

## Automatic Speech Recognition

What we have accomplished?

- Built an automatic large vocabulary continuous speech recognition (LVCSR) system to convert speech to text on-the-fly. LVCSR system is based on context-dependent hidden Markov model.

Why is it useful?

- Useful for automatic annotation of multimedia data consisting of audio only or audio-visual information. Overcomes manual annotation procedures which are tedious and expensive

What is our corpus?

- Our LVCSR system was trained and tested on news broadcasts in Illinois covering a wide range of topics such as politics, weather, education, sports, obituary etc.
- Speaking style of news readers somewhere between read versus spontaneous

Item	Train	Test
Audio Data Size	117 min	25 min
Total Words	22051	4613
Unique Words	3925	1444
OOV (out-of-vocab) words	320	85
Unique OOV words	180	45
OOVs per 100 words	1.45	1.84
OOVs per min. of audio	2.74	3.40

Performance of LVCSR system:

Total Words Tested	Words incorrectly detected by LVCSR	Word Error Rate	Insertions	Deletions	Substitutions
4613	1983	43 %	409	230	1344

Examples of True vs LVCSR hypothesis:

- True: *!SIL MEMORIALS MAY BE MADE TO THE G. C. M. S. EDUCATION FOUNDATION OUR LADY OF LOURDES CATHOLIC CHURCH OR WESTMINSTER VILLAGE IN BLOOMINGTON !SIL*

LVCSR: *!SIL MEMORIALS MAY BE MADE TO THE **SHE SEE** M. S. **AT TO KITCHEN** FOUNDATION OUR LADY OF LOURDES CATHOLIC CHURCH OR **WEST CONCERN** VILLAGE IN BLOOMINGTON !SIL*

- True: *!SIL EVERYONE IN THE WORKFORCE UNION AND NON UNION ENJOYS THE FRUIT OF WHAT ORGANIZED LABOR HAS FOUGHT FOR AND WON OVER THE YEARS !SIL AS J. F. K. SAID A RISING TIDE LIFTS ALL BOATS !SIL*

LVCSR: *!SIL EVERYONE IN THE WORK FORCE UNION **ON** UNION **IN TO WASTE AFRAID** OF WHAT ORGANIZED LABOR HAS **FIVE** FOR **AN ONE OF** THE YEARS AS **STAFF KAYE SAD OF RAISING TIRED THAT SOLD LET'S** !SIL*