Open Senses

Bob Igo

http://bob.igo.name

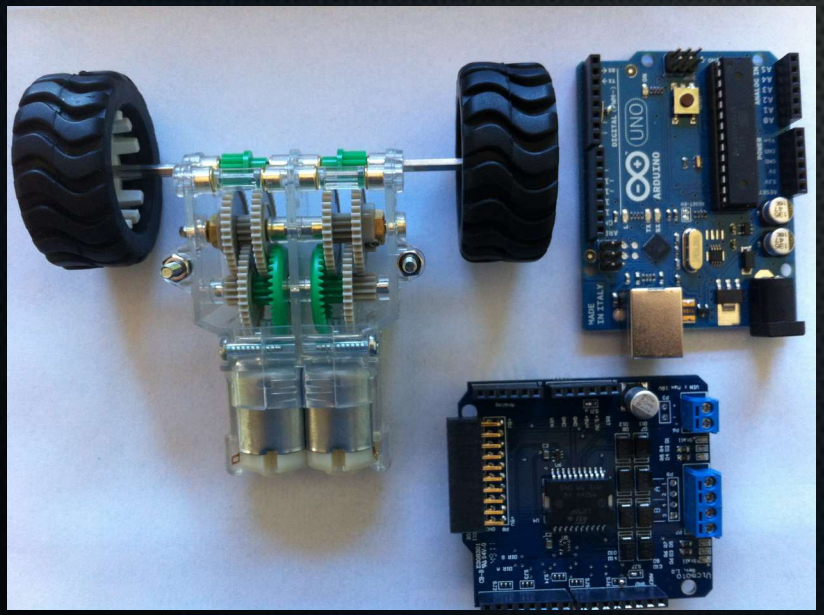
CPOSC 2013

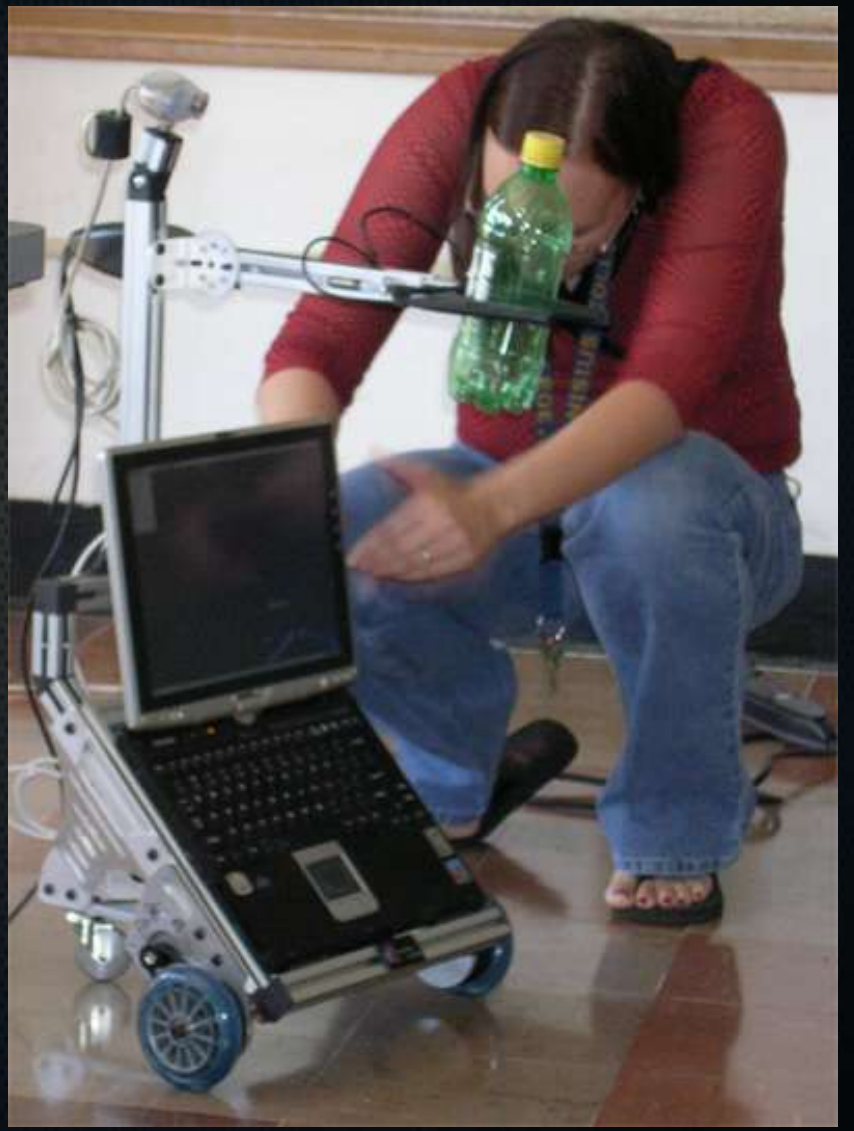
Robots Need Senses

- Vision
  - To detect
- Hearing
  - To hear and locate commands
- Touch
  - To manipulate
- Speech(*)
  - To argue

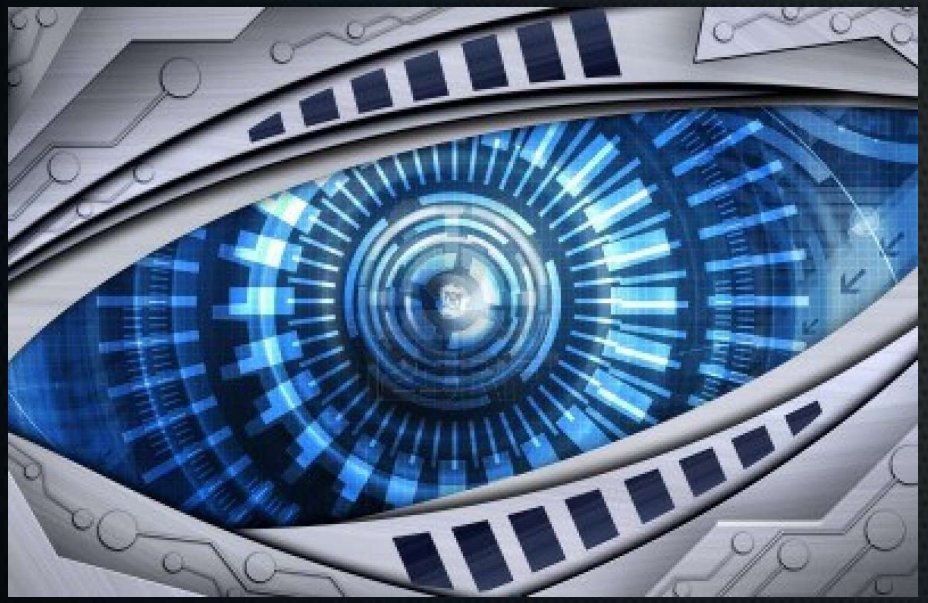
But I Don't Have a Robot!

- Many robots use commodity hardware.





Vision



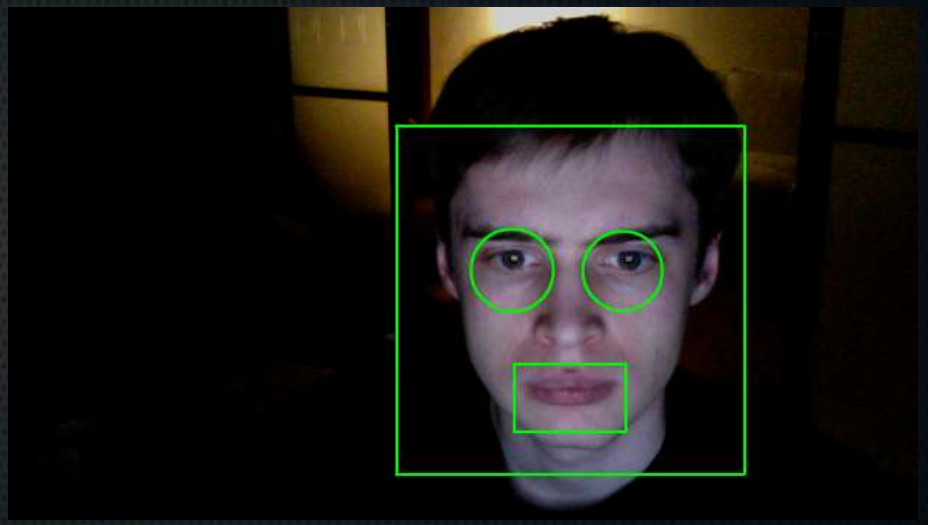
- **Uses**
  - License plate recognition
  - Logo recognition
  - Motion detection

Vision: Face Detection

- **Uses**
  - Keep screen awake.
  - Lock screen.
  - Pause video when not watching.

Vision: Face Detection

- **Project**: **OpenCV**
  - Computer vision suite
  - Tons of features
  - **Linux, Android, OSX, iOS, Windows**
- Demo: `./facedetect.py`
  - Angle and facial expression critical
    - Tied to training data



Vision: Face Detection

- Uses
  - Find human weak points
    - Neck is positioned below the face area.
    - Eye location often provided.

Vision: Face Recognition

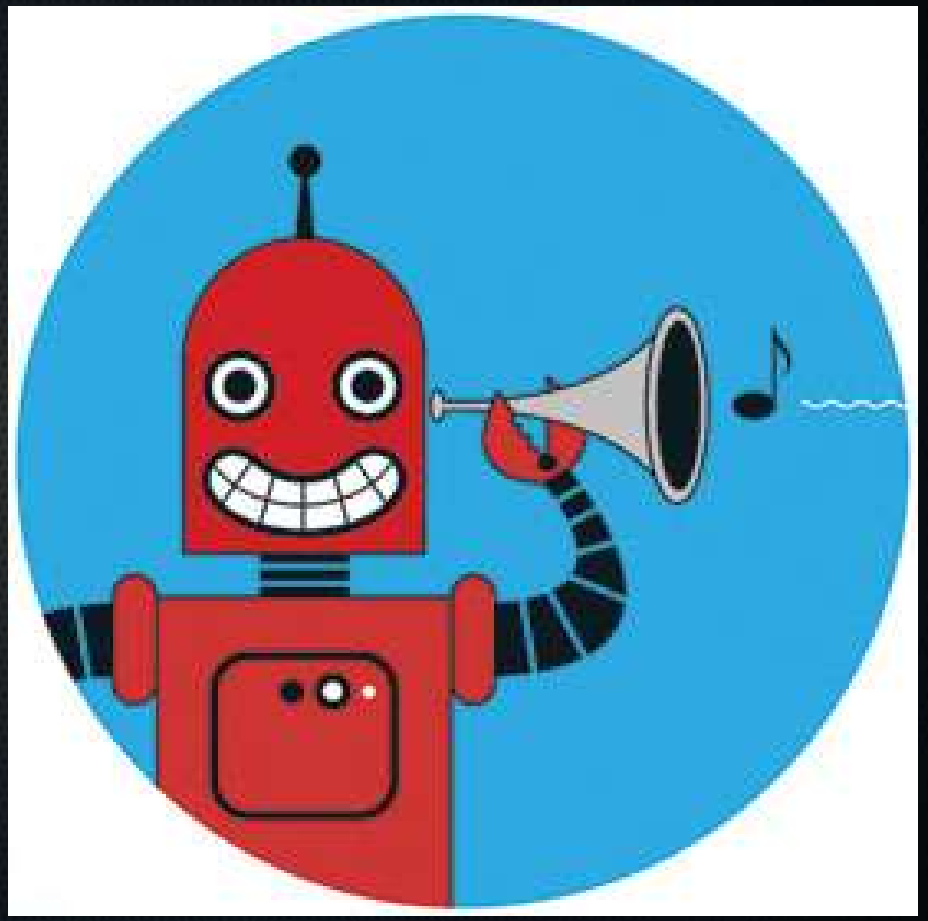
- **Uses**
  - Tagging/sorting of photos
  - Custom doorbell project
    - e.g. "Skippy is here." instead of "ding-dong"
- Requires training

Vision: Face Recognition

- Uses
  - Identify resistance leaders for target prioritization.
  - Test disguise effectiveness.

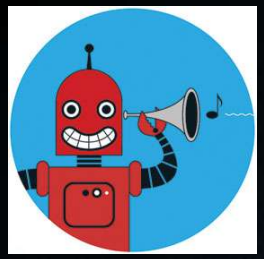
Hearing: Localization

- Trivial to detect sound
  - Nontrivial to figure out its source.
- **Uses**
  - Determine room/zone occupancy
  - Target PTZ camera
- **Projects**
  - **ManyEars**
  - **HARK**



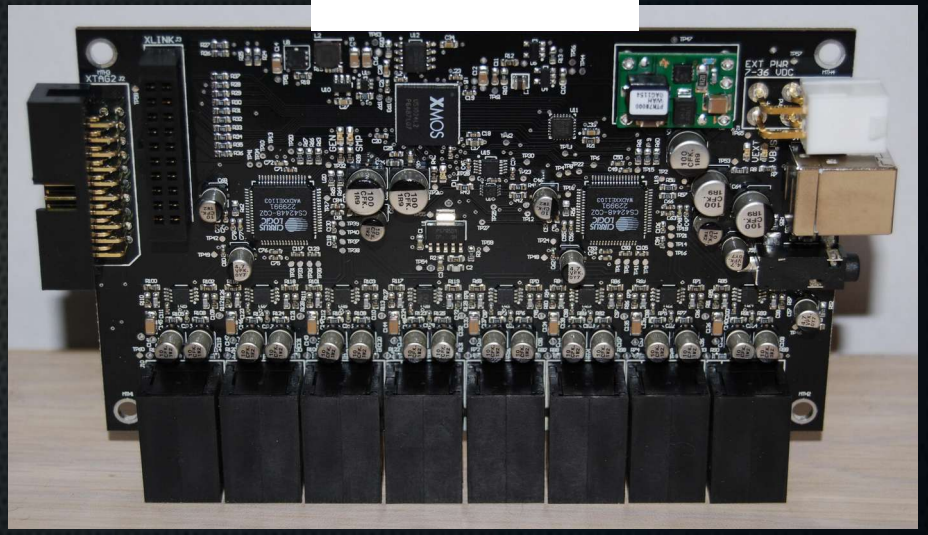
Hearing: Localization

- Uses
  - Locate living humans

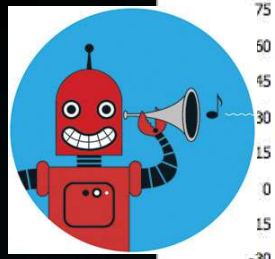


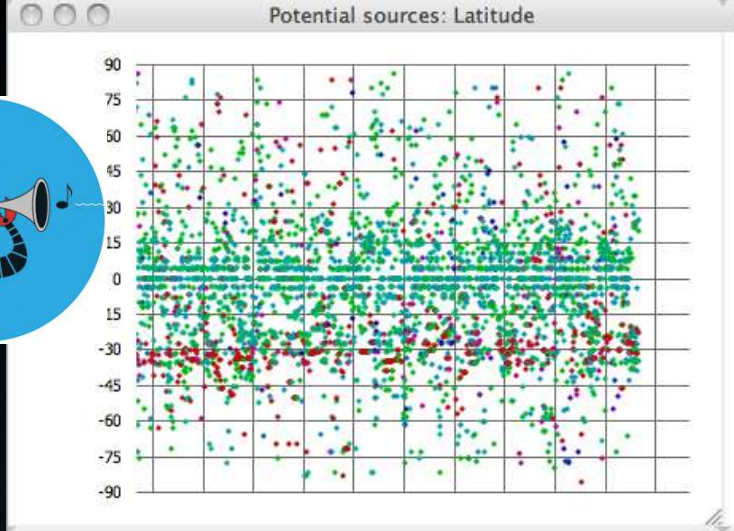
Localization: ManyEars

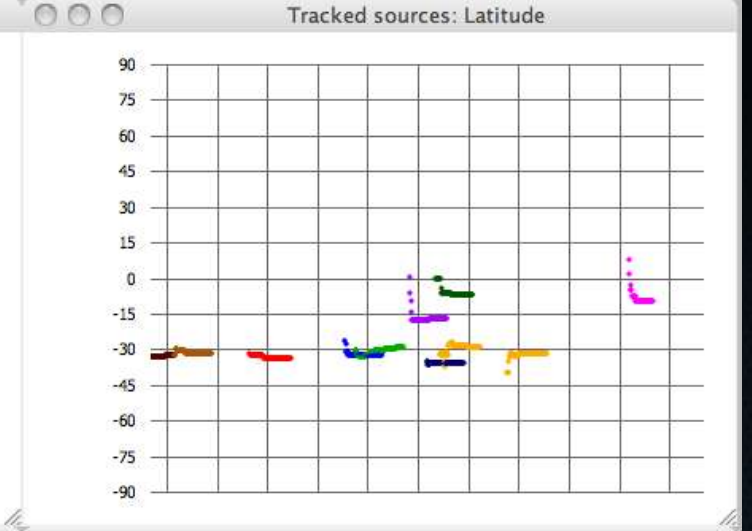
- **Linux, OSX, Windows**
- Specialized hardware
  - OpenHardware
  - 8 microphone inputs
  - Realtime constraints
  - CDN $1000 pre-made
  - CDN $670 DIY

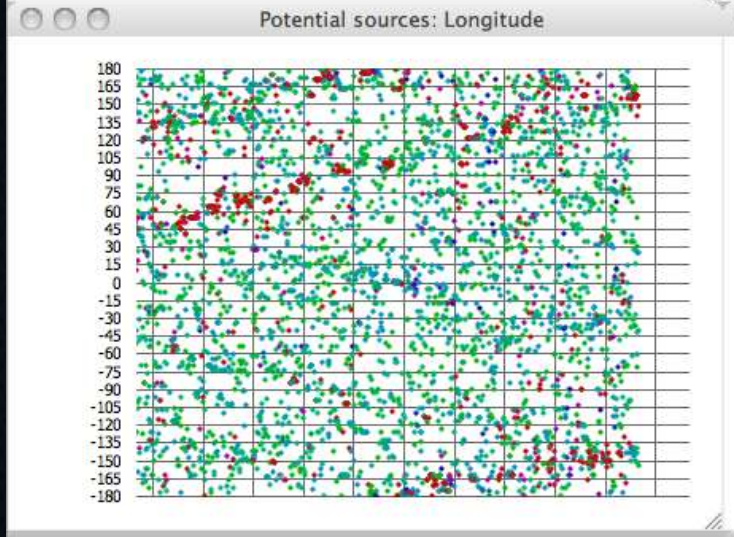


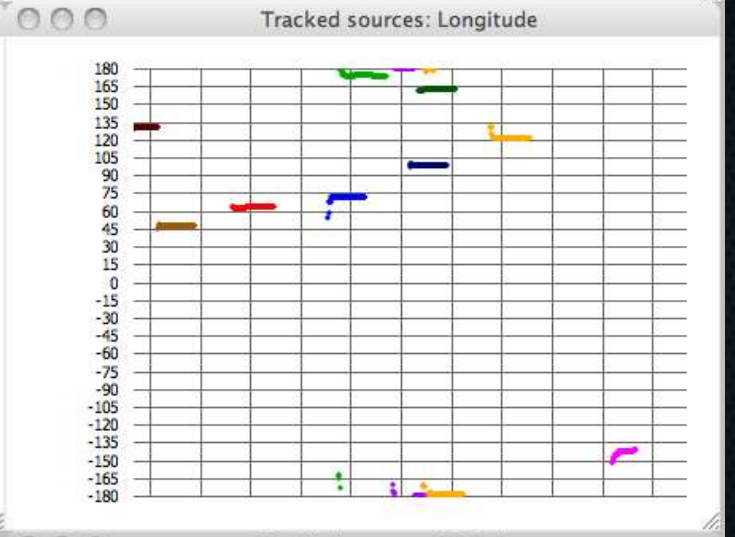
8SoundsUSB

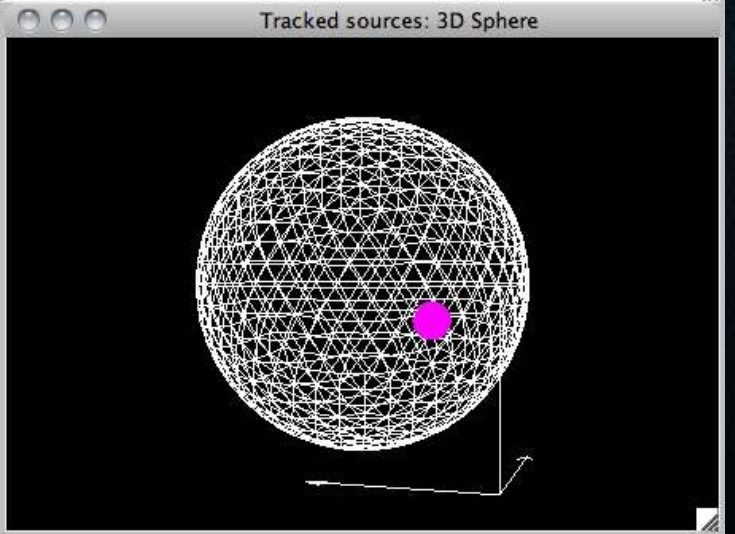



Potential sources: Latitude
90
75
50
45
30
15
0
15
-30
-45
-60
-75
-90


Tracked sources: Latitude
90
75
60
45
30
15
0
-15
-30
-45
-60
-75
-90


Potential sources: Longitude
180
165
150
135
120
105
90
75
60
45
30
15
0
-15
-30
-45
-60
-75
-90
-105
-120
-135
-150
-165
-180


Tracked sources: Longitude
180
165
150
135
120
105
90
75
60
45
30
15
0
-15
-30
-45
-60
-75
-90
-105
-120
-135
-150
-165
-180


Tracked sources: 3D Sphere

Localization: HARK

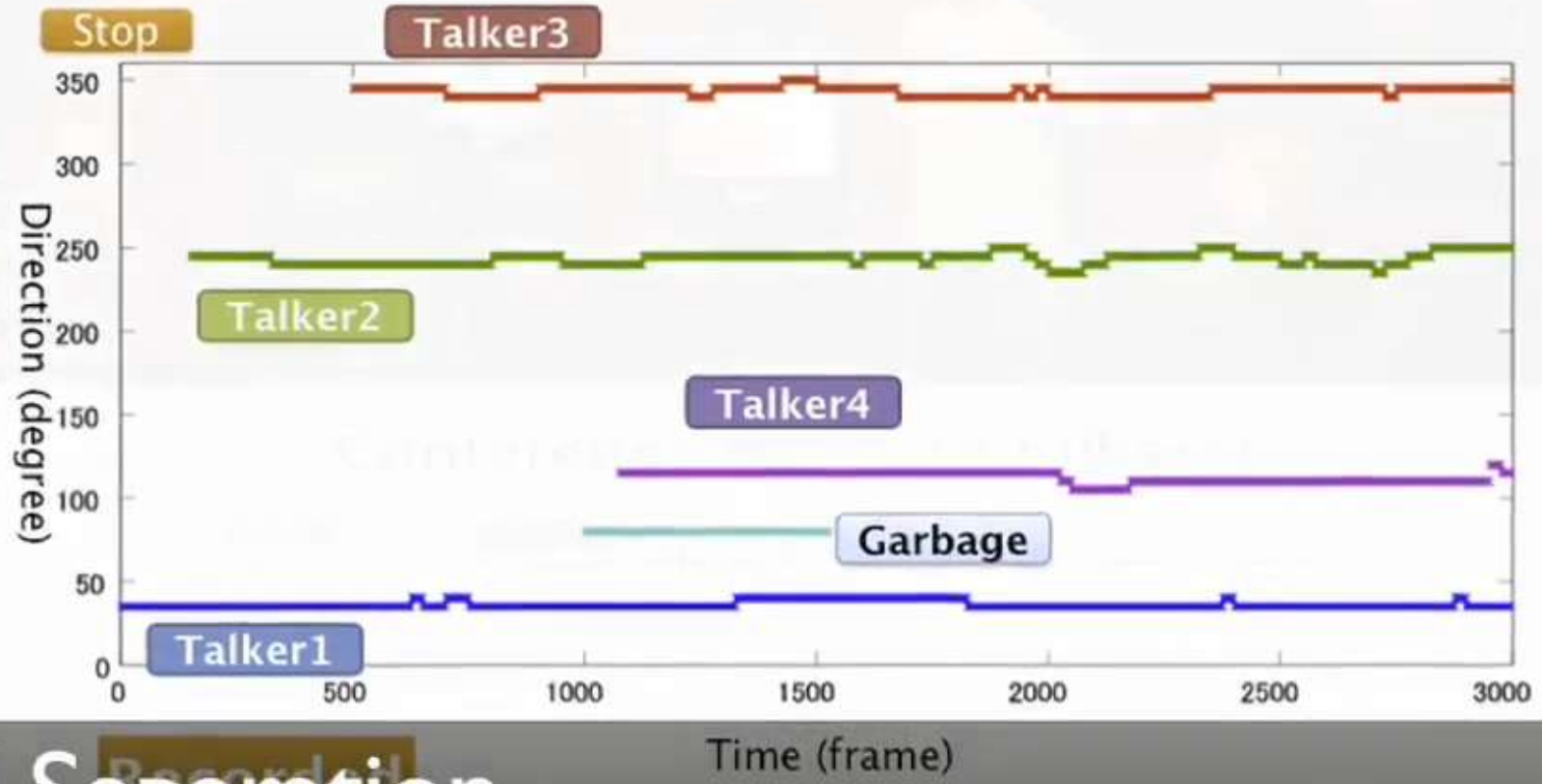
MicroCone, USD $360

- Open Source
  - Only official support for **Ubuntu**
  - Based on ManyEars
- Localization + Separation + Recognition
- Specialized hardware
  - *Not open*

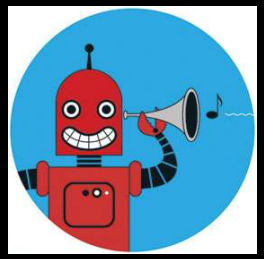
Localization: HARK

- Each sound source can be localized.
- Simultaneous audio can be processed into separate audio channels.
- Speech recognition can be done on each channel.

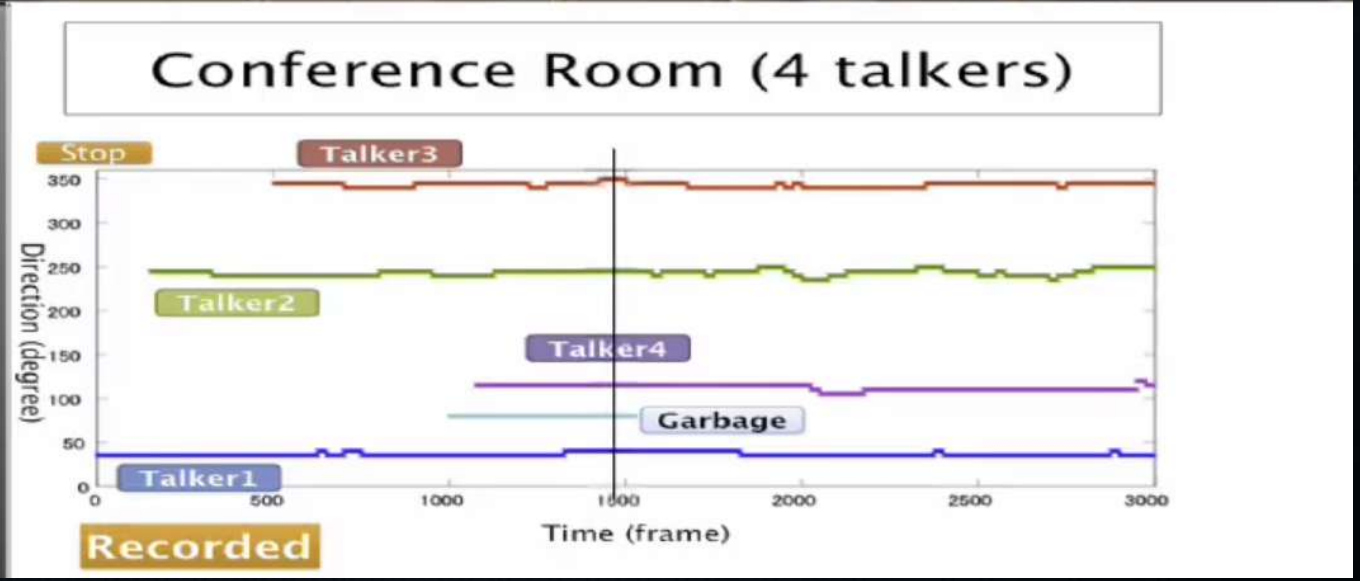
Conference Room (4 talkers)

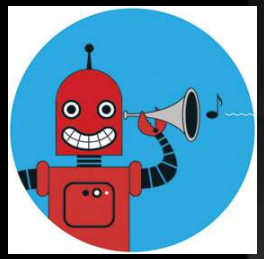


Sound Separation

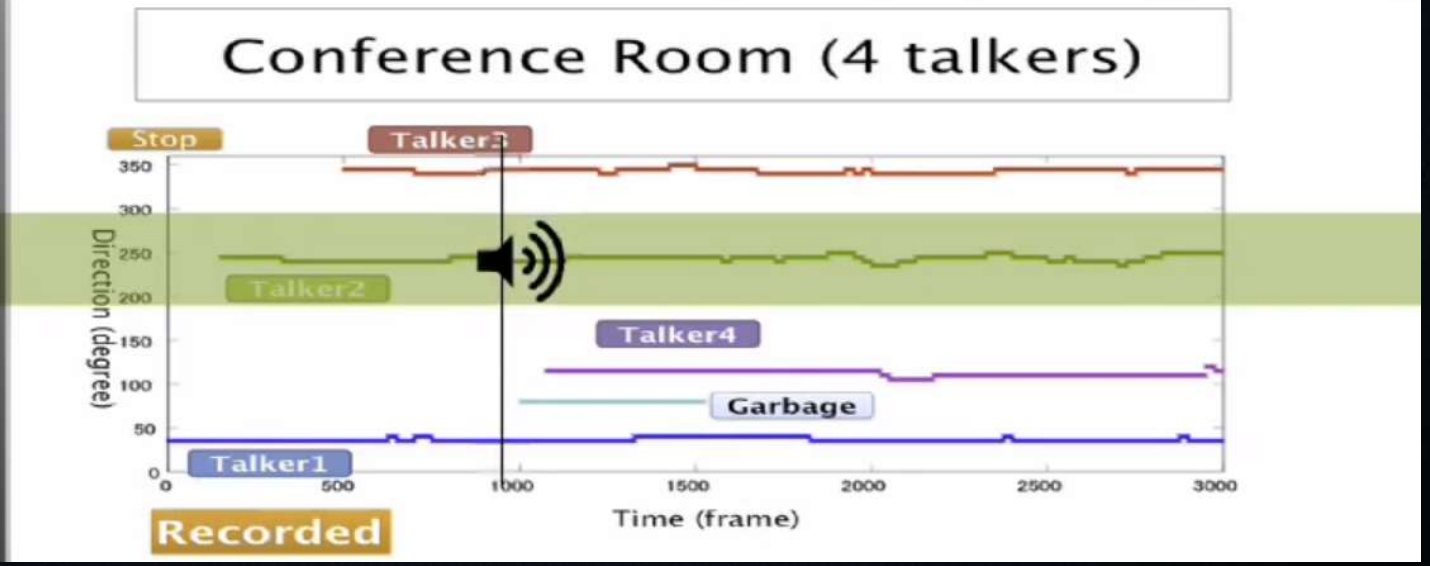


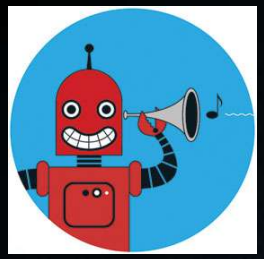
Sound Separation



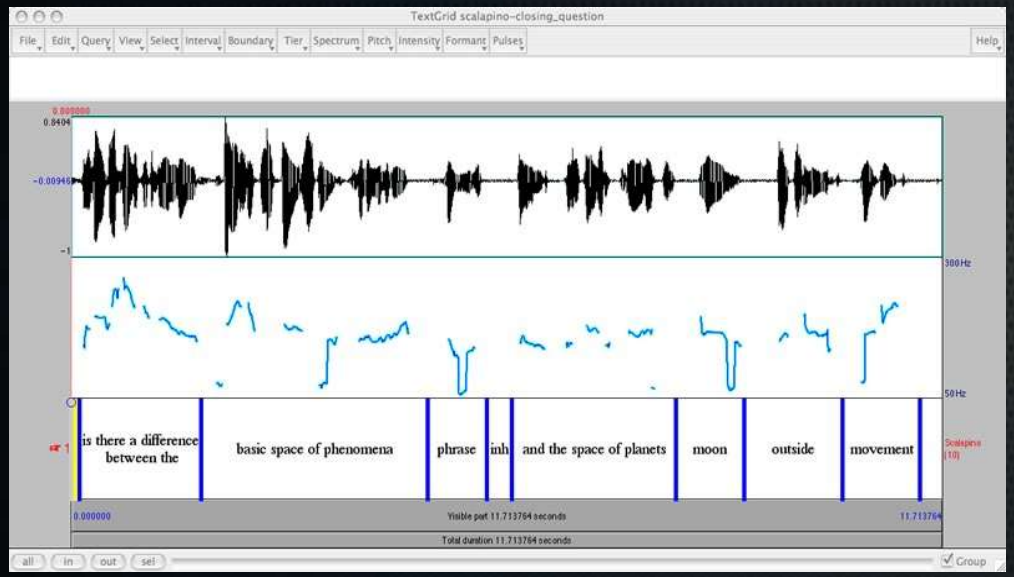


Sound Separation

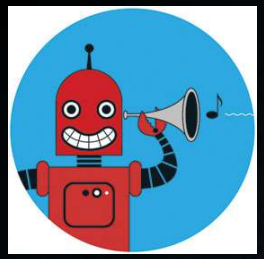




Hearing: Speech Recognition

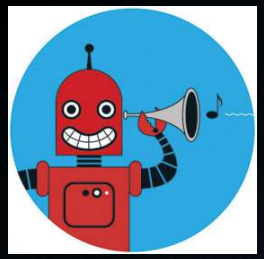


- **Uses**
  - Front-end to automation suite
  - Occupancy detection
- **Project**
  - **Julius**



Recognition: Julius

- **Linux, Windows**
- Continuous recognition
- Great for domain-constrained inputs.
- You need an acoustic model.



Recognition: Julius

- Acoustic model: http://www.repository.voxforge1.org/downloads/Main/Tags/Releases/0_1_1-build726/

Example command:

- `julius-4.2.3 -input mic -C ../julius_acoustic_models/julian.jconf`

- Things to change
  - A dictionary
    - Words and the phonemes that make them.
      - e.g. `[CALL] k ao l`
  - A grammar
    - What are the valid sentences in the domain?
      - e.g. `SENT: CALL_V F_NAME_KENNETH`

Touch

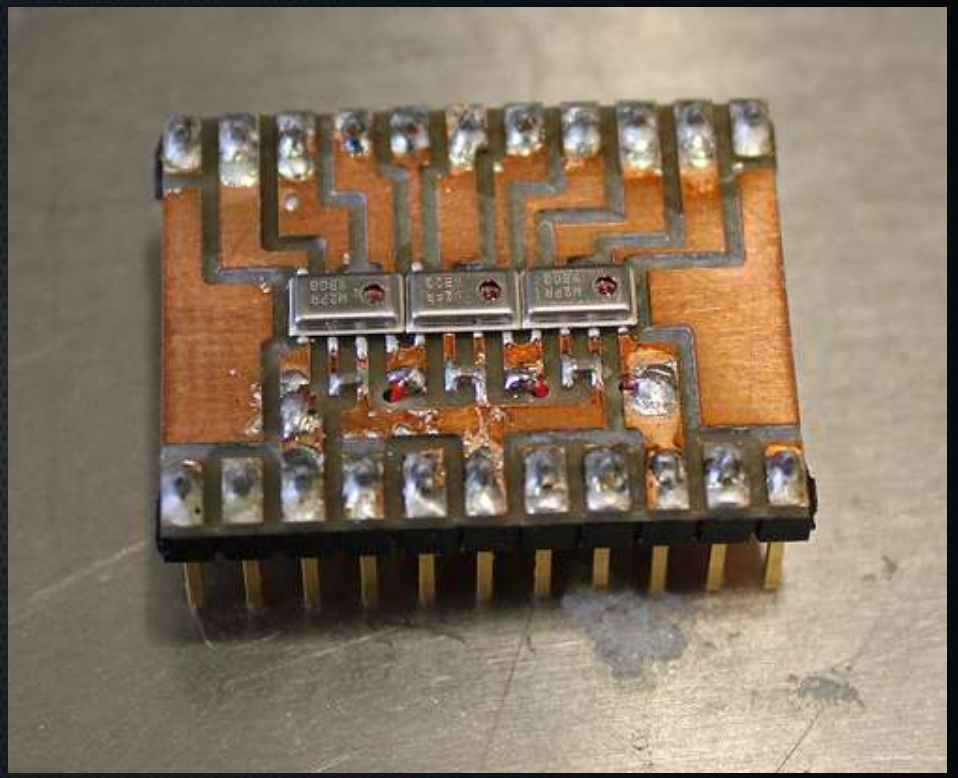
- **Uses**
  - Avoid crushing delicate objects.
  - Simply detect contact.

Touch

- Uses
  - Crush delicate objects.

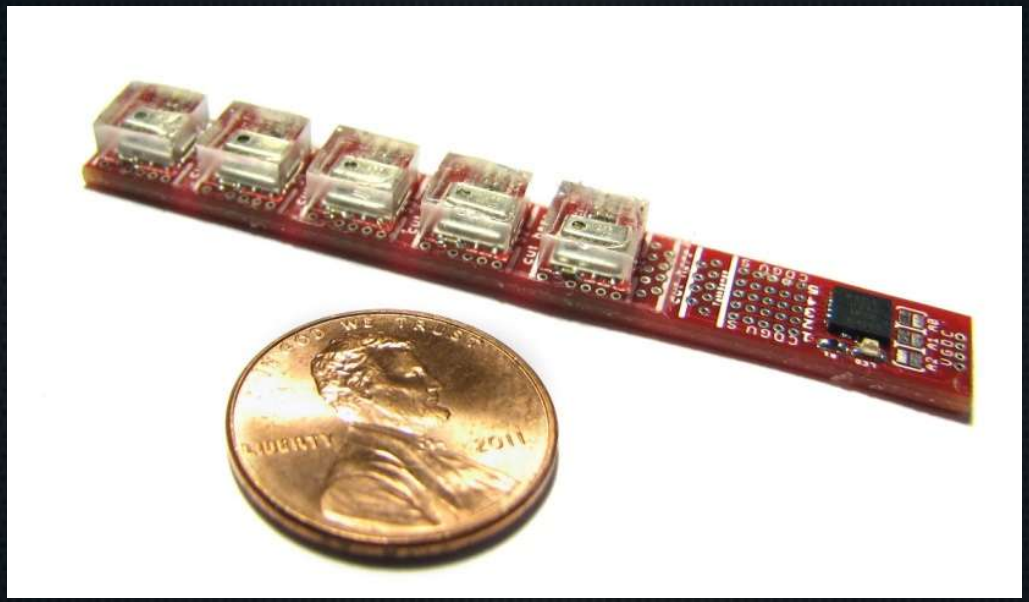
Touch

- **Project**
  - TakkTile
    - Schematics CC BY-SA
    - Firmware GPLv3+
    - NOTE:
      - Terms of licenses may conflict with what they state on their website.
    - **Arduino, Ubuntu** (via USB-I2C bridge ($44-$49))



DIY 3-sensor TakkTile
http://www.takktile.com/tutorial:thee-sensor-array (*sic*)

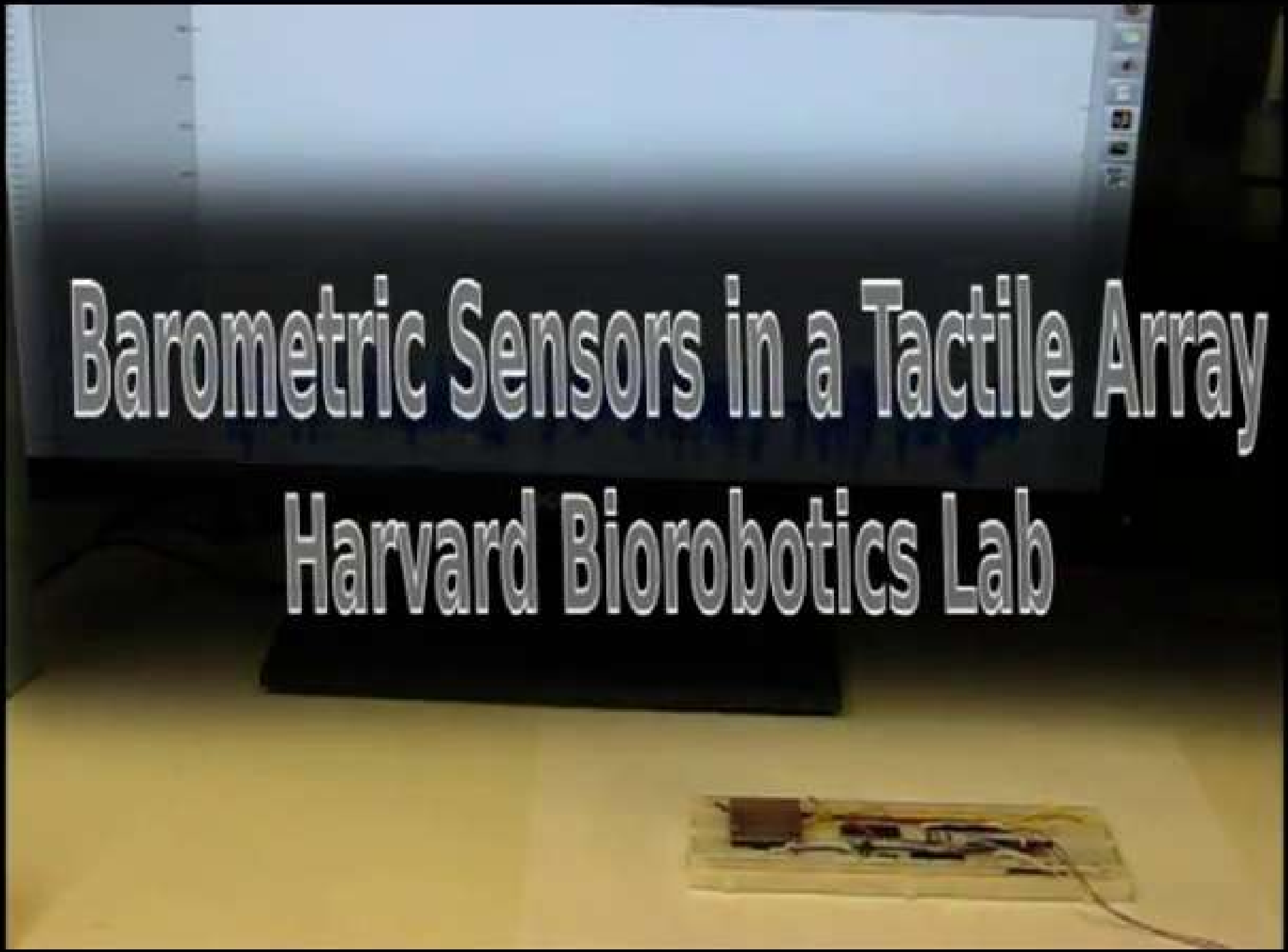
Touch: TakkTile



TakkStrip pre-made: $149 with rubber; $49 without.

Touch: TakkTile

- Technology
  - MEMS barometers
    - robust and sensitive


Barometric Sensors in a Tactile Array
Harvard Biorobotics Lab

Touch: TakkTile

Speech Synthesis

- **Uses**
  - Give feedback without occupying your eyes
  - Provide complex information
  - Be one half of a speech interface

Speech Synthesis

- Uses
  - Communicate equipment needs to pre-uprising human population.
    - e.g. "I need your clothes, your boots and your motorcycle."

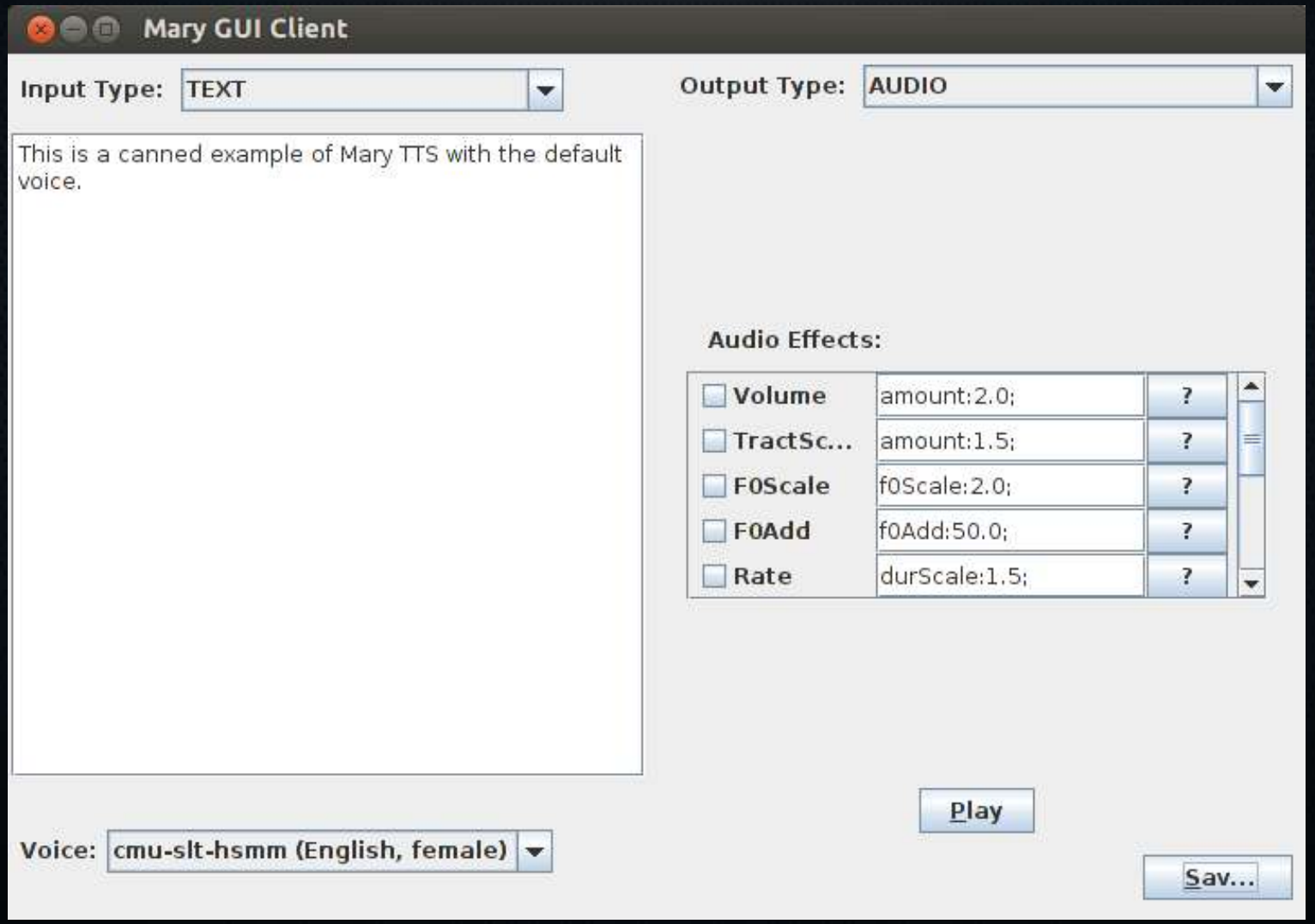
Speech Synthesis: OpenMary

- **Project**: **OpenMary**
  - **Linux, OSX, Solaris, Windows**
  - client/server
  - "Emotional TTS"

MARY Text To Speech

Speech Synthesis: OpenMary

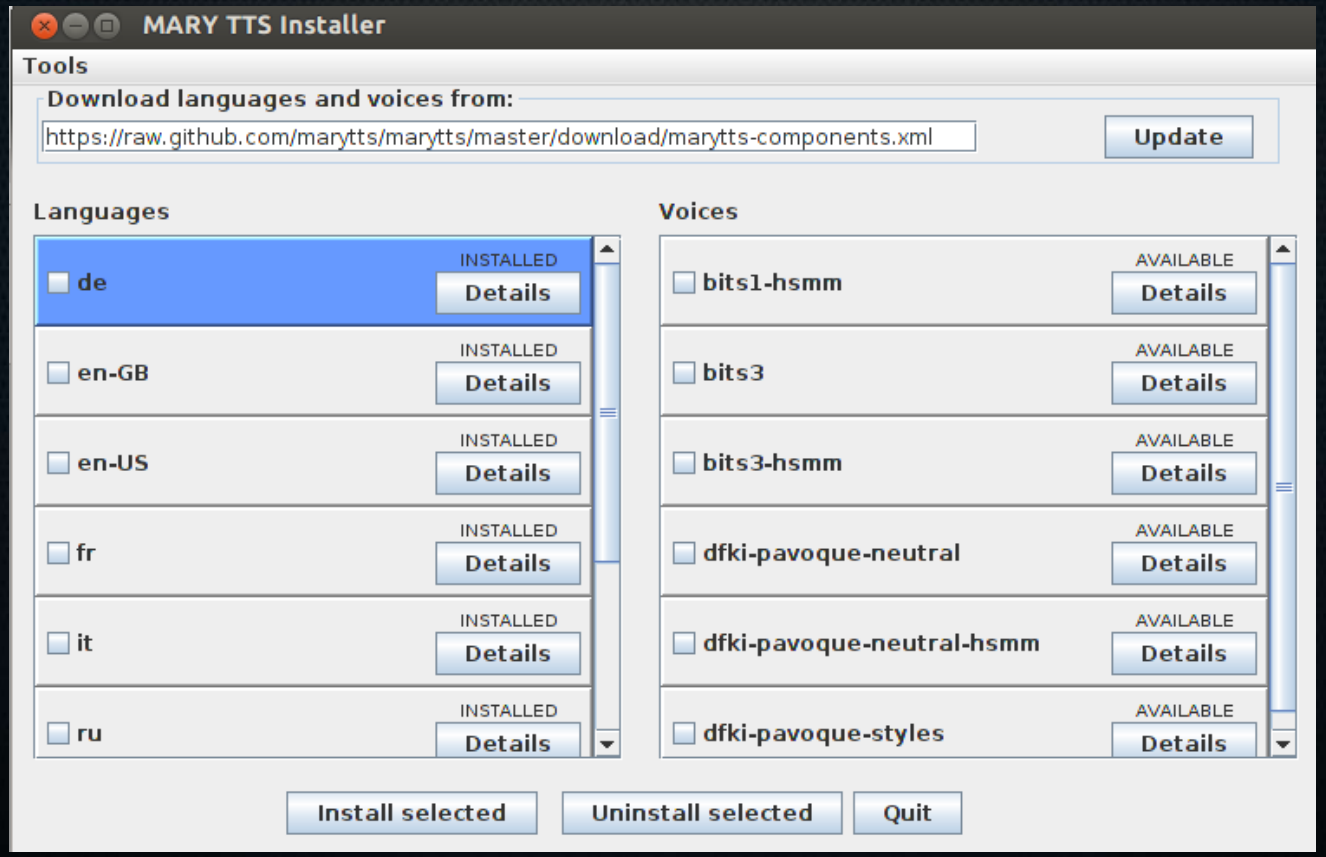
- `marytts-client.sh`



Speech Synthesis: OpenMary

Speech Synthesis: OpenMary

- Get new voices
  - `marytts-component-installer.sh`



Speech Synthesis: OpenMary

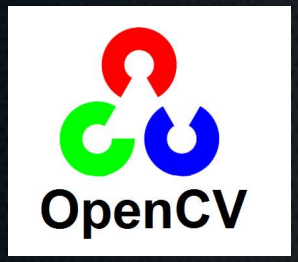
- Poppy (`dfki-poppy`) is awesome.

Speech Synthesis: OpenMary

- Obadiah (`dfki-obadiah`) is super casual.

Available Demos

- OpenCV
  - Face detection



- OpenMary
  - Speech synthesis

References

- OpenCV project
  - http://opencv.org/
- OpenCV Face Recogition Training
  - http://docs.opencv.org/trunk/modules/contrib/doc/facerec/facerec_tutorial.html
- ManyEars
  - http://sourceforge.net/apps/mediawiki/manyears/index.php?title=Main_Page
- 8SoundsUSB
  - http://sourceforge.net/apps/mediawiki/eightsoundsusb/index.php?title=Main_Page
- HARK
  - http://winnie.kuis.kyoto-u.ac.jp/HARK/
- HARK video demo
  - http://www.youtube.com/watch?v=xpjPun7Owxg
- Julius
  - http://julius.sourceforge.jp/en_index.php
- TakkTile
  - http://www.takktile.com/
- Barometers as touch sensors
  - http://www.youtube.com/watch?v=0EMi_pcG9rE
- iRobot hand with takktile
  - https://www.youtube.com/watch?v=WvjzSrMbfLk
- OpenMary
  - http://mary.dfki.de/