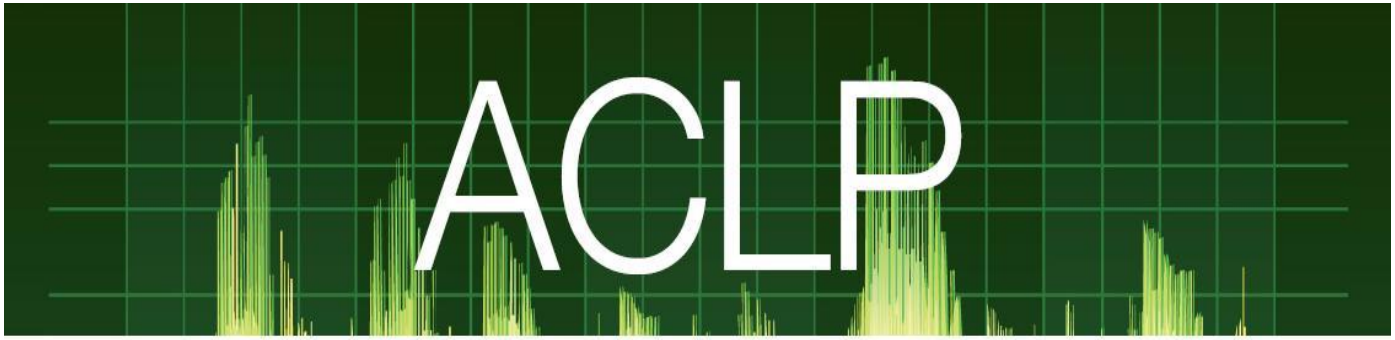
Creating Word Hypotheses Using Anchors



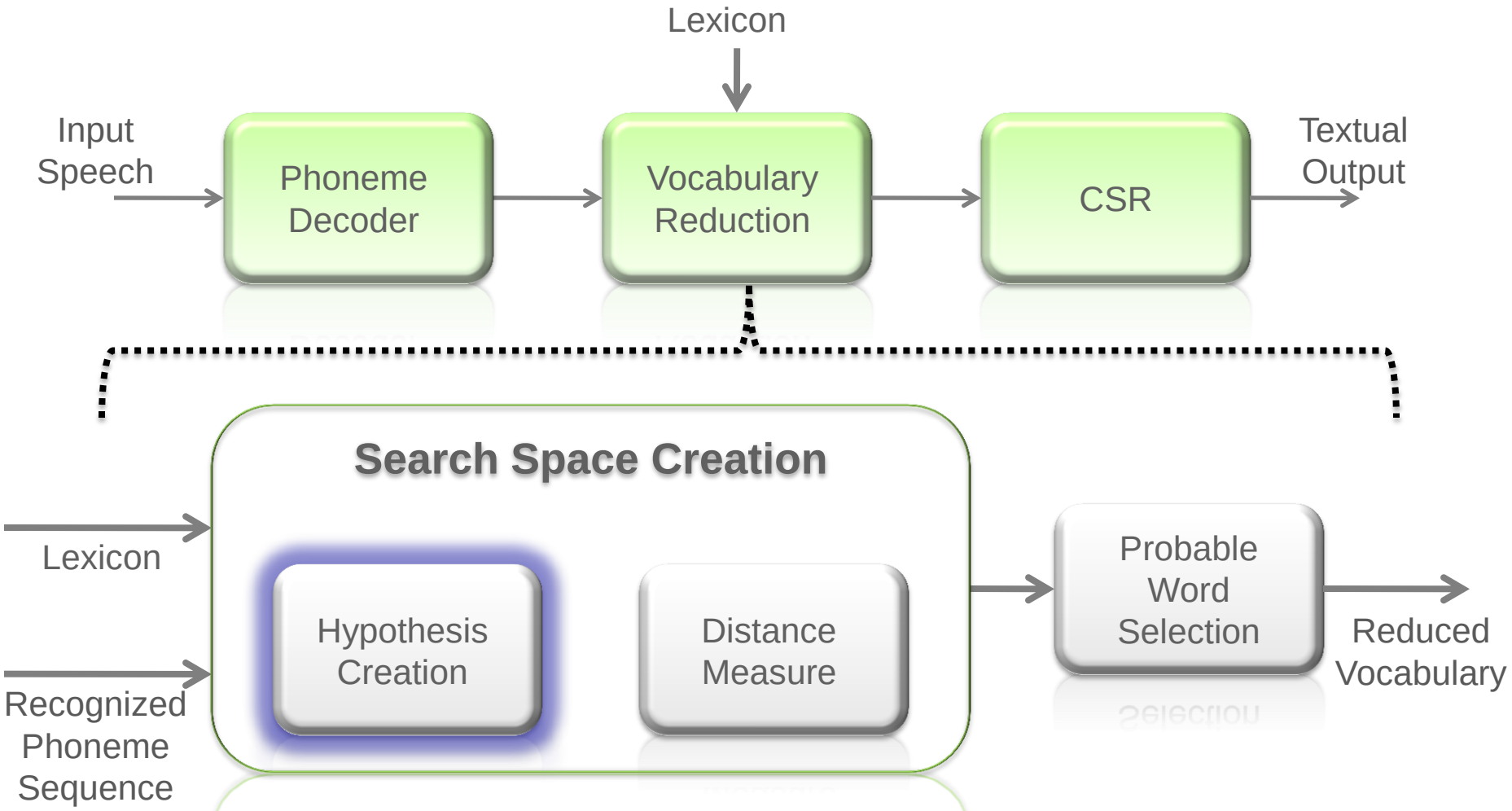
Afeka Center for Language Processing (ACLP)

Ella Tetariy

Speech Recognition Day – 8/6/2010

AFEKA TEL-AVIV ACADEMIC COLLEGE OF ENGINEERING

Multi-Stage LVCSR



Overview

Problem:

- 100k LVCSR – Reduced Vocabulary Exhaustive Process
- Search Problem for >100k Lexicon
  - Increased computational complexity and memory requirements
  - Reduced recognition performance

Research Goal:

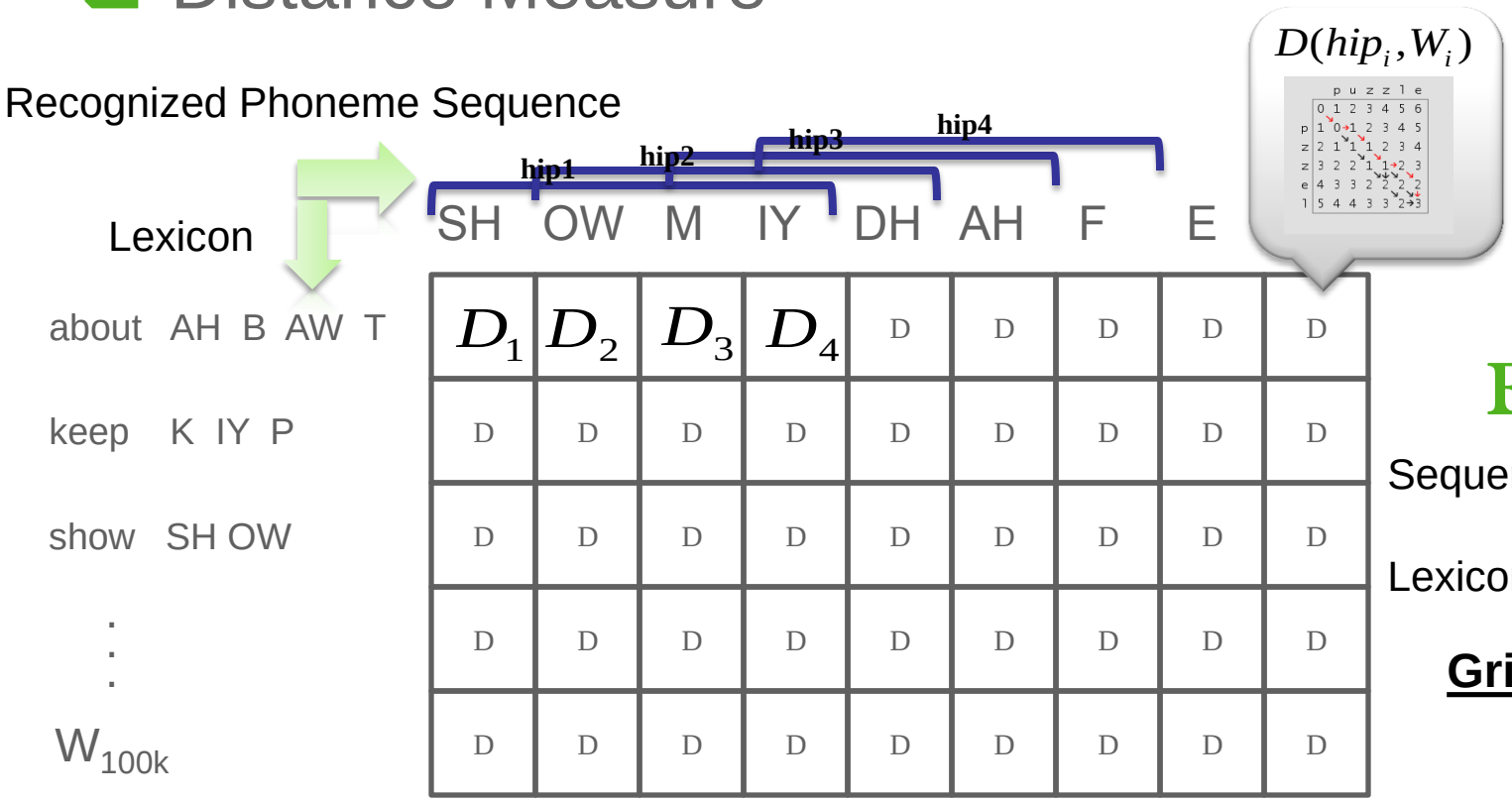
- Reduce Search Space - to support a large Lexicon
- Minimum Performances Loss

Direction:

- Anchor-Based Hypothesis Creation

Exhaustive Search Space Creation

- Hypothesis Creation
- Distance Measure



Example:

Sequence – 250 Phones
+20 Sec
Lexicon size = 100k Words

Grid size ≈ 25M

Testing DB

Authentic VM messages

# of Sentences	2485
Avg. Sentence Time (sec)	20+ sec
Words in Sentence	70-80
# Phones in Sentence	250
Lexicon Size	100k

Exhaustive Search Space:

Grid Size	Grid Creation Time (sec)

Anchor-Based Search

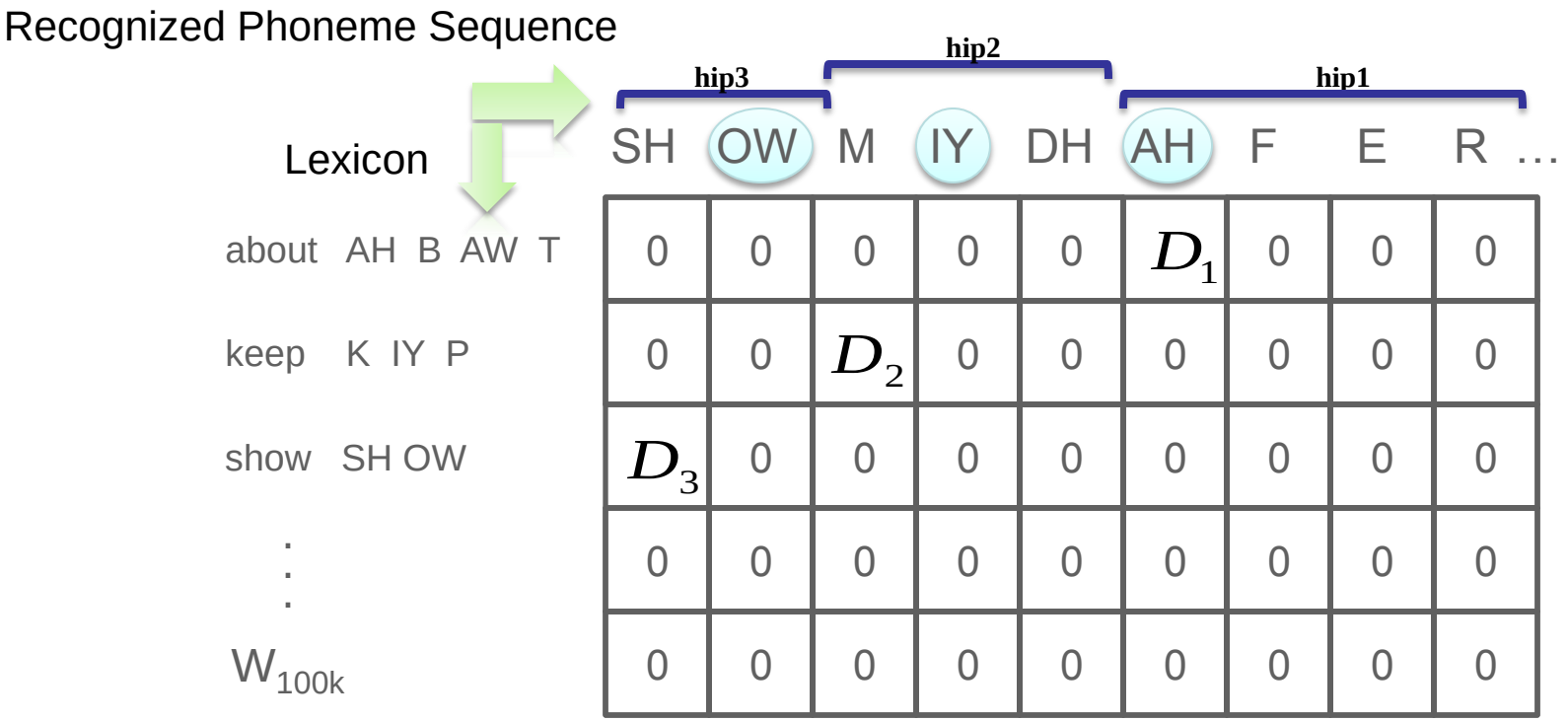
Phonemes with high confidence recognition results

- **Anchor-Based Search -** Smart search for reducing the search space by building hypotheses around anchor points

The goal is to develop an algorithm that will significantly reduce search perimeters with minimal effect on recognition performance

Word Hypothesis Creation
- Around Anchors

- Only anchor-based hypotheses will be evaluated

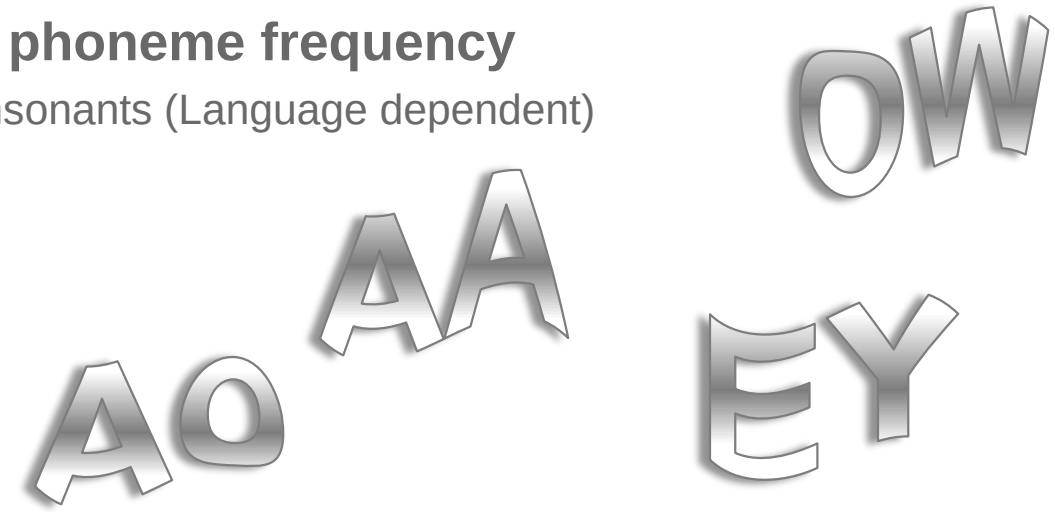


Possible Anchor Types

- Global Anchor Points
  - Linguistic Basis
  - Statistical Basis
- Dynamic Anchor Points
  - Input Sequence
  - Lexicon Word

Global Anchor Points
Linguistic Basis

- **Dependent on phoneme frequency**
  - Vowels vs. consonants (Language dependent)



ΔGrid Size	ΔGrid Creation Time (sec)	ΔCoverage

Global Anchor Points
Statistical Basis

- Based On Phoneme Recognition Engine Performance

(I,D,S,C)

ΔGrid Size	ΔGrid Creation Time (sec)	ΔCoverage

Dynamic Anchor Points
Best Anchor Per Partial Phoneme String

- Windowing Mechanism
- For Each Window - Anchor decision based on MAX(Quality Measure)

Phoneme Sequence:

ΔGrid Size	ΔGrid Creation Time (sec)	ΔCoverage

Dynamic Anchor Points
Best Anchor Per Lexicon Word

- **Anchor points are lexicon word specific - based on MAX (Quality Measure)**

Lexicon Word: slowly S L OW L IY

Phoneme Sequence:

SH OW M IY DH AH F E R S T F L AY T T UW L OW S AE …

ΔGrid Size	ΔGrid Creation Time (sec)	ΔCoverage

Results Summary

Exhaustive Search	Grid Size	Grid Creation Time (sec)
#Phones in Sentence - 250 Lexicon Size – 100k		

Anchor Type	ΔGrid Size	ΔGrid Creation Time (sec)	ΔCoverage
Global – Linguistic Basis			
Global – Statistical Basis			
Dynamic – Input Sequence			
Dynamic – Lexicon Word			

- 90% Search Reduction with 5% Coverage Loss

Summary

- Vocabulary Reduction
  - Exhaustive Search Space - Search Problem for >100k Lexicon
- Creating Word Hypotheses
  - Using Anchor Points
  - Possible Anchor Types
- Test Results
- Next Step