Nemours Children’s Health System

- **Clinical**
  - Hospitals in Wilmington & Orlando
  - Clinics & satellites in DE, FL, PA, NJ, MD, GA
  - Over 700 physicians
  - Over 30 pediatric specialties
  - Over 1.2 million encounters per year
  - Around 300,000 unique patients seen per year

- **Research**
  - Almost $9M in NIH funding for 2014
  - 11 Research Centers
  - ~40 Laboratories
CPASS

- **Four Labs + Bioinformatics**
  - Auditory Physiology & Psycho-acoustics – Head, Morlet
  - Balance & Vestibular Disorders – Head, O’Reilly
  - Craniofacial Outcomes Research – Head, Vallino
  - Speech Research Lab – Head, Bunnell
    - Acoustic Phonetics
    - Speech Perception/Production
    - Clinical Speech Technology
      - Applications involving speech recognition & synthesis technology
  - Bioinformatics – Head, Bunnell
Clinical Speech Technology

- **Utterance Verification / Classification**
  - Acoustic Phenotyping
  - Used as functional hearing evaluation
  - Auditory/verbal therapy
  - Used as objective speech intelligibility measure
Speech Synthesis for Assistive Technology

- **Problem – Speech Generating devices are:**
  - Limited in choice of voices
  - Impersonal
  - Lacking expressiveness

- **Existing Solutions**
  - Voice banking
  - Voice conversion/creation
  - Parametric synthesis to modulate prosodic/expressive features
This is a demonstration of the ModelTalker Speech Synthesis System.
Going forward for personal voices...

**User Needs**
- More Expressive!
  - Lacking in Unit Selection without massive amounts of data
  - Not modeled well in statistical parametric synthesis
- More Natural
  - Issue particularly for parametric synthesis
  - Fewer ‘glitches’ in unit selection
- Lower Barriers to creation
  - Fewer hours of recording
  - Improved morphing

**Research Needs**
- Phonetics/Phonology
  - Improve acoustic models of emotion and expressiveness
  - Improve models of the time-varying structure of speech (!)
- Engineering
  - Improvements in signal processing to model voice/vocal-tract interaction
- Capture the features of an individual’s speech in a small number of dimensions than can be manipulated in expressively useful ways.