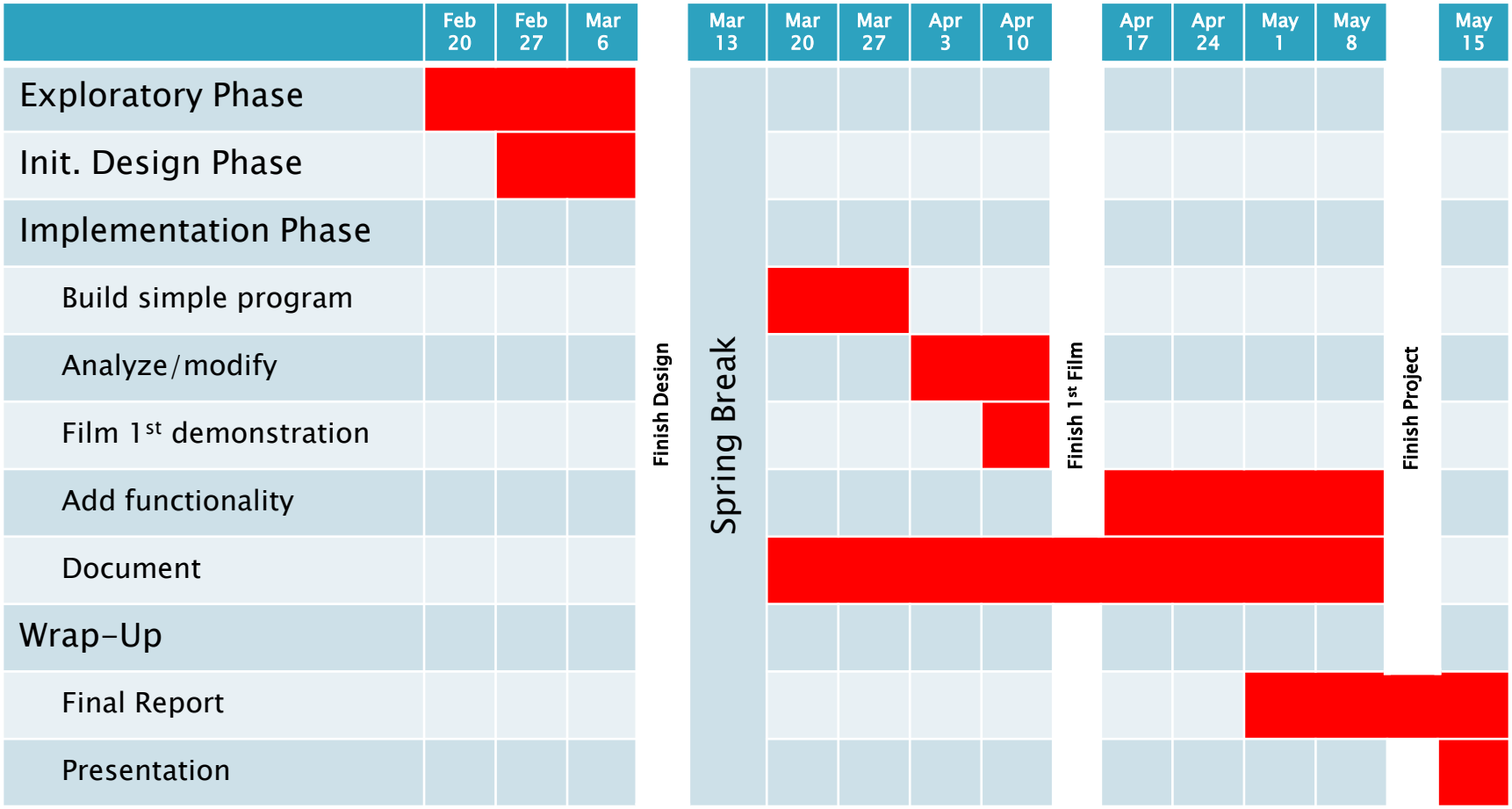
- Get a very simple program working that can do only one or two things
- Analyze our voice control process when used in practice and improve/modify accordingly
  - E.g.: should we ask for confirmation? is our process the best one in practice? safety issues?
  - Documentation is key here!
- Film first demonstration(s)
- Repeatedly add to program capabilities, analyze, document and film whenever we have a good idea for a demonstration

# Dependencies

| | Being done to resolve… | Affects.. | Resolve by… |
|---|---|---|---|
| **Access to Mock OR** | Allison Morrow | Exploratory Phase | 2/21 |
| **NDA w. Intuitive Surgical** | Alyssa | Exploratory Phase | 2/21 |
| **JHED Access to Mock OR computer** | Anton | Exploratory Phase | 2/27 |
| **Time with Anton to put necessary software on computers** | Anton | Design Phase | 3/6 |
| **Sphynx 4 C++ Wrapper** | CS undergrad should be done in 2 weeks | Design Phase | 3/12 |
| **Video Camera** | DMC has available to reserve | Implementation Phase | April |

# Timeline and Milestones

# Timeline


Feb 20
Feb 27
Mar 6
Mar 13
Mar 20
Mar 27
Apr 3
Apr 10
Apr 17
Apr 24
May 1
May 8
May 15
Exploratory Phase
Init. Design Phase
Implementation Phase
Build simple program
Analyze/modify
Film 1st demonstration
Add functionality
Document
Wrap-Up
Final Report
Presentation
Finish Design
Spring Break
Finish 1st Film
Finish Project

# Milestones

- March 13
  - Finish design and list of states and commands
  - Finish software architecture of speech-to-command program
- April 17
  - Finish first video demonstration
- May 15
  - Wrap up all coding
- May 19
  - Poster presentation

# Management Plan

# Management Specifics

- Weekly meetings with Anton Deguet
- We will be meeting at least twice per week, depending on week-to-week schedule
- During the 1st two phases, our weekly schedule will revolve mostly around when the mock OR is available
- Partner programming in the beginning to learn software
- Weekly on Sunday: assessment of progress and adjustment of schedule

# Thank You

- Questions/Comments?

# Bibliography


- A. Kapoor, A. Deguet, and P. Kazanzides, "Software components and frameworks for medical robot control," in Robotics and Automation, 2006. ICRA 2006. Proceedings 2006 IEEE International Conference on, 2006, pp. 3813–3818.
- A. Deguet, R. Kumar, R. Taylor, and P. Kazanzides, “The cisst libraries for computer assisted intervention systems,” in MICCAI Workshop on Systems and Arch. for Computer Assisted Interventions, Midas Journal, Sep 2008.
- Liu, Peter X., A.D. C. Chan, and R. Chen. "Voice Based Robot Control." *International Conference on Information Acquisition (ICIA)* (2005): 543. Web. <http://ieeexplore.ieee.org.proxy3.library.jhu.edu/stamp/stamp.jsp?tp=&arnumber=1635148>.
- Sevinc, Gorkem. *INTEGRATION AND EVALUATION OF INTERACTIVE SPEECH CONTROL IN ROBOTIC SURGERY*. Thesis. Johns Hopkins University, 2010. Print.
- Schalkwyk, J., Hetherington, L., Story, E. "Speech Recognition with Dynamic Grammars Using Finite-State Transducers." Proc. Eurospeech. Geneva, Switzerland, 2003.