

CLASR: Cross-Language Automatic Speech Recognition

John Henderson

781-271-2849 • jhndrsn@mitre.org

MSR

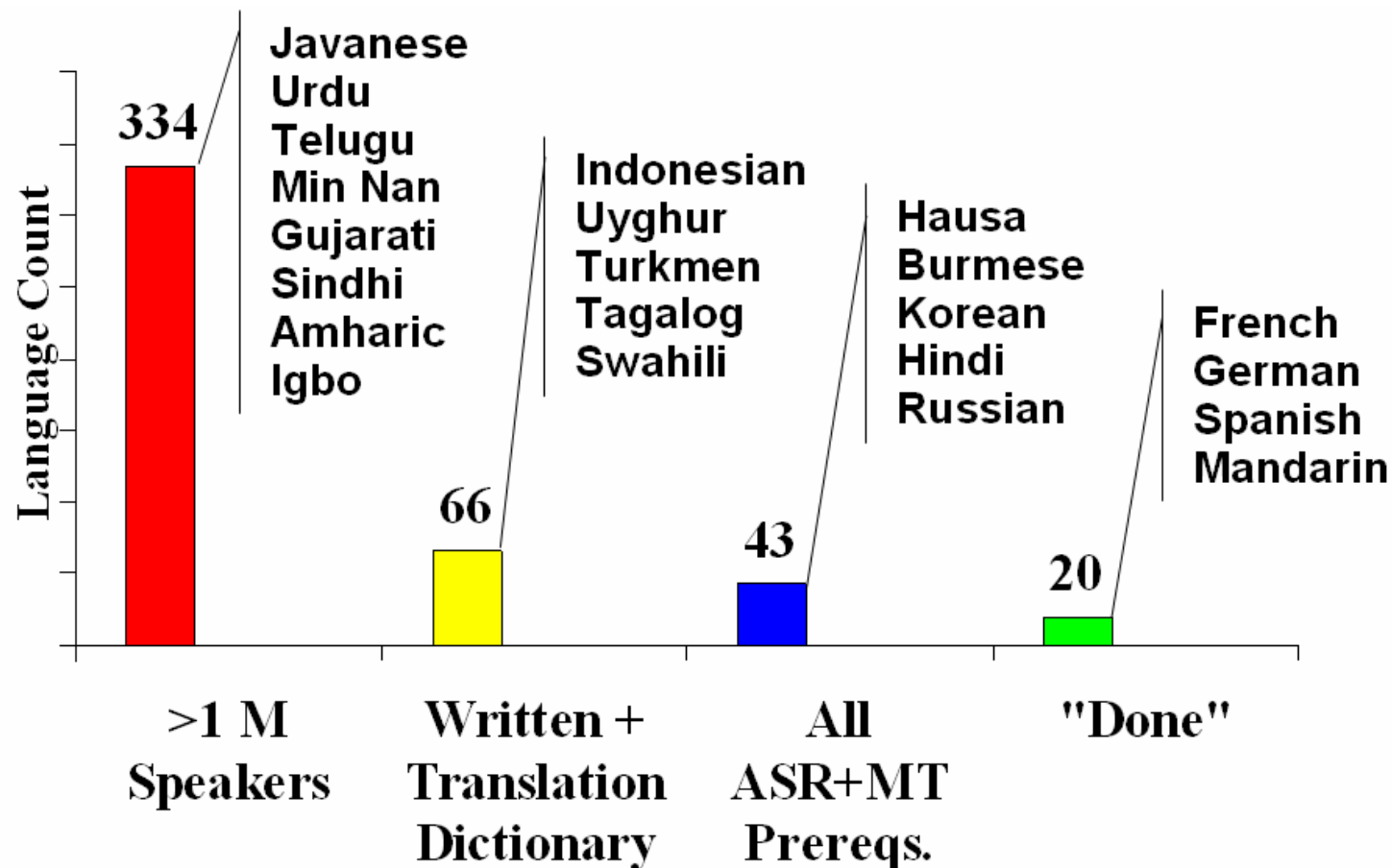
 MITRE
Technology
Program

Problem

- **More than 300 languages have a million speakers.**
- **Fewer than 70 languages have standardized written forms and extensive written texts.**
- **There are too many languages and too few linguistic resources to create enough speech translation systems.**

Background

Resources needed to develop speech recognition and machine translation systems are not common for rare, high-value languages.



Objective

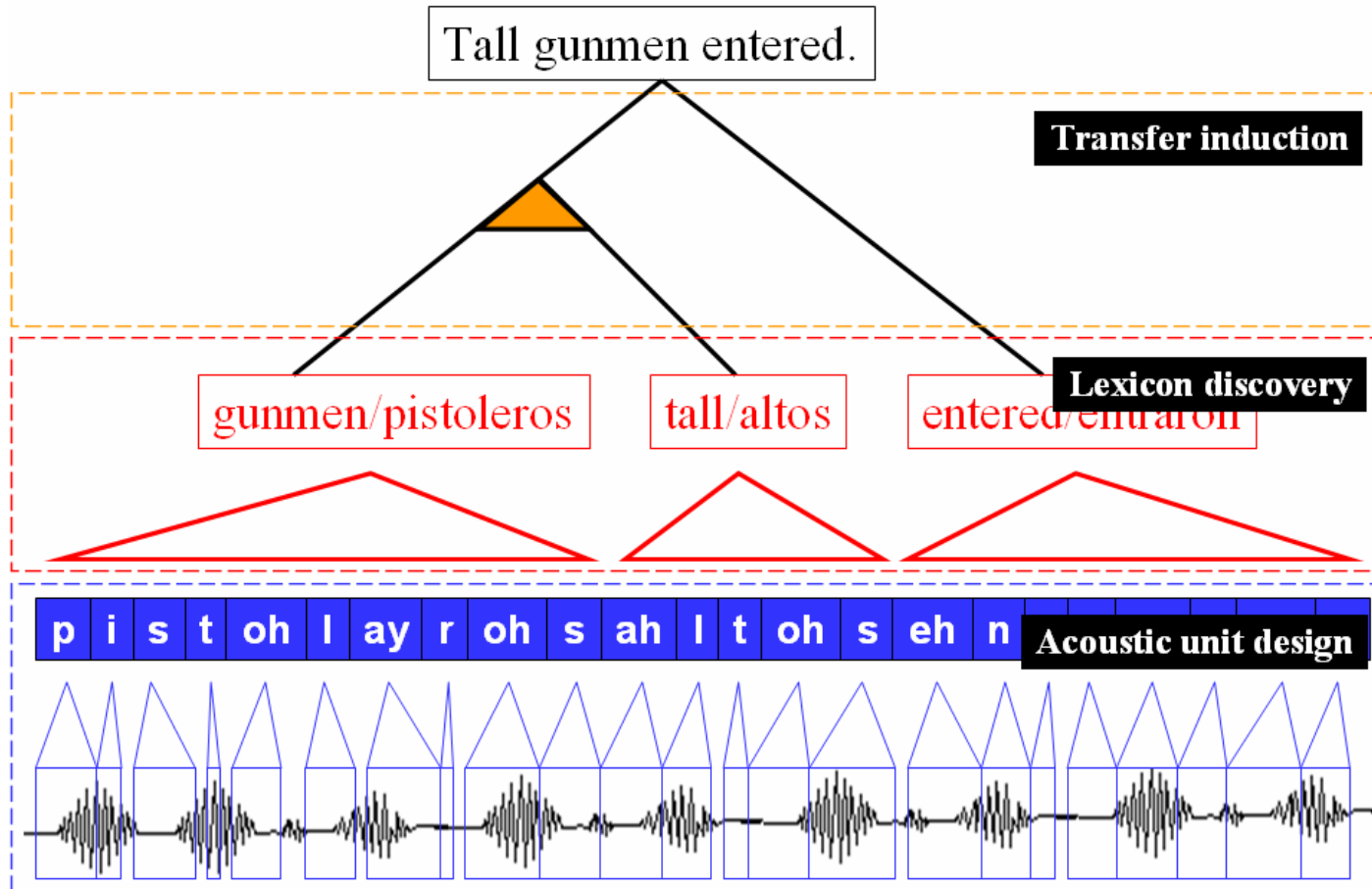
- **We are developing a system for spoken language translation of languages that lack significant written resources.**
 - **Acoustic units must be induced.**
 - **Bilingual pronunciation dictionaries must be derived.**
 - **Word reordering must be learned.**

Activities

- **Year 1: Preliminaries**
- **Year 2: Prototype of novel joint model system**
 - **Alignment of multilingual audio and text**
 - **Baseline systems**
 - **First single-stage CLASR prototype**
- **Year 3: Incorporate best available MT and ASR models into prototype**

Highlight

The speech translation process is composed of multiple modeling layers.



Demonstration

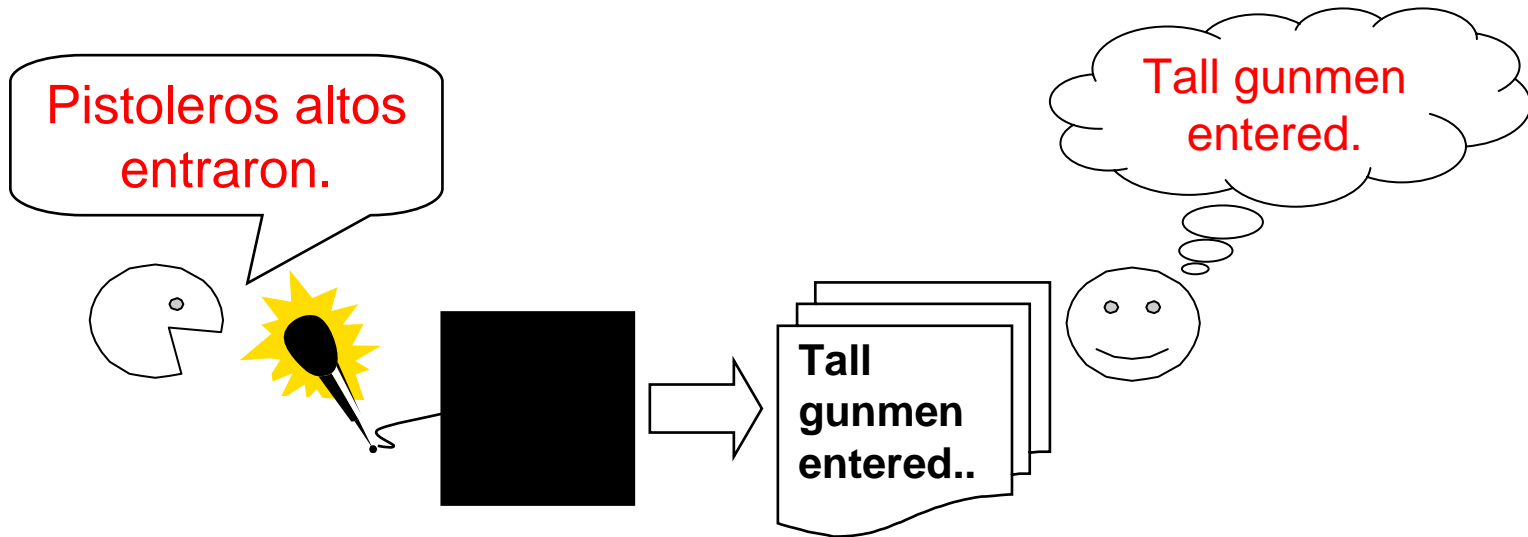
Mandarin dataset alignment from foreign audio to English text and sample output.

Text Transcript	江泽民首先对卢兹参议员近年为对中美关系的改善特别是在无条件延长对华最惠国待遇方面做出的努力表示赞赏
Acoustic	j i a n g z e m i n s h o u x i a n d u i l u z i c a n y i y u a n j i n n i a n w e i d u i z h o n g m e i g u a n x i d e g a i s h a n t e b i e s h i z a i w u t i a o j i a n y a n c h a n g d u i h u a z u i h u i g u o d a i y u f a n g m i a n z u o c h u d e n u l i b i a o s h i z a n s h a n g
Sample output	jiang zemin first of all for the improvement of the conditions for 对华最惠 countries on the efforts made appreciate that the status of sino-us relations are in particular the senator to 卢兹 years
Reference	First jiang zemin expressed appreciation for efforts made by senator lutz in recent years to improve the sino-us relations especially the unconditional extension of china's most-favored nation trade status

Impacts

- **Universal Access** to massive audio collections
 - Support audio triage: coarse, rough characterization of foreign language audio
 - Enable global infectious disease monitoring
- **New approach** to an old problem
 - Joint model of ASR & MT is closer to the real optimization problem for many scenarios
 - Results will widen the bridge between ASR & MT research communities

Future Plans



**The goal.
For all foreign languages.**