Speaker Identification (SID) in Low-Rate Coded Speech

Andrew Catellier and Stephen Voran
{acatellier,svoran}@its.bldrdoc.gov
Institute for Telecommunication Sciences
Telecommunications Theory Division

Effect of Transmission on Multimedia Quality of Service
17-19 June 2008 – Prague, Czech Republic

ITS

SID: Motivation

- New equipment for first responders
- Anecdotal complaints about system performance
  - Speaker emotional state, and speaker identity obscured
- Legitimate concerns!

SID: Motivation

- To help improve the platform, these problems must be measured
- Two experiments were designed to measure the problem
- We'll be talking about SID today

SID: Experiment Design

- Specifications:
  - Unfamiliar talkers
  - Clips with and without prosodic information (short and long clips)
  - Six simulated communication systems
  - Manageable experiment length

SID: Experiment Design

- Realization:
  - Tactical Speaker Identification Database
    - Used three males and three females
  - Three clip lengths: sentence, four digits, two digits
  - MELP, IMBE 7.2, 3.6 kbps (with and without impairments), MNRU
  - 360 total clips, experiment length around one hour

SID: Experiment Design

- C1 clips produced by resampling at 8 kHz, filtering (160-3640 Hz bandpass), and then normalized to -26 dB below clipping
- Low-rate vocoders in C2, C3, C5 and C6 are similar to those used in Public Safety communication systems
- C5 and C6 have additional transmission impairments

Condition (C)	Description
C1	Null (no further processing)
C2	IMBE Codec, 7.2-kbps gross 4.4-kbps net
C3	MELP Codec, 1.2-kbps net
C4	MNRU, $Q = 6$ dB SNR
C5	IMBE Codec, 3.6-kbps gross 2.45-kbps net 7% BER, random
C6	C5+Packet Impairments+C5
	Packet Impairments: create 60 ms packets, delete 10% of packets at random, insert the same number of empty packets at random and apply PLC to them

Table 1. Six conditions used in the experiment.

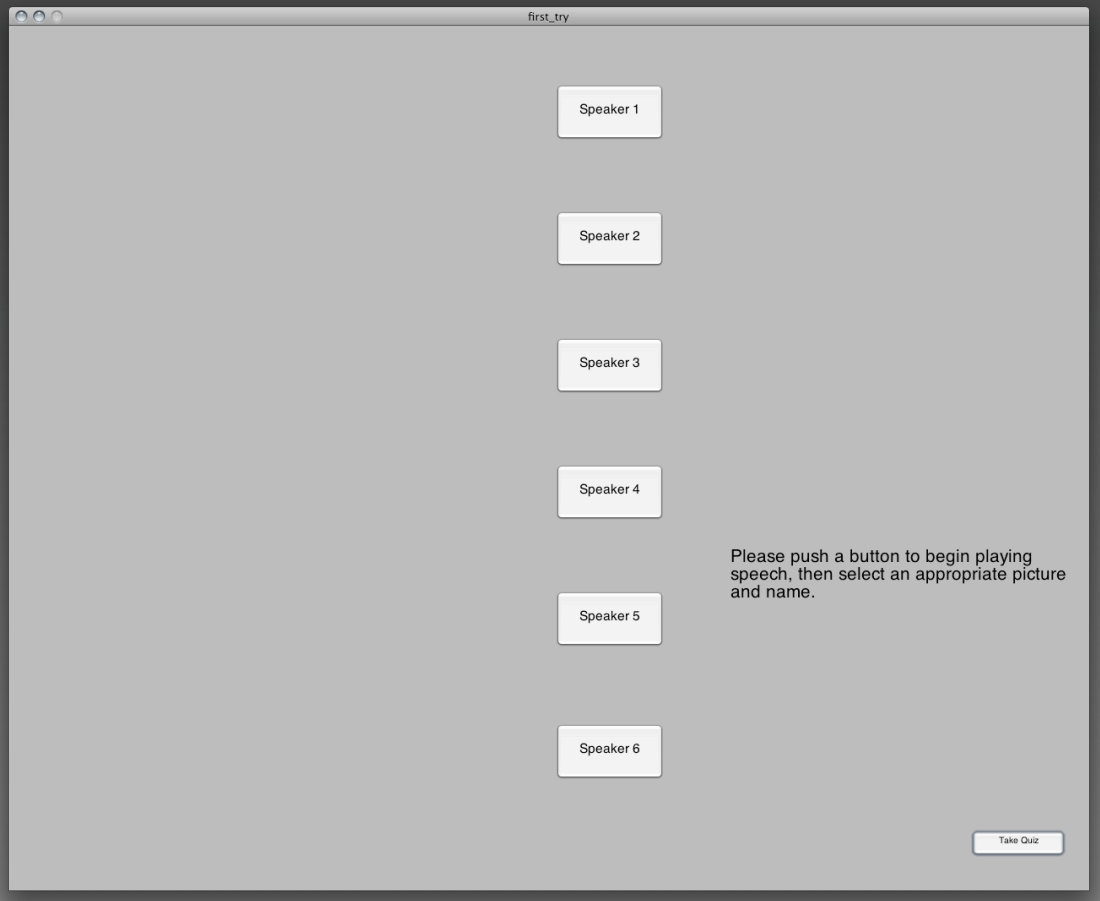
SID: Experiment Design

- Problems:
  - Training listeners to accurately recognize any given speaker
  - False confidence in training
  - Mid-test mistraining

SID: Experiment Design

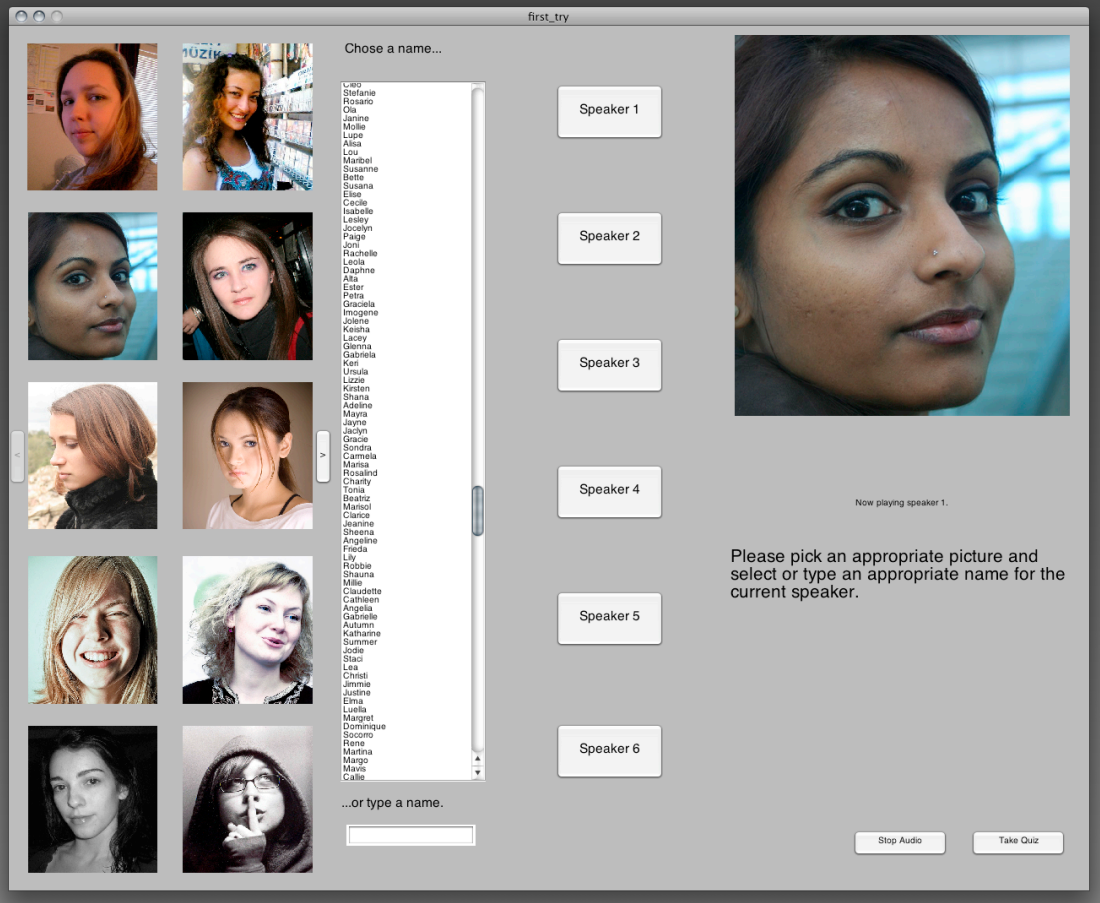
- Training:
  - Used set of clips where speakers were giving directions (semi-spontaneous)
  - Allowed listeners to assign appropriate memory aids: a name and a face

SID: Experiment Design

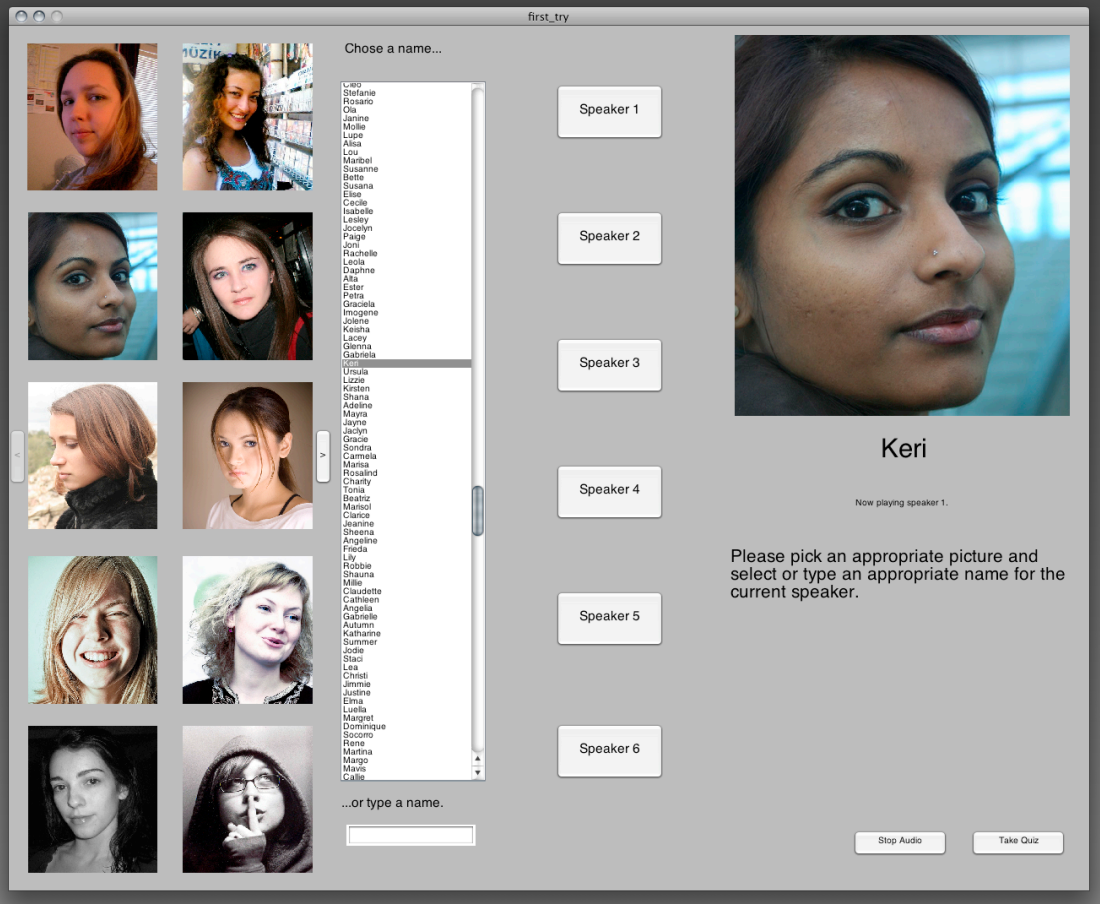


SID: Experiment Design

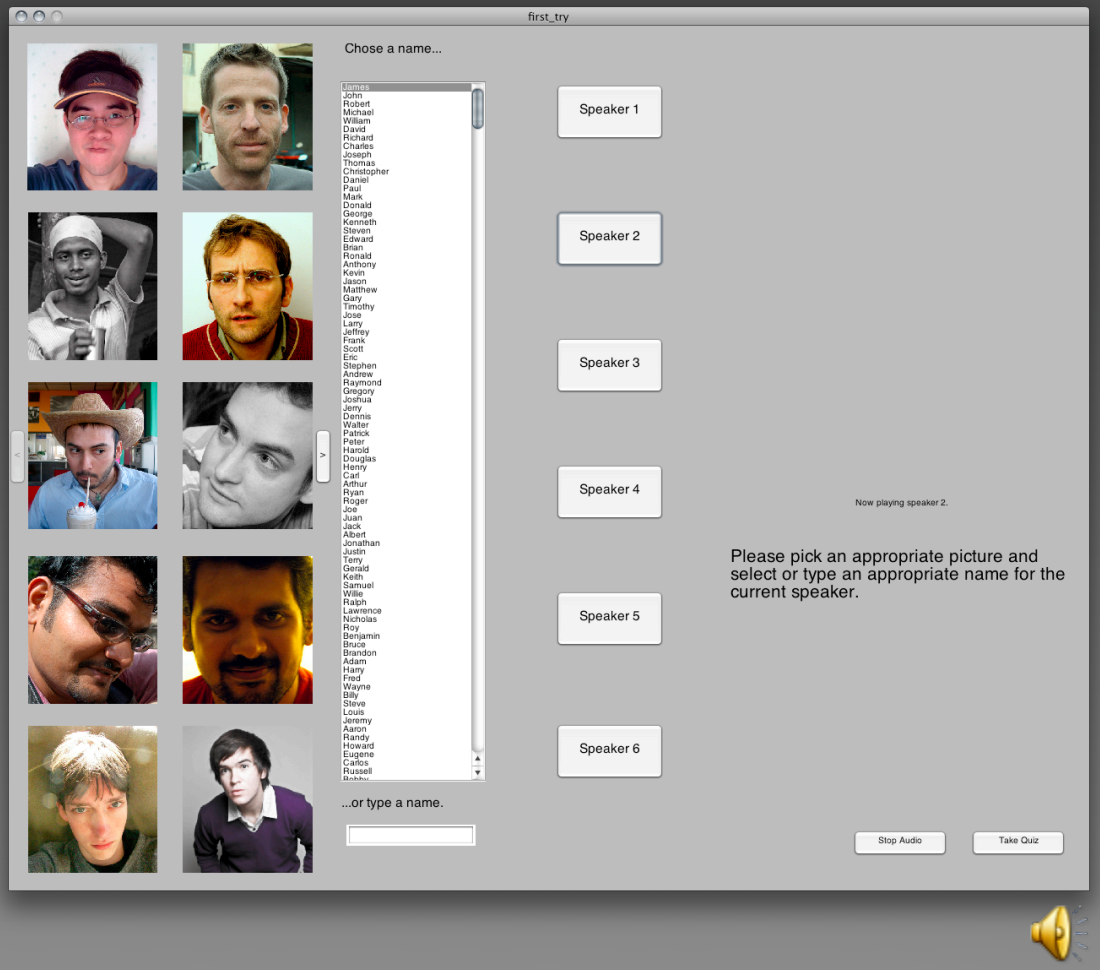
SID: Experiment Design



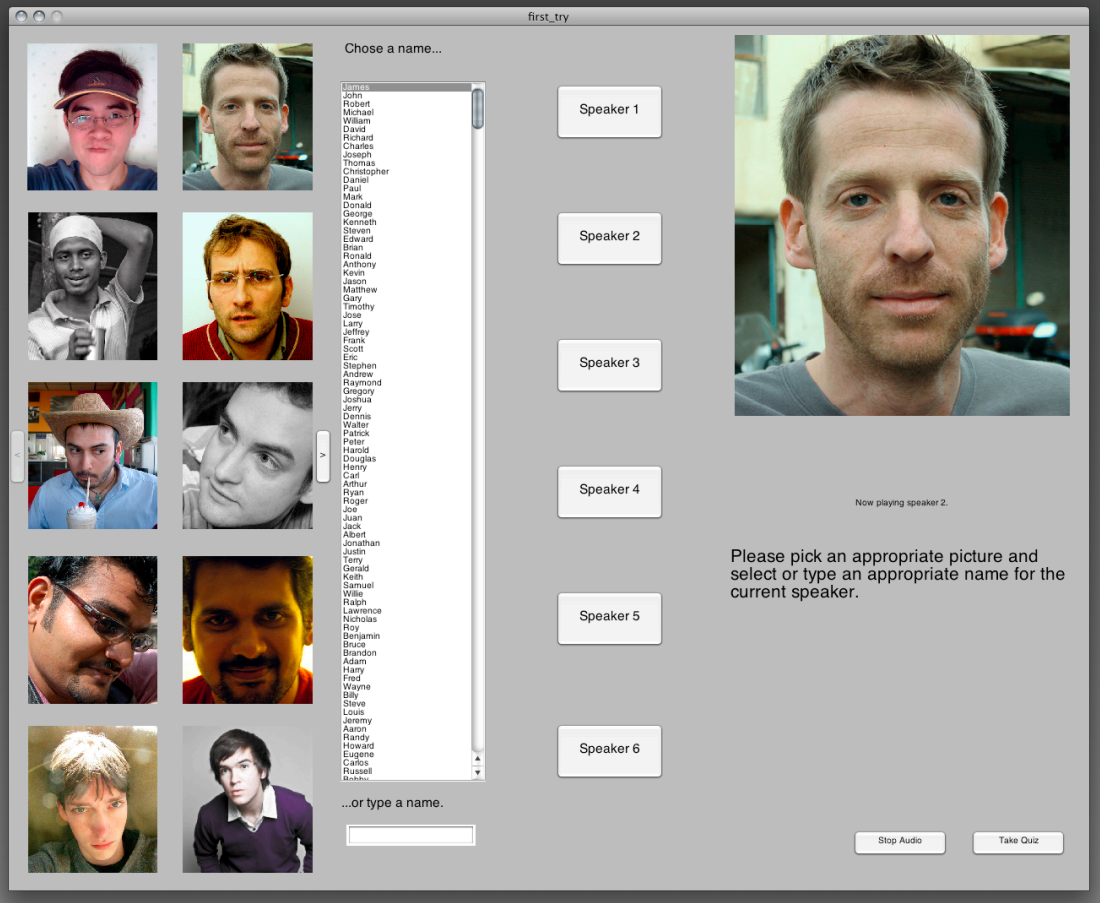
SID: Experiment Design



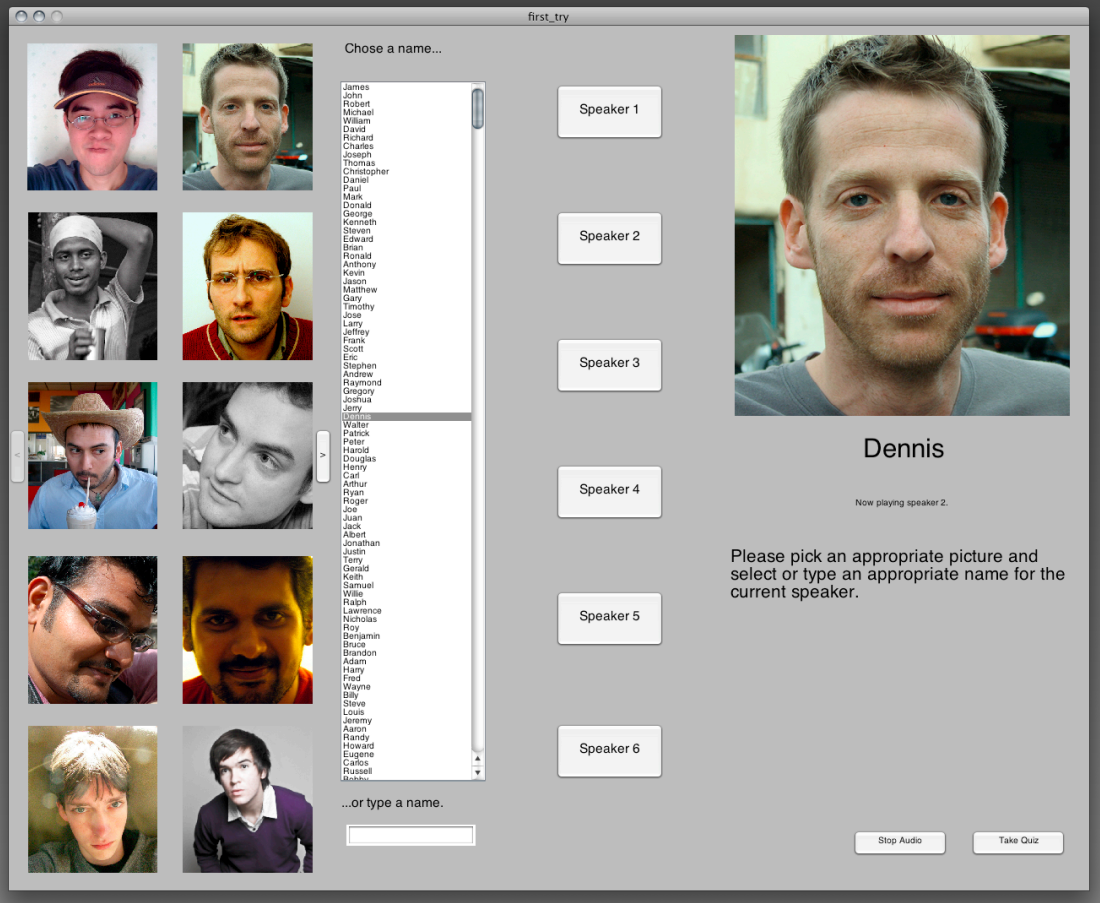
SID: Experiment Design



SID: Experiment Design



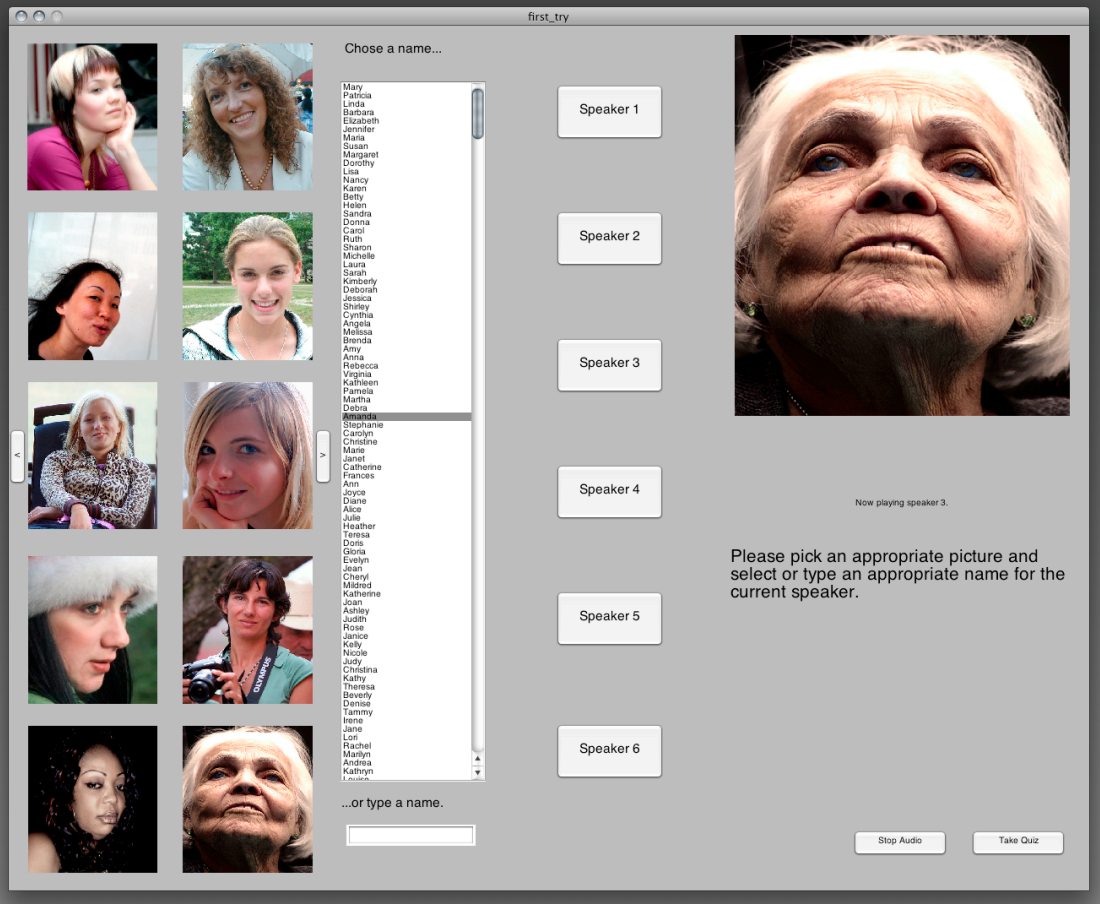
SID: Experiment Design



SID: Experiment Design

SID: Experiment Design

SID: Experiment Design



SID: Experiment Design