



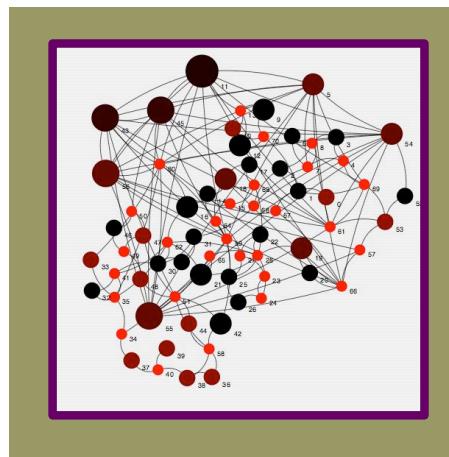
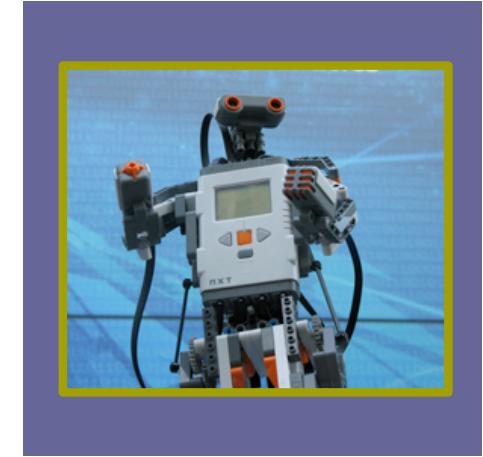
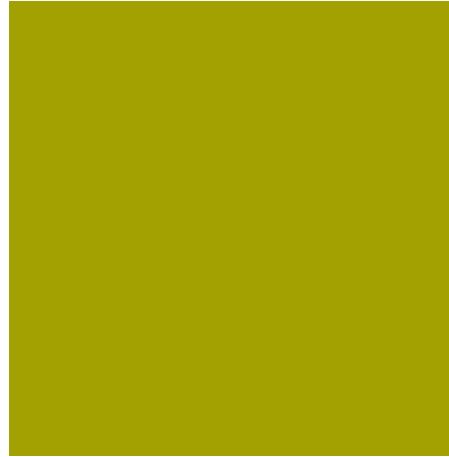
COMP40020

Human Language Technologies

'Twas brillig and the slithy
toves
April 2019



Prof. Julie Berndsen
School of Computer Science
Julie.Berndsen@ucd.ie



Contents:

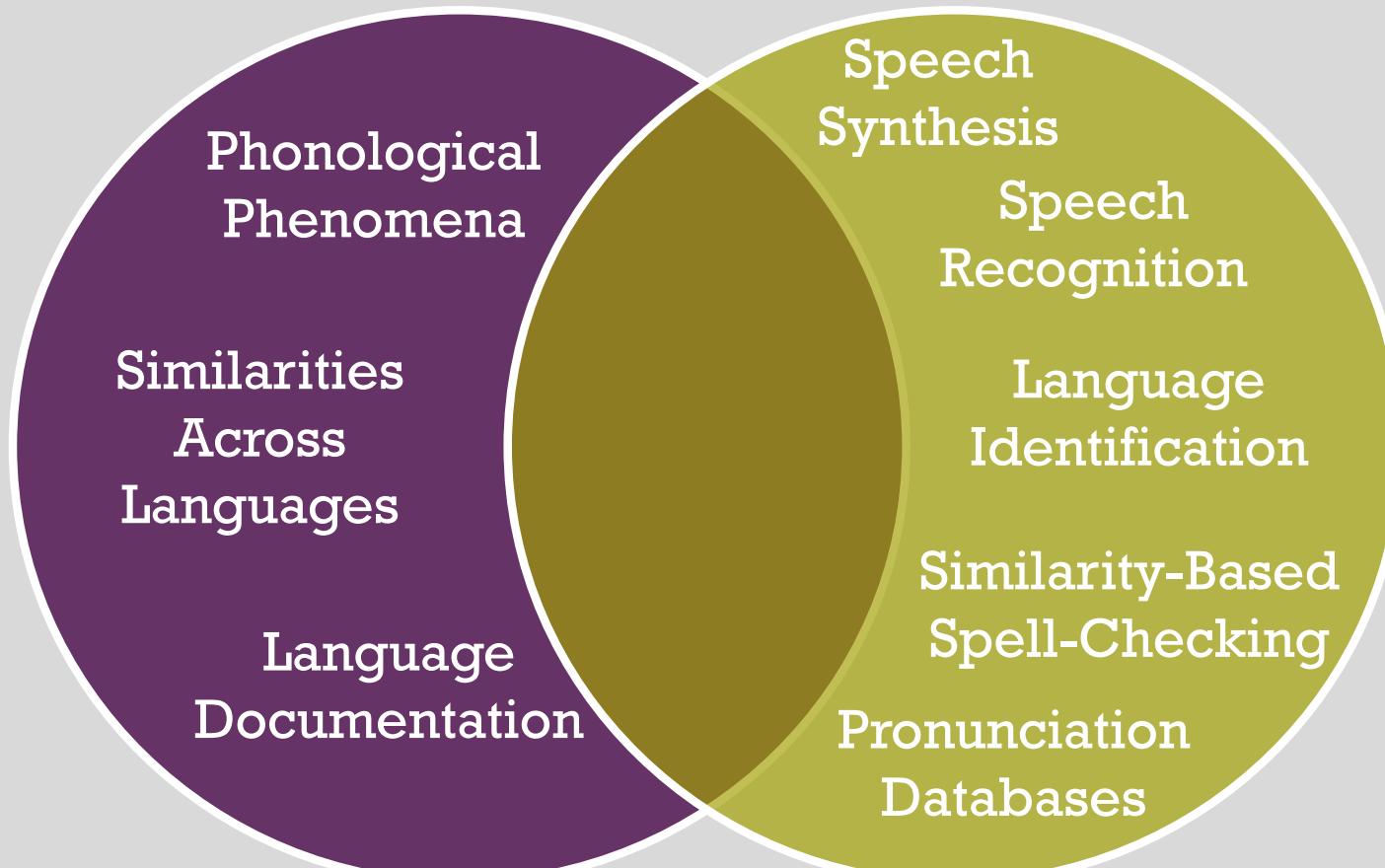
- Motivations for Computational Phonology Research
- Phonological Parsing
- Syllabification

Aim:

- To give an overview of knowledge-based and data-driven computational models of phonology and in particular the use of finite state techniques in phonological analysis.

+ Computational Phonology Research

HLT12



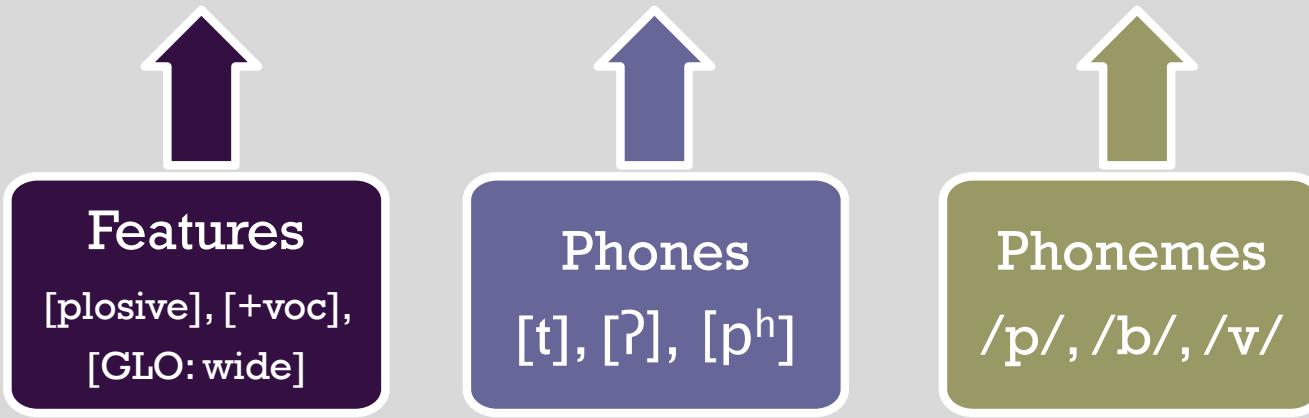
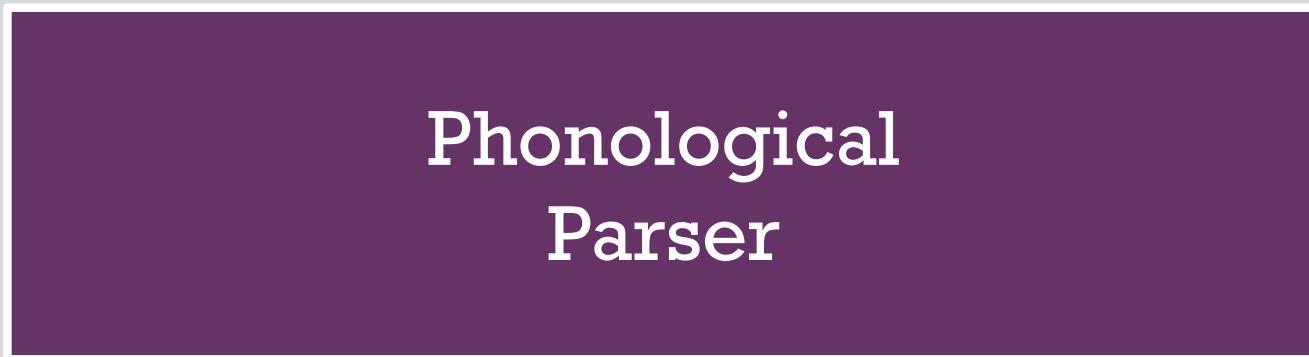
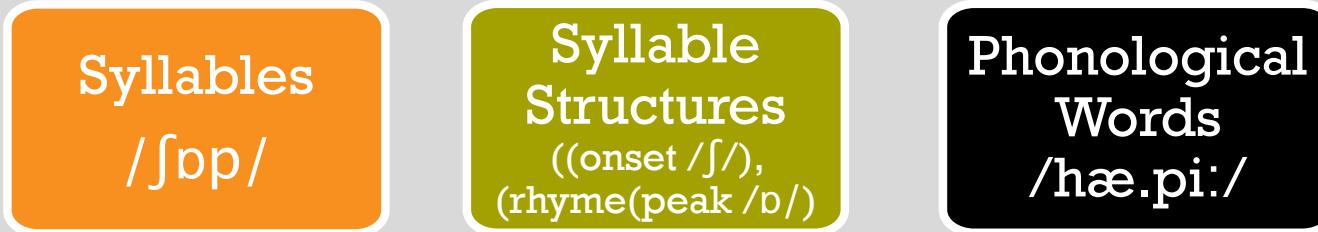
Linguistic Explanation

Technology Applications



+ Phonological Parsing

HLT12



+ Phonological Parsing

HLT12

Syllables
/ʃɒp/

Syllable Structures
((onset /ʃ/),
(rhyme(peak /ɒ/))

Phonological Words
/hæ.pi:/

Phonological Parser

Phonological “Grammar”

Features
[plosive], [+voc],
[GLO: wide]

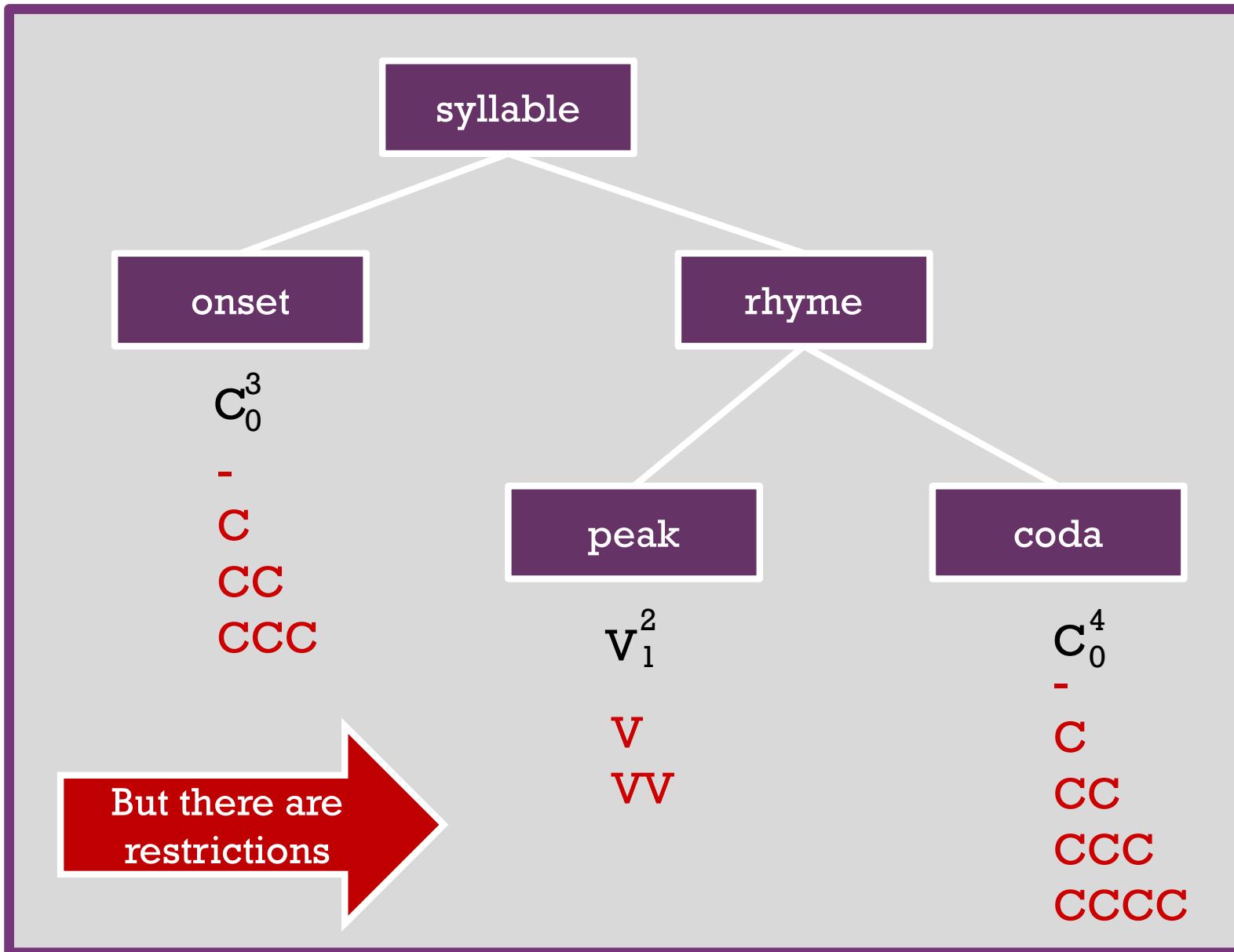
Phones
[t], [ɾ], [pʰ]

Phonemes
/p/, /b/, /v/



+ Recall: Syllable Structure

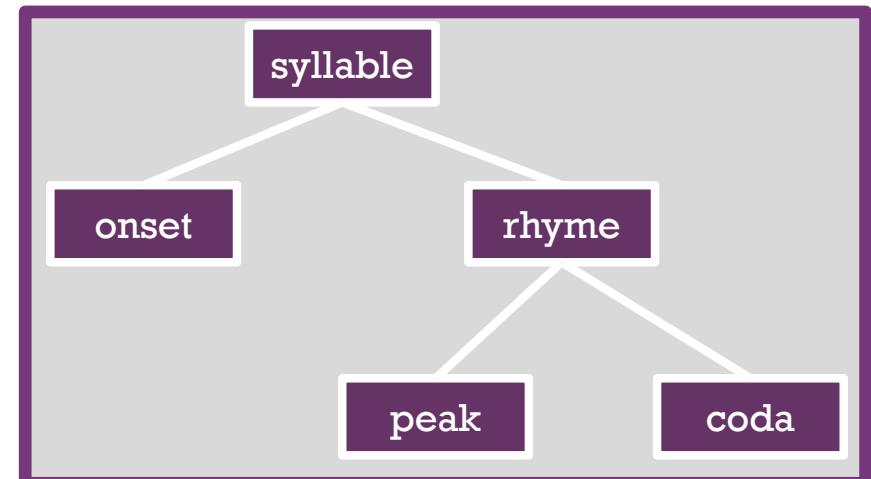
HLT12



+ Phonological Grammar

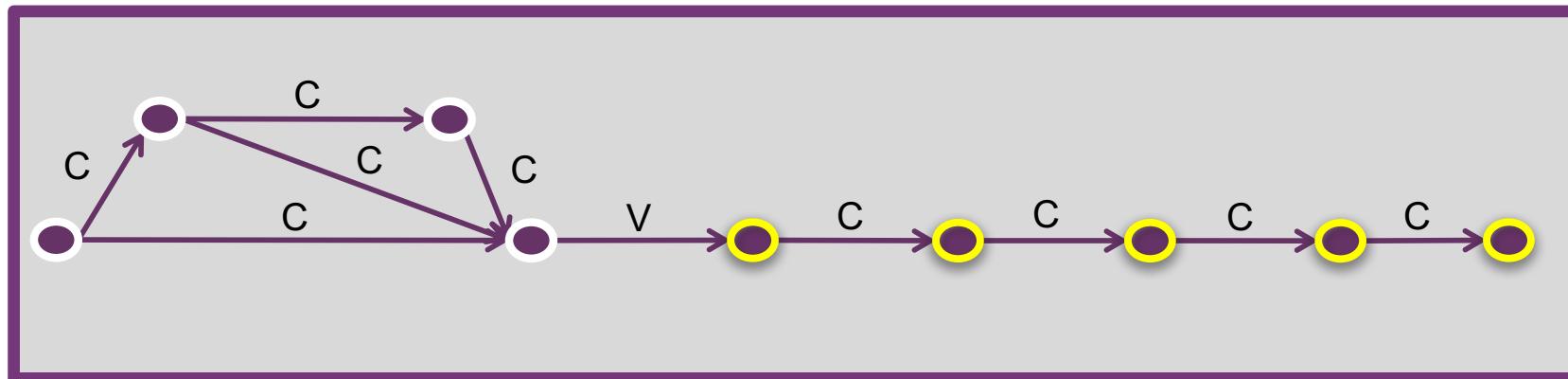
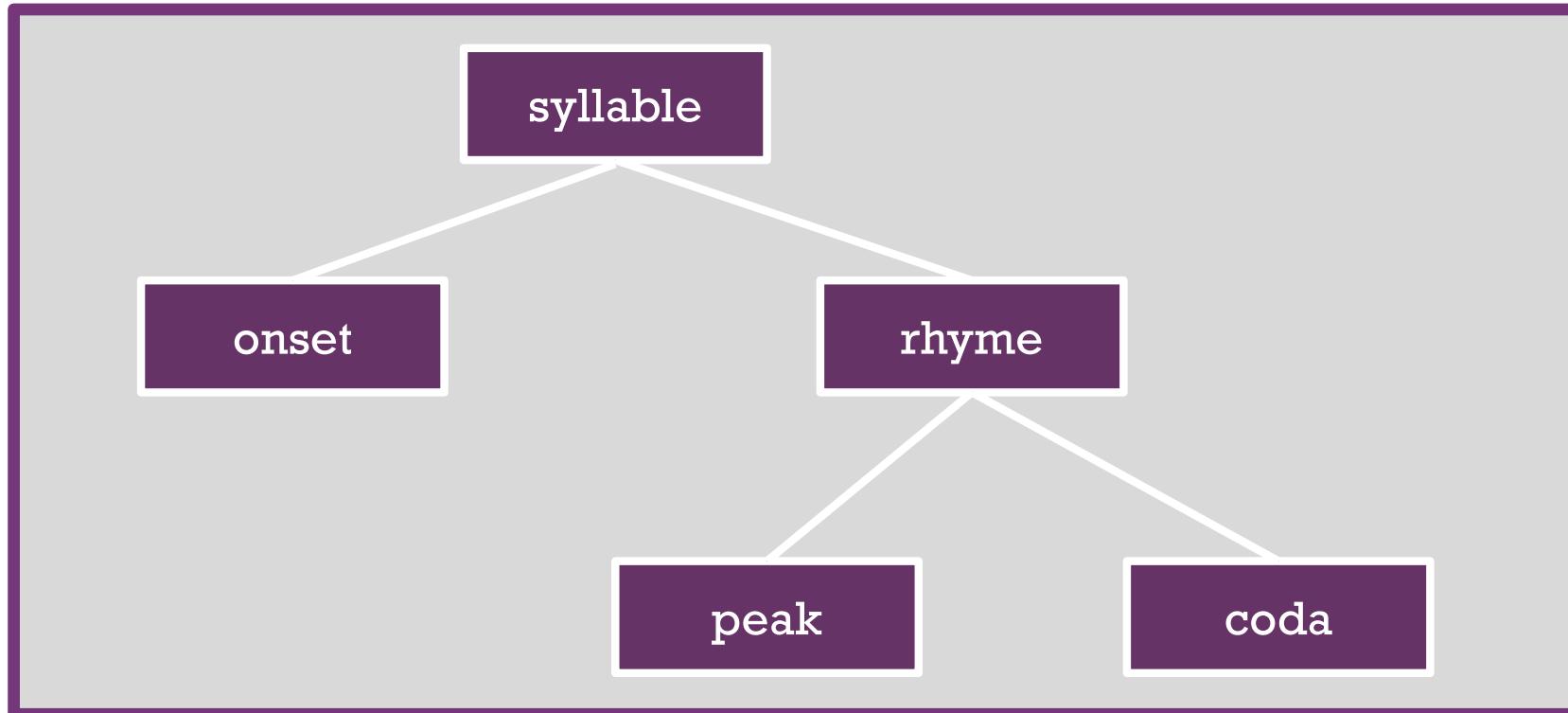
HLT12

- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → C C C | C C | C
- PEAK → V
- CODA → C C C C | C C C | C C | C
- C → p | t | k | b | d | g | f | v | θ | ð | s | z | ʃ | ʒ | h | j | w | l | r
| m | n | ɳ | tʃ | dʒ
- V → i | i | e | ɜ: | ə | ʊ | u: | ɔ: | æ | ʌ | ɒ | ɑ: | eɪ | ɔɪ | aɪ | eə |
ɪə | ʊə | əʊ | aʊ



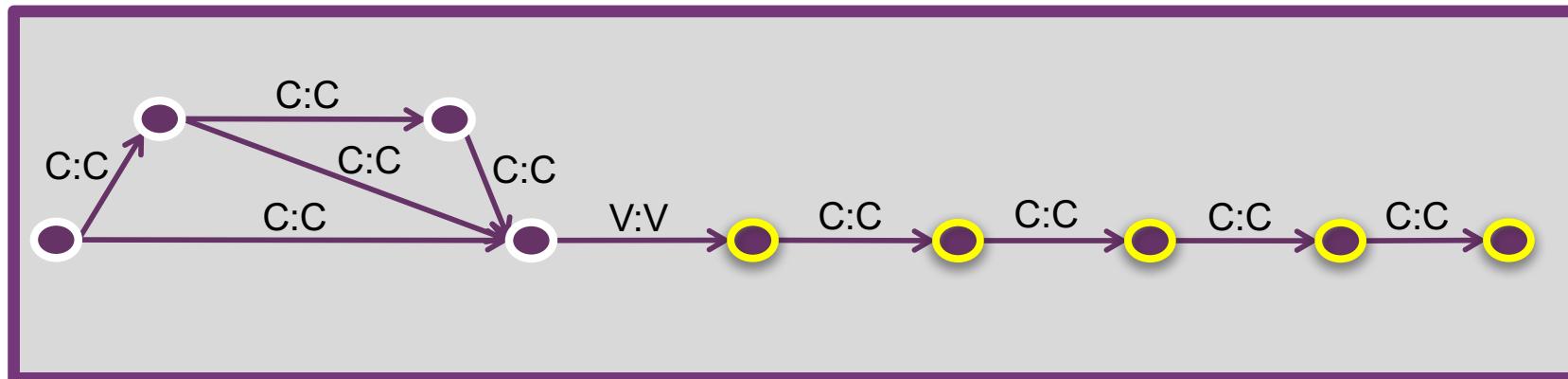
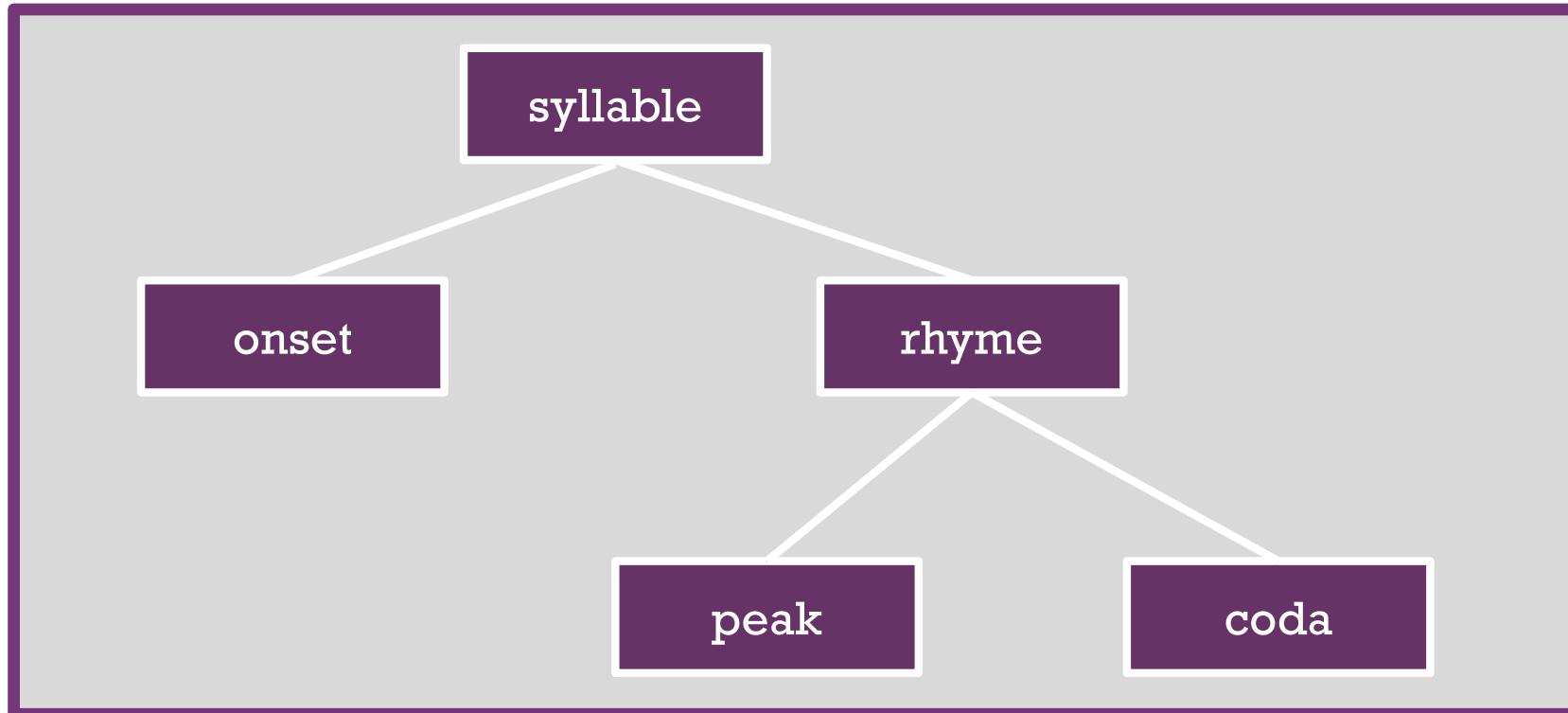
+CV-Automaton

HLT12



+CV-Transducer

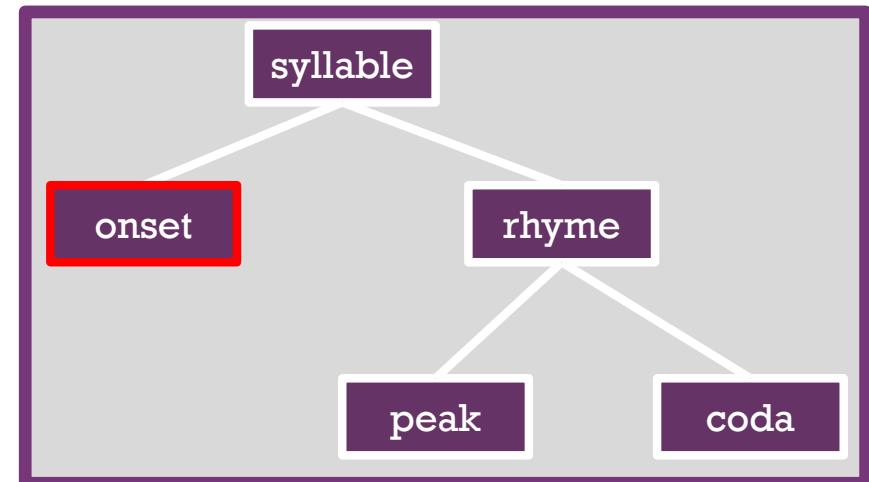
HLT12



+Onset Restrictions: Grammar

HLT12

- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → s C1 C2 | C C | C
- C1 → p | t | k
- C2 → l | r
- C → p | t | k | b | d | g | f | v | θ | ð | s | z | ʃ | ʒ | h | j | w | l | r
| m | n | ɳ | tʃ | dʒ
- V → i | i | e | ɜ: | ə | ʊ | u: | ɔ: | æ | ʌ | ɒ | ɑ: | eɪ | ɔɪ | aɪ | eə |
ɪə | ʊə | əʊ | aʊ



s p r

...



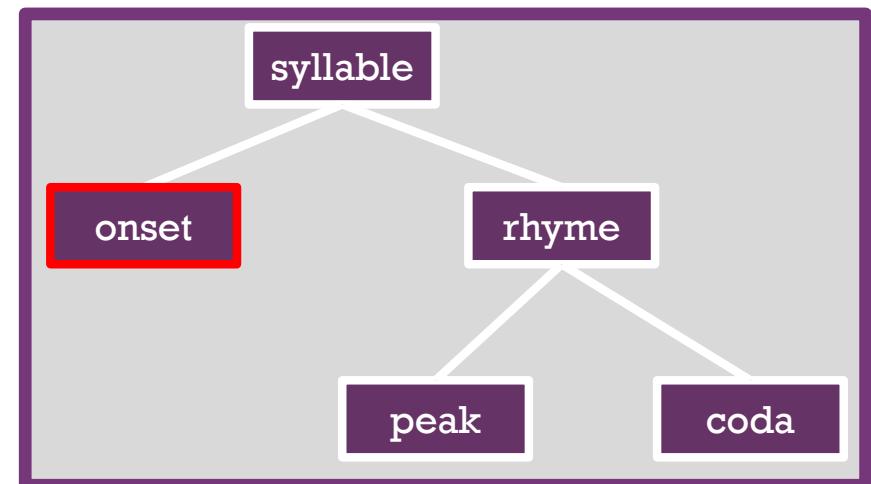
ONSET



+Onset Restrictions: Grammar

HLT12

- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → s C1 C2 | C C | C
- C1 → p | t | k
- C2 → l | r
- C → p | t | k | b | d | g | f | v | θ | ð | s | z | ʃ | ʒ | h | j | w | l | r
| m | n | ɳ | tʃ | dʒ
- V → i | i | e | ɜ: | ə | ʊ | u: | ɔ: | æ | ʌ | ɒ | ɑ: | eɪ | ɔɪ | aɪ | eə |
ɪə | ʊə | əʊ | aʊ



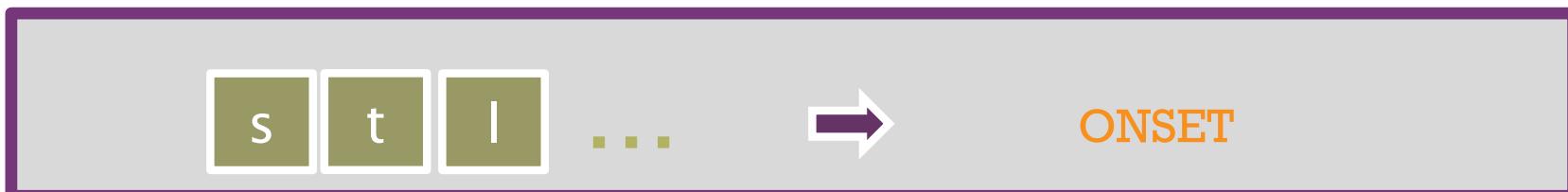
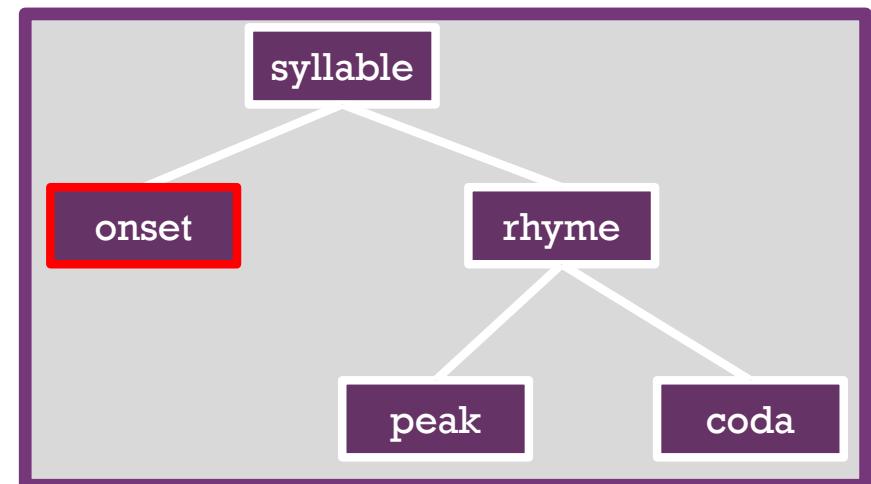
ONSET



+Onset Restrictions: Grammar

HLT12

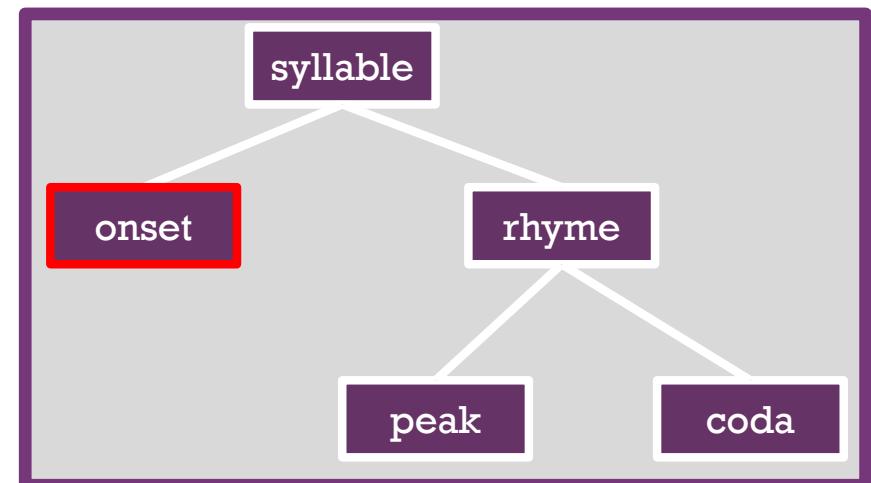
- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → s C1 C2 | C C | C
- C1 → p | t | k
- C2 → l | r
- C → p | t | k | b | d | g | f | v | θ | ð | s | z | ʃ | ʒ | h | j | w | l | r
| m | n | ɳ | tʃ | dʒ
- V → i | i | e | ɜ: | ə | ʊ | u: | ɔ: | æ | ʌ | ɒ | ɑ: | eɪ | ɔɪ | aɪ | eə |
ɪə | ʊə | əʊ | aʊ



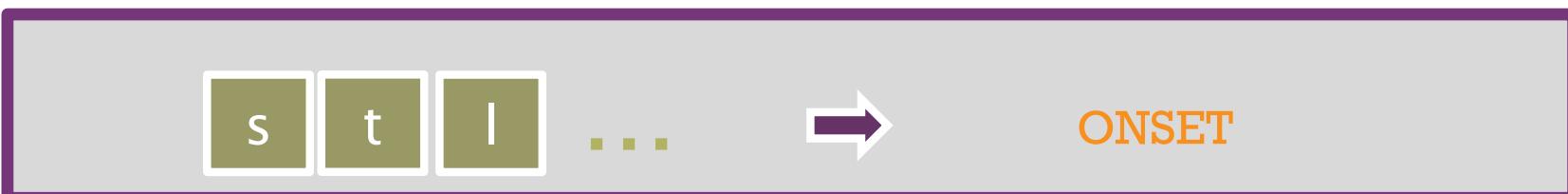
+Onset Restrictions: Grammar

HLT12

- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → s C1 C2 | C C | C
- C1 → p | t | k
- C2 → l | r
- C → BUT: /stl/ is not a possible onset in English
- V → i | i | e | ə | ʊ | u: | ɔ: | æ | ʌ | ɒ | ɑ: | ei | ɔɪ | ai | eə | ɪə | ʊə | əə | au



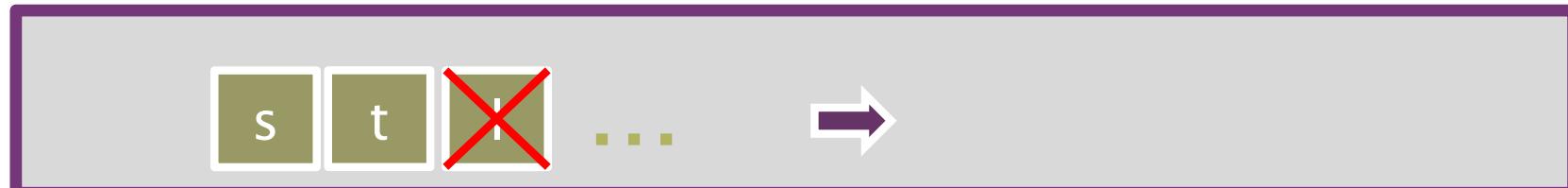
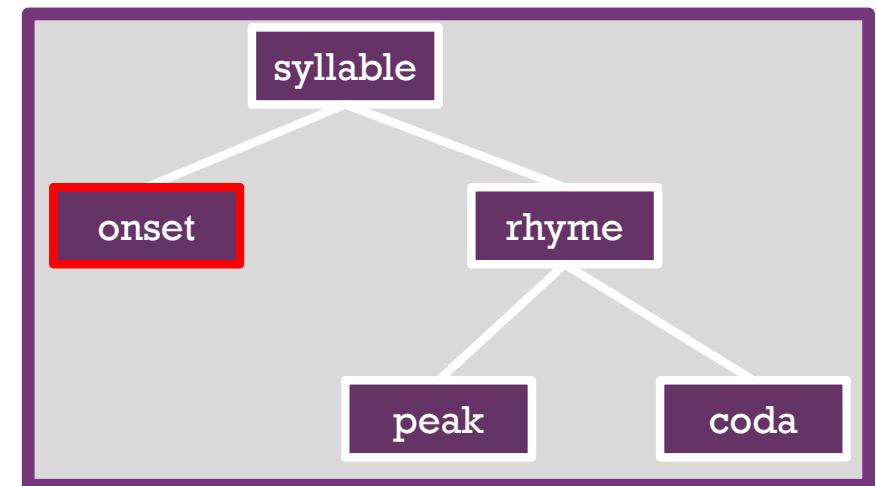
| θ | ð | s | z | ʃ | ʒ | h | j | w | l | r



+Onset Restrictions: Grammar

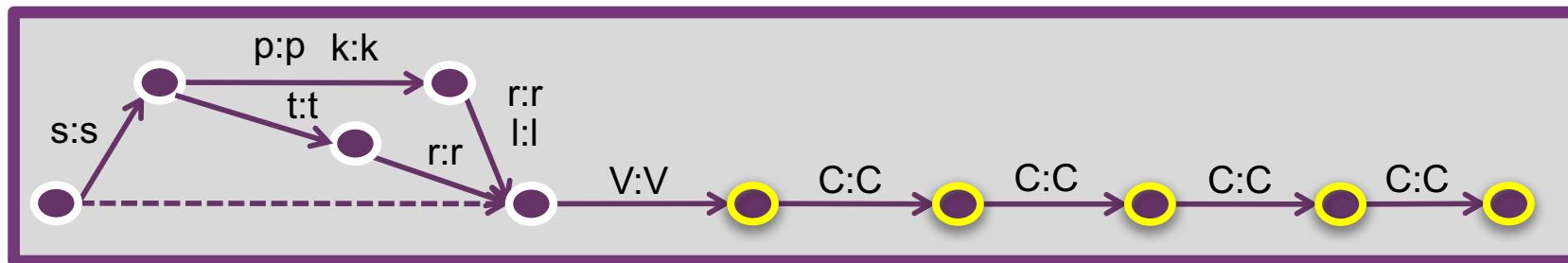
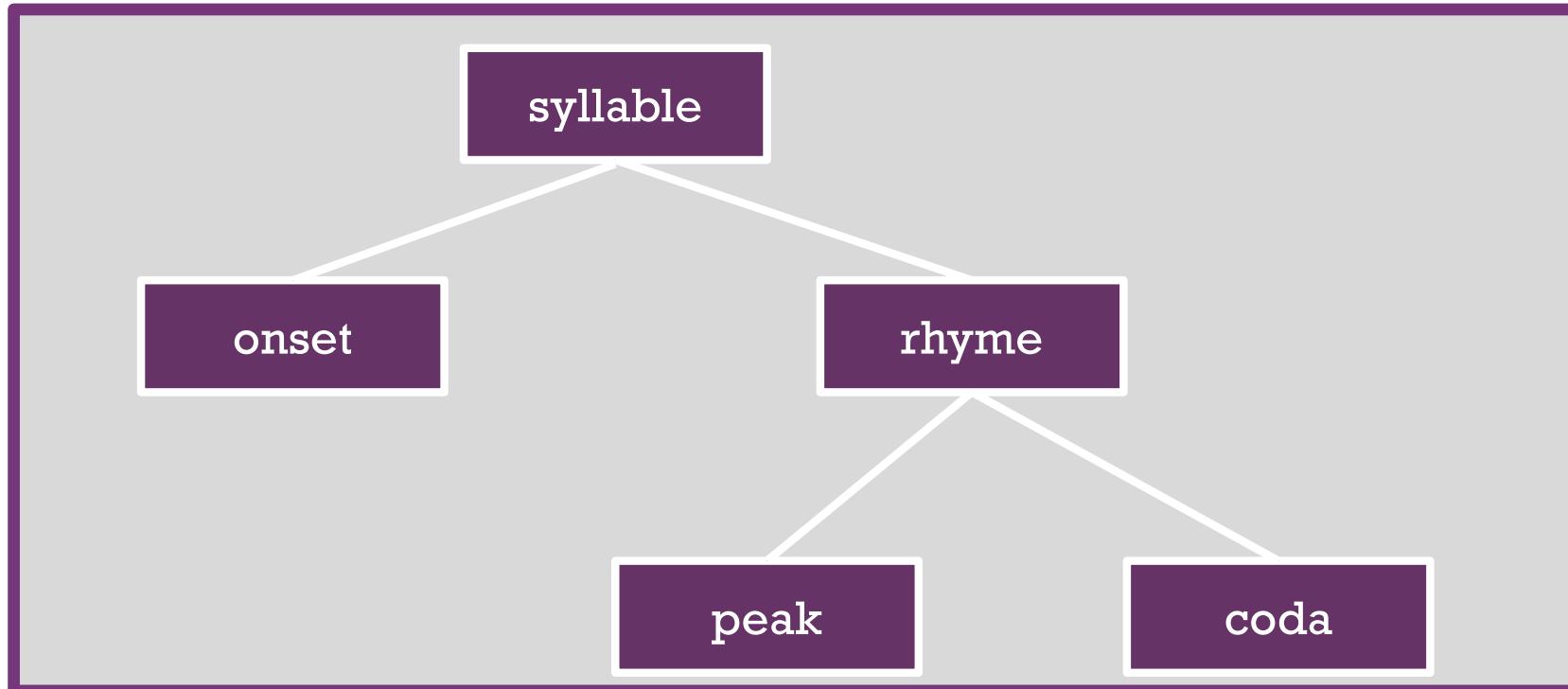
HLT12

- WORD → SYLLABLE*
- SYLLABLE → ONSET RHYME
- RYHME → PEAK CODA
- ONSET → s C1 C2 | s C1a C2a | ...
- C1 → p | k
- C1a → t
- C2 → l | r
- C2a → r



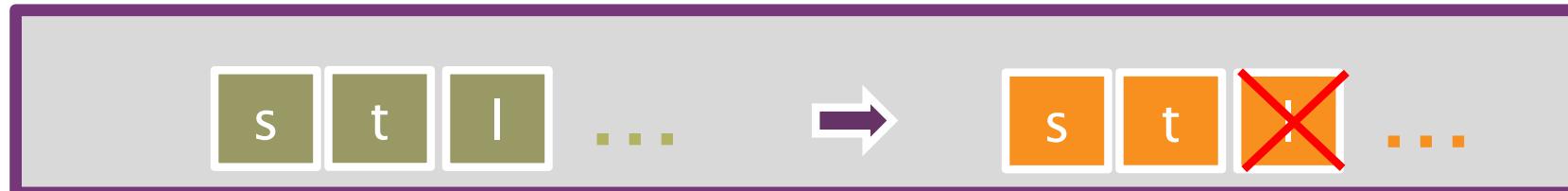
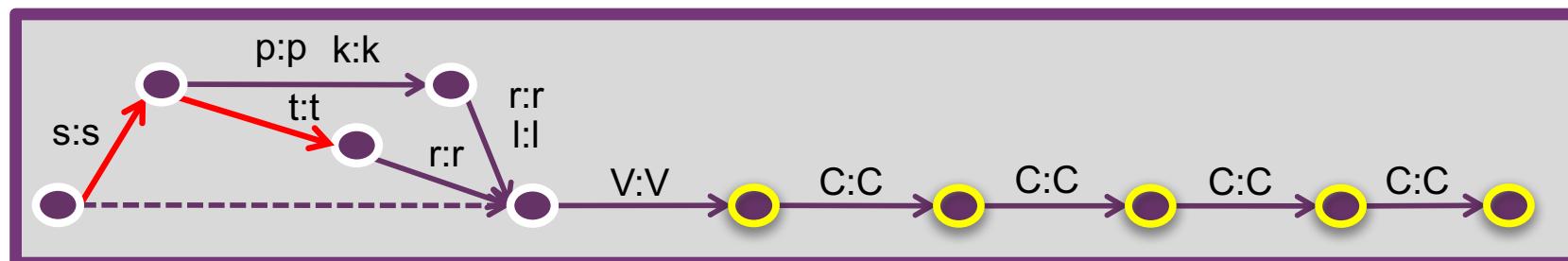
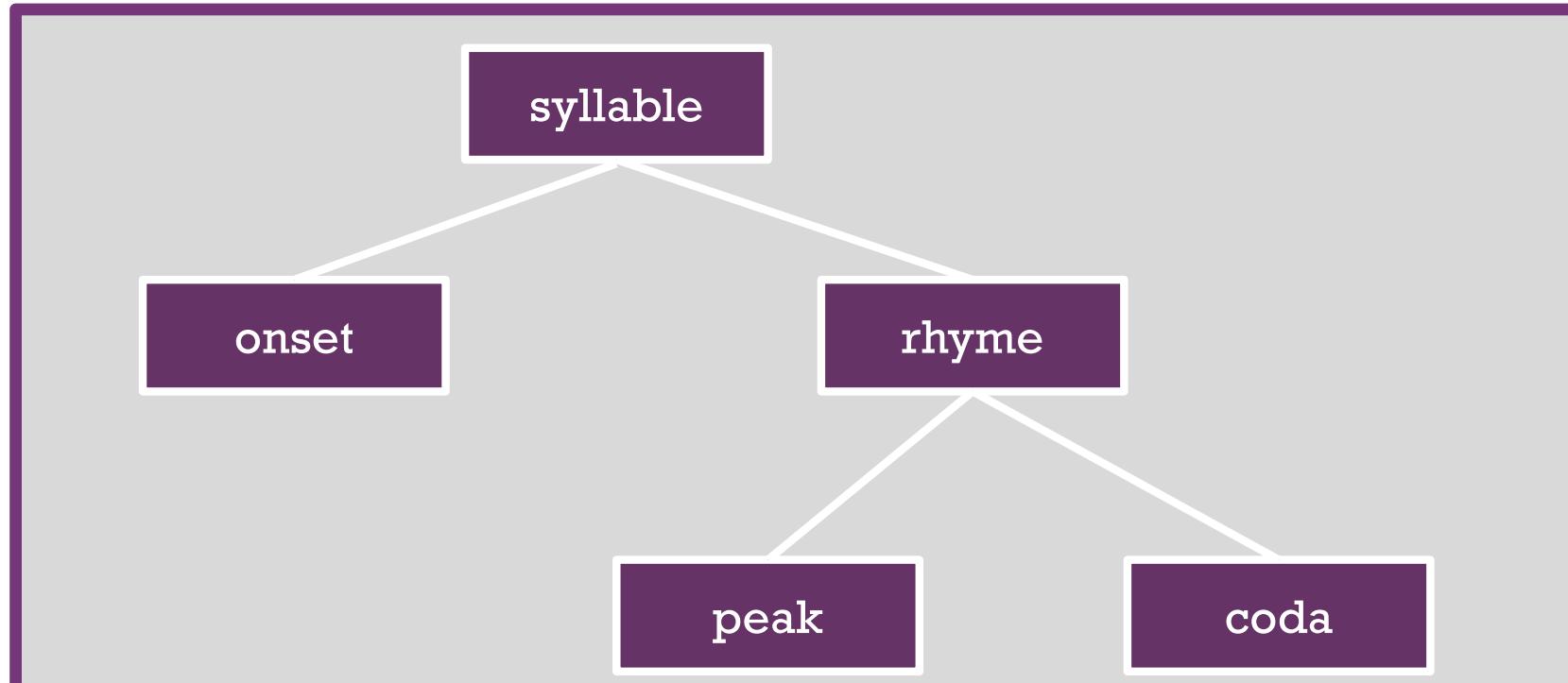
+Onset Restrictions: Transducer

HLT12



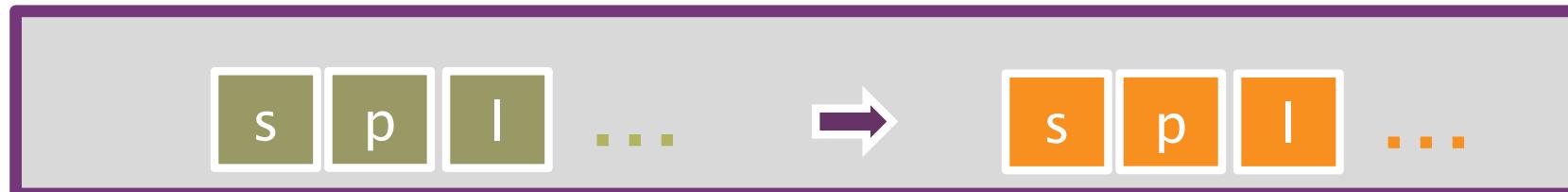
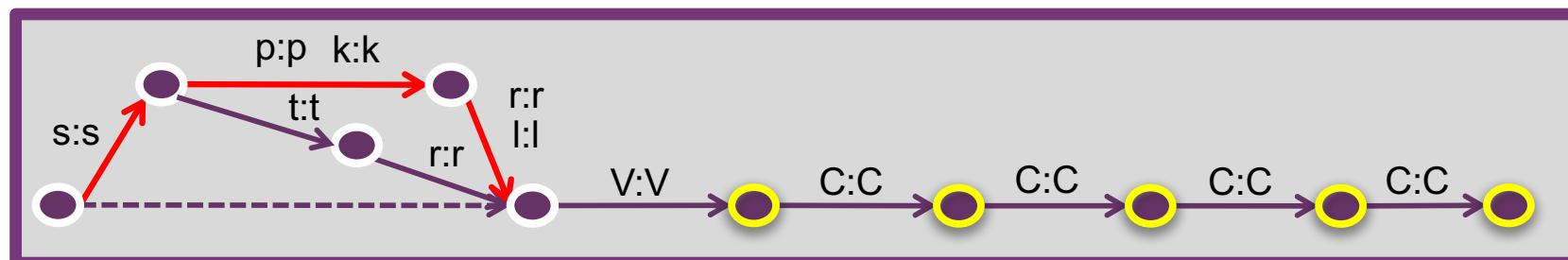
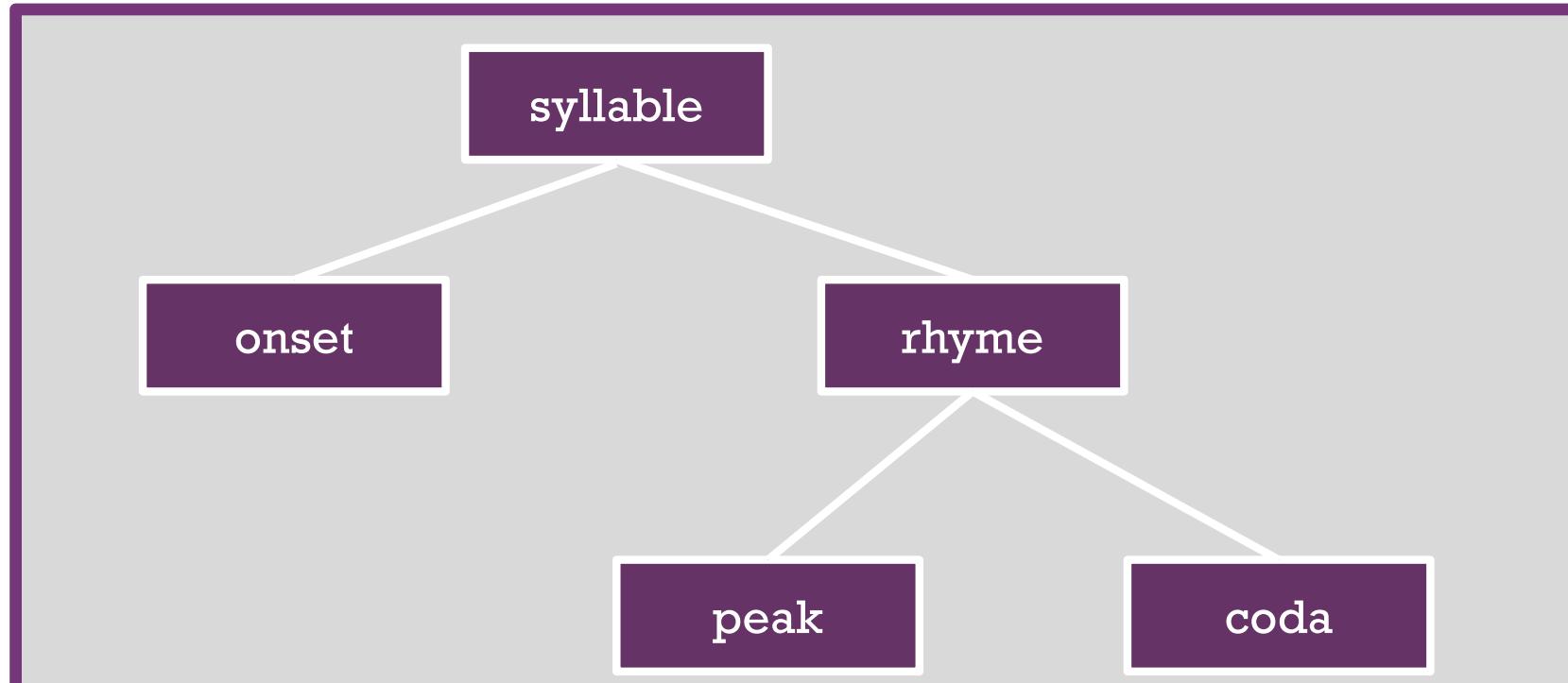
+Onset Restrictions: Transducer

HLT12



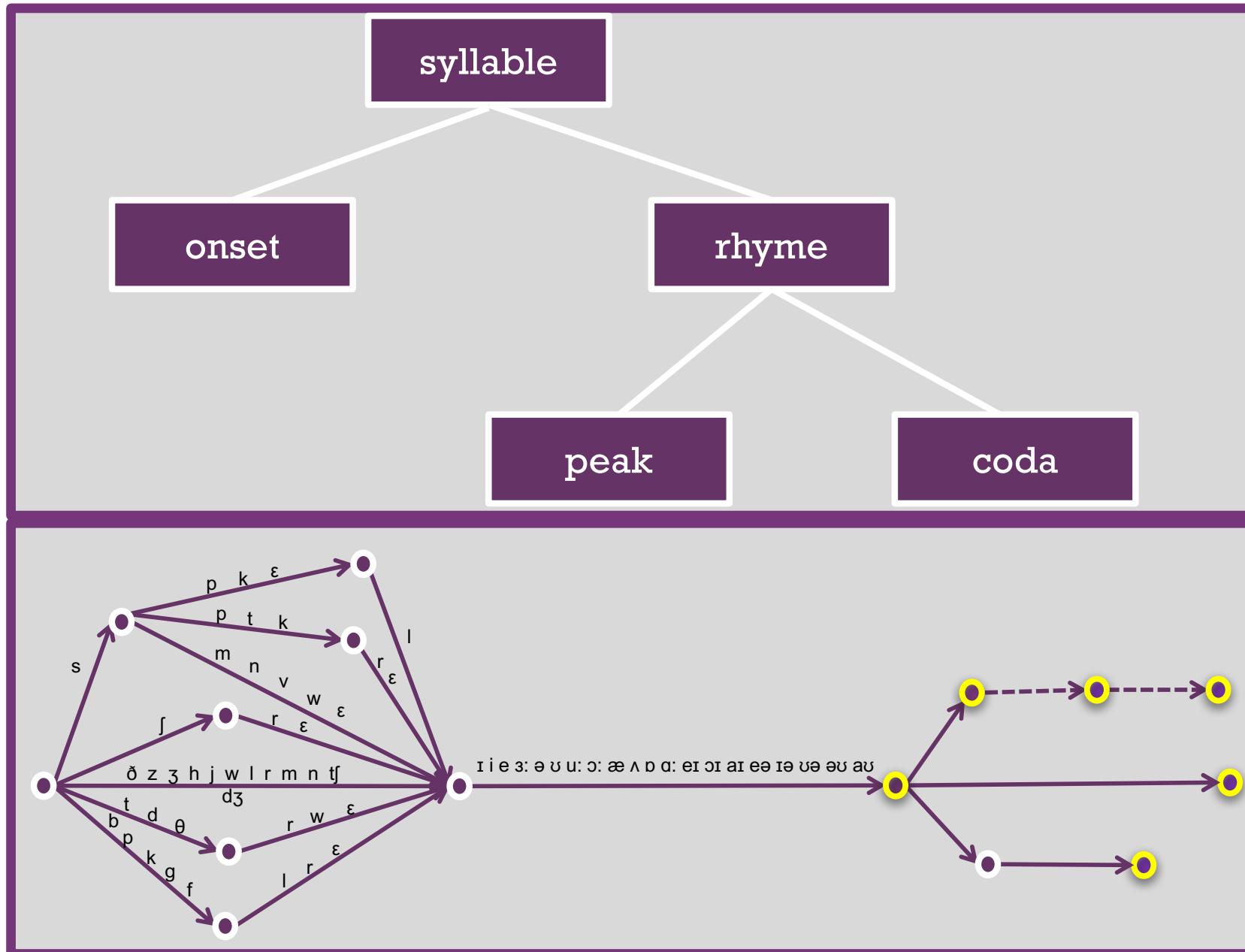
+Onset Restrictions: Transducer

HLT12



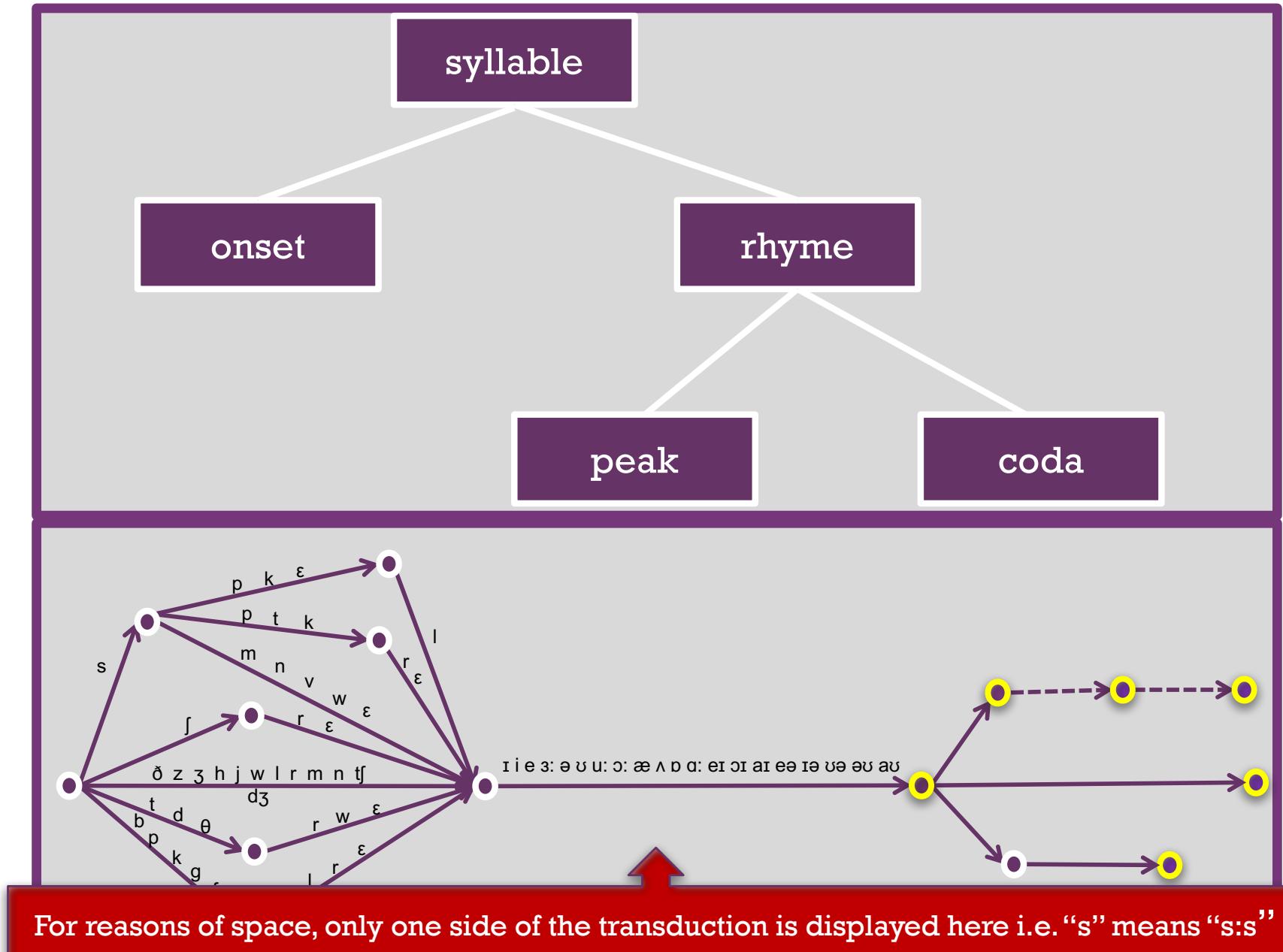
+ Phonotactic Transducer

HLT12



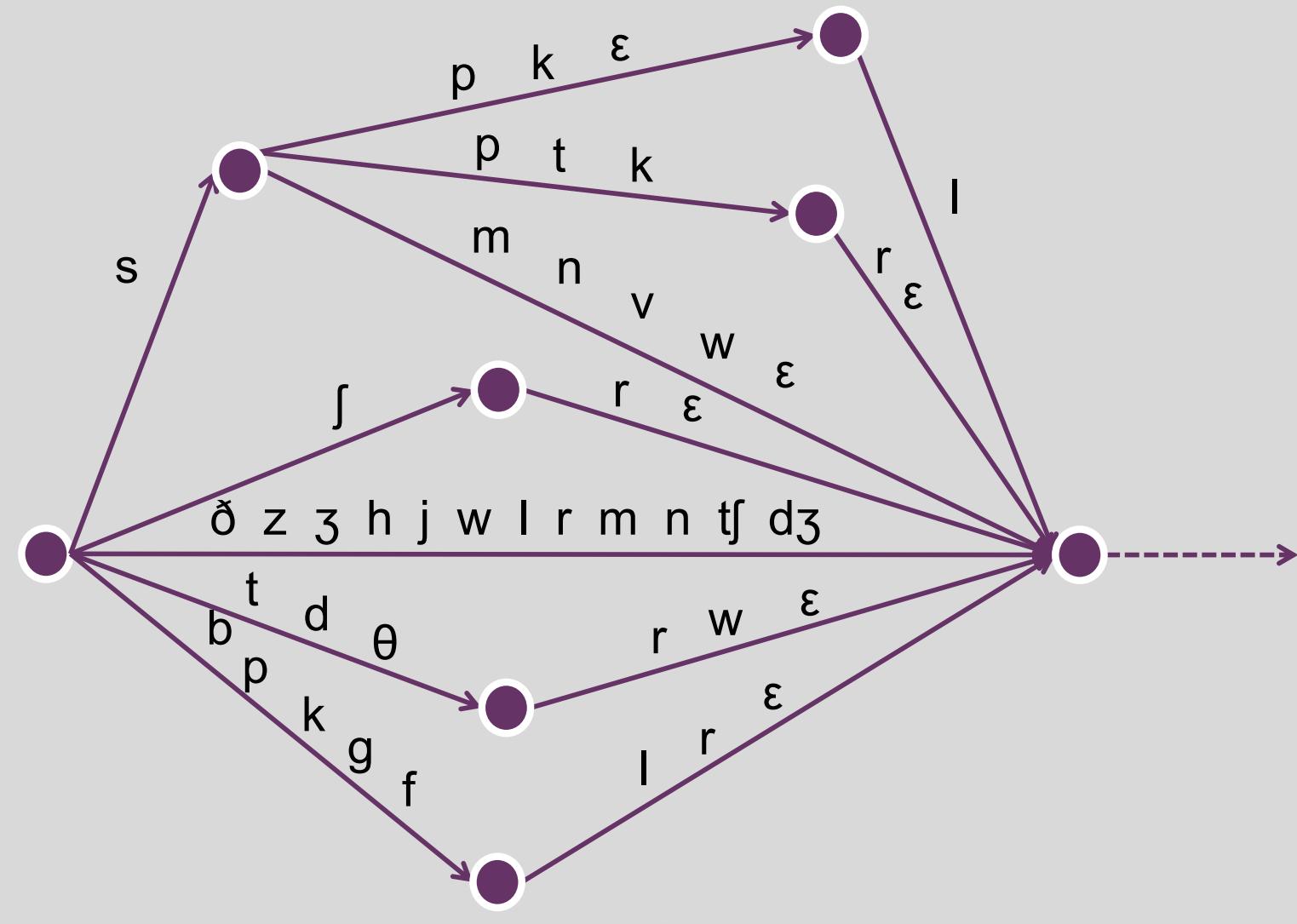
+ Phonotactic Transducer

HLT12



+ Phonotactic Transducer: Onset

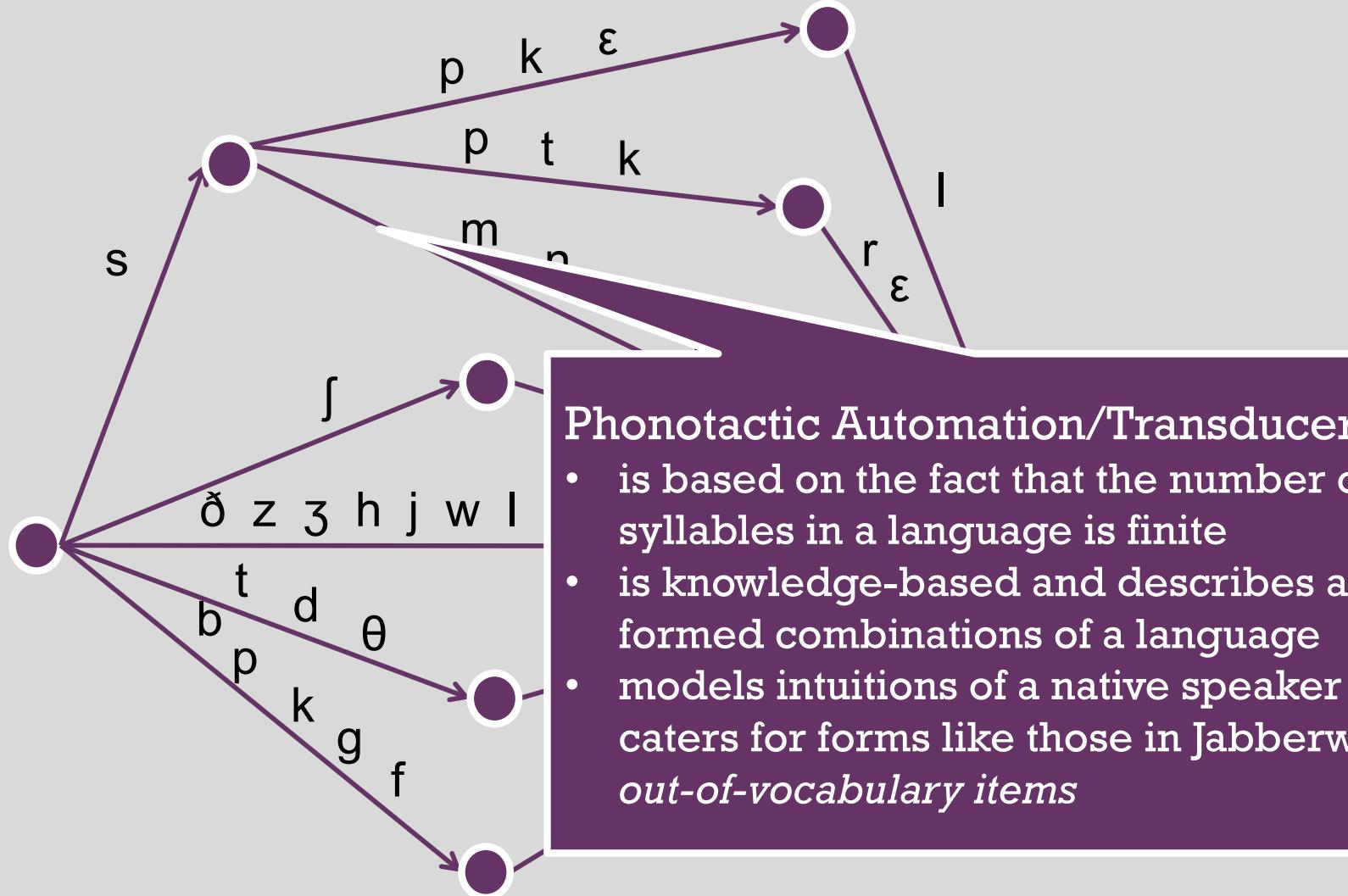
HLT12



For reasons of space, only one side of the transduction is displayed here i.e. "s" means "s:s"

+ Phonotactic Transducer: Onset

HLT12



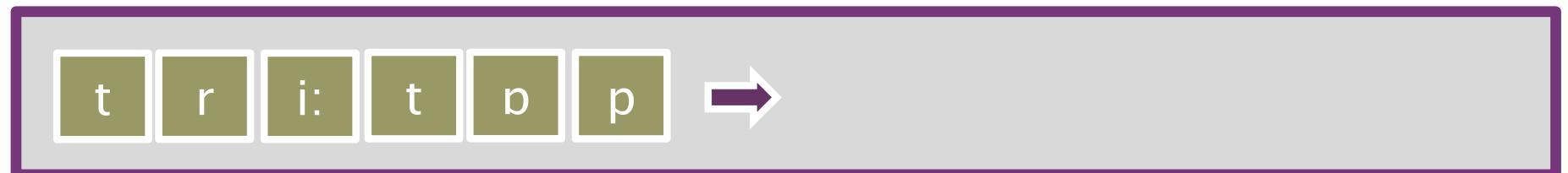
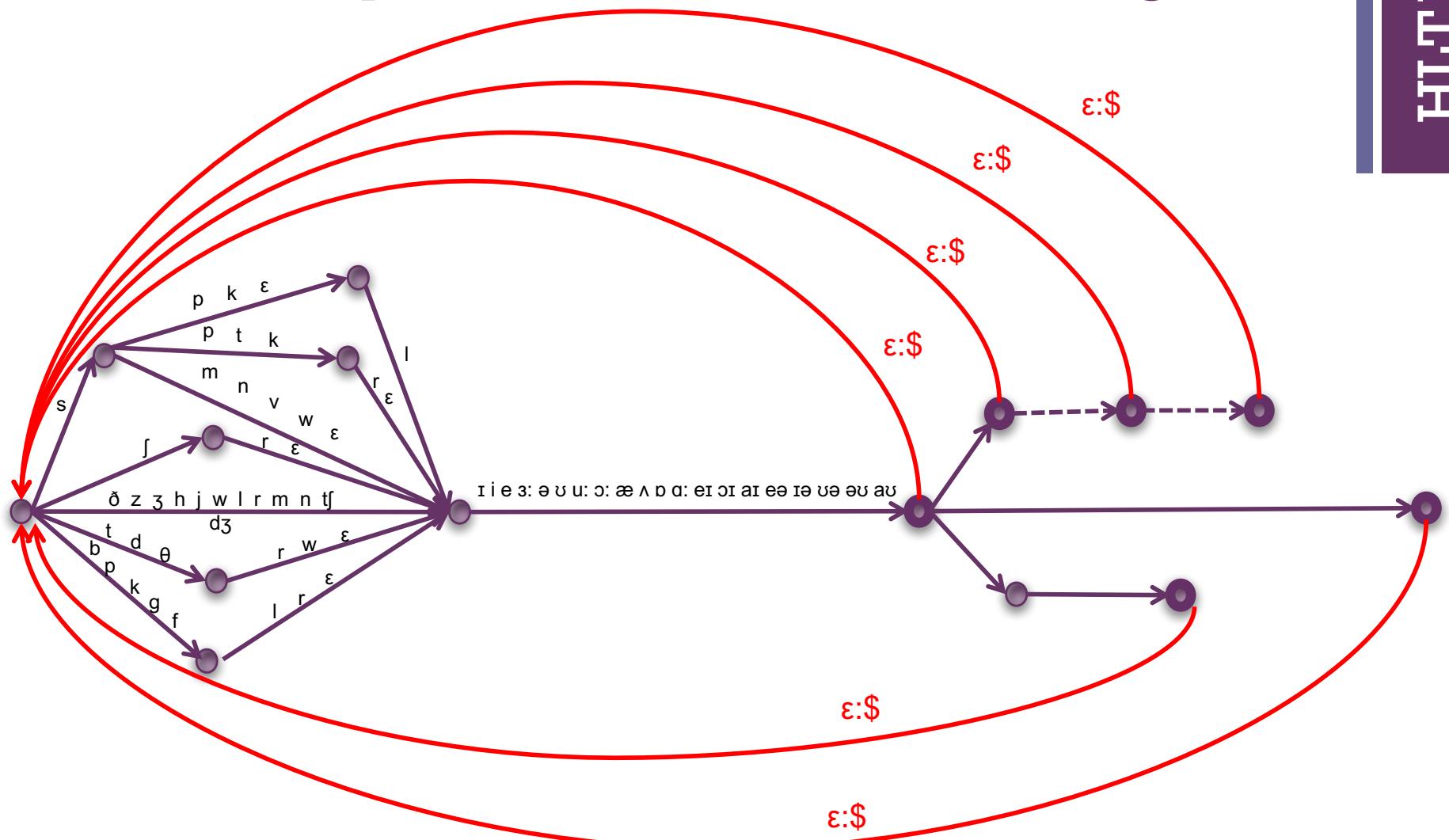
Phonotactic Automation/Transducer:

- is based on the fact that the number of syllables in a language is finite
- is knowledge-based and describes all well-formed combinations of a language
- models intuitions of a native speaker and thus caters for forms like those in Jabberwocky → *out-of-vocabulary items*

For reasons of space, only one side of the transduction is displayed here i.e. “s” means “s:s”

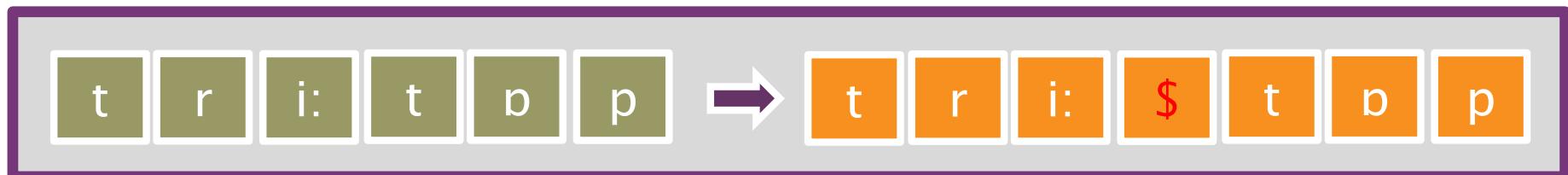
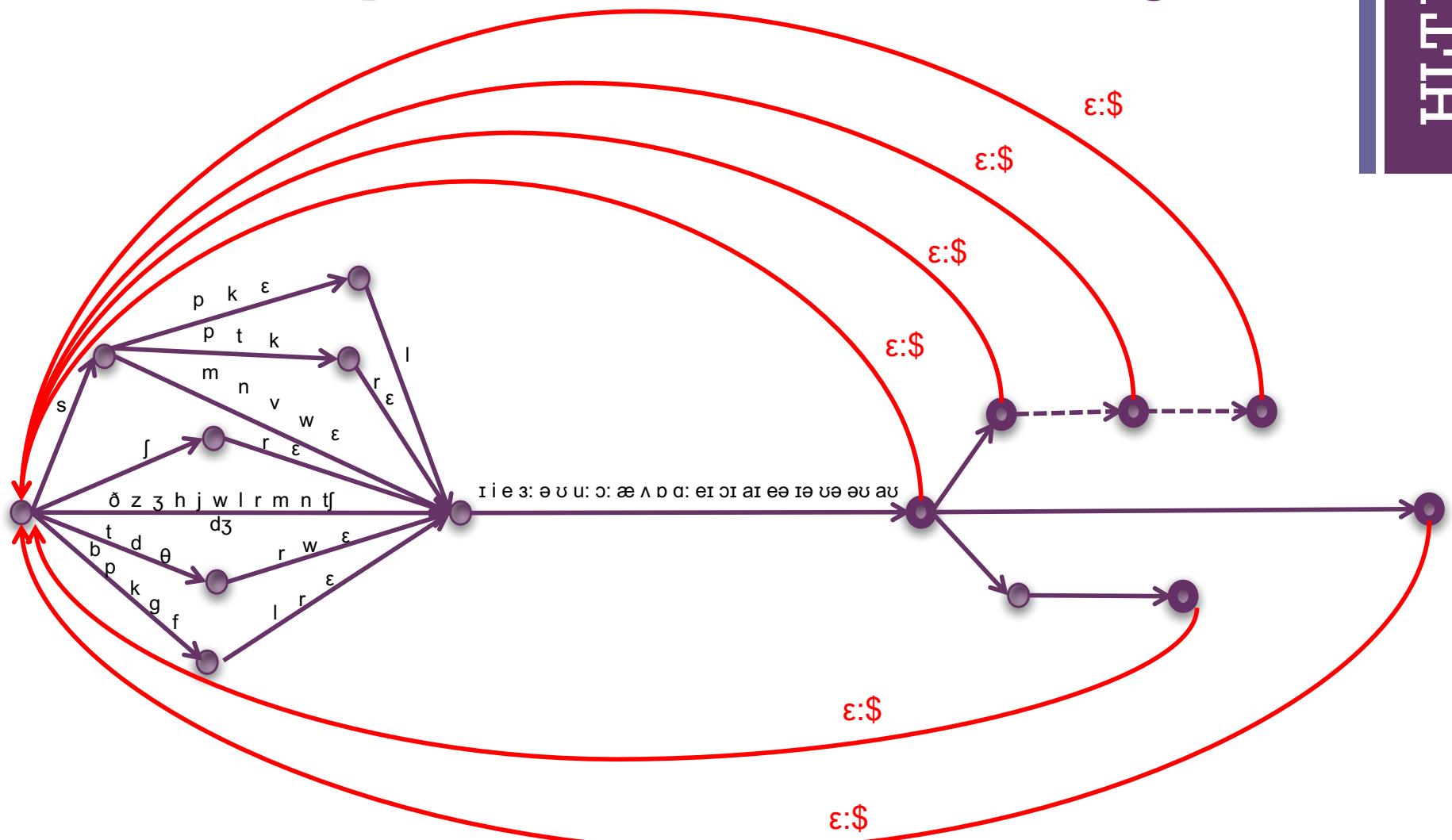
+ Boundary Placement \approx Parsing

HLT12



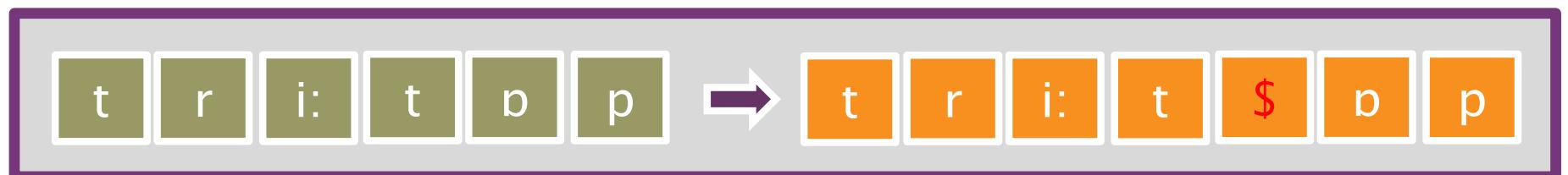
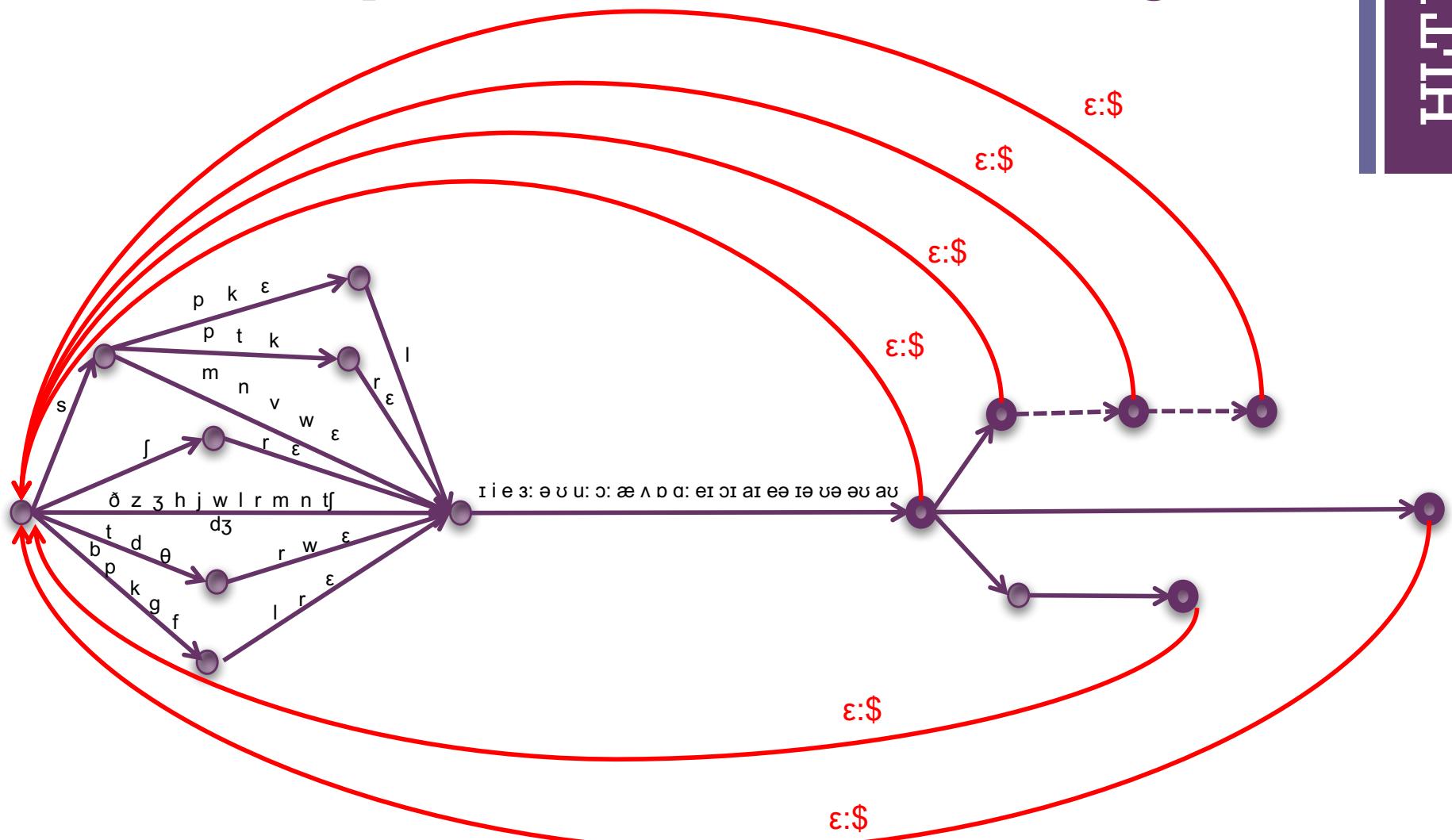
+ Boundary Placement \approx Parsing

HLT12



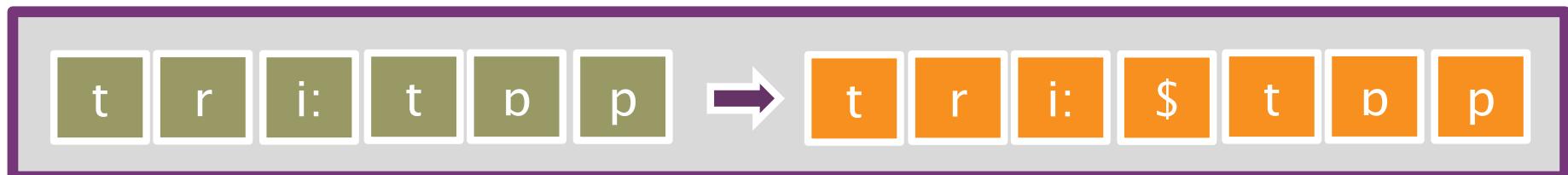
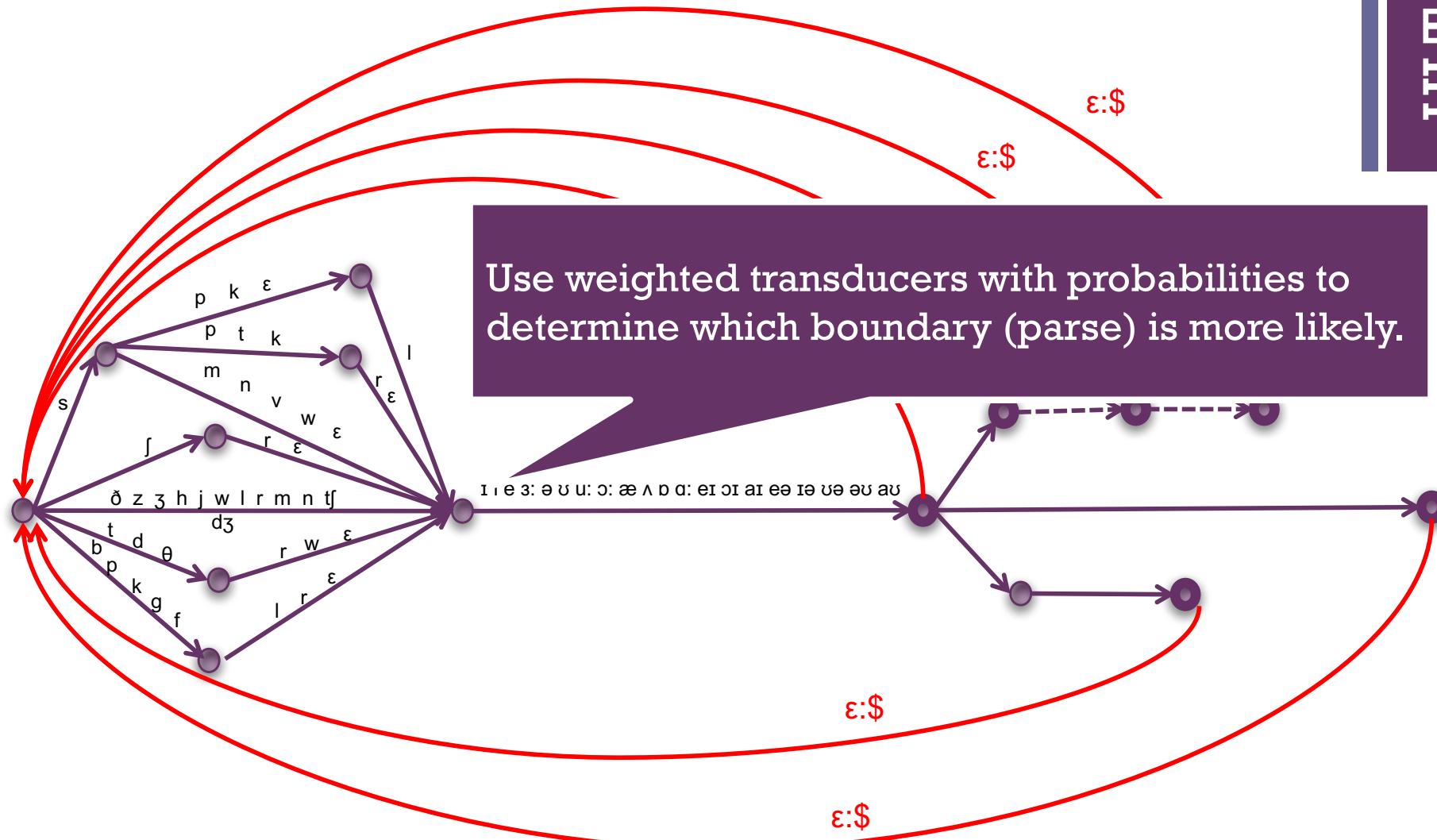
+ Boundary Placement \approx Parsing

HLT12



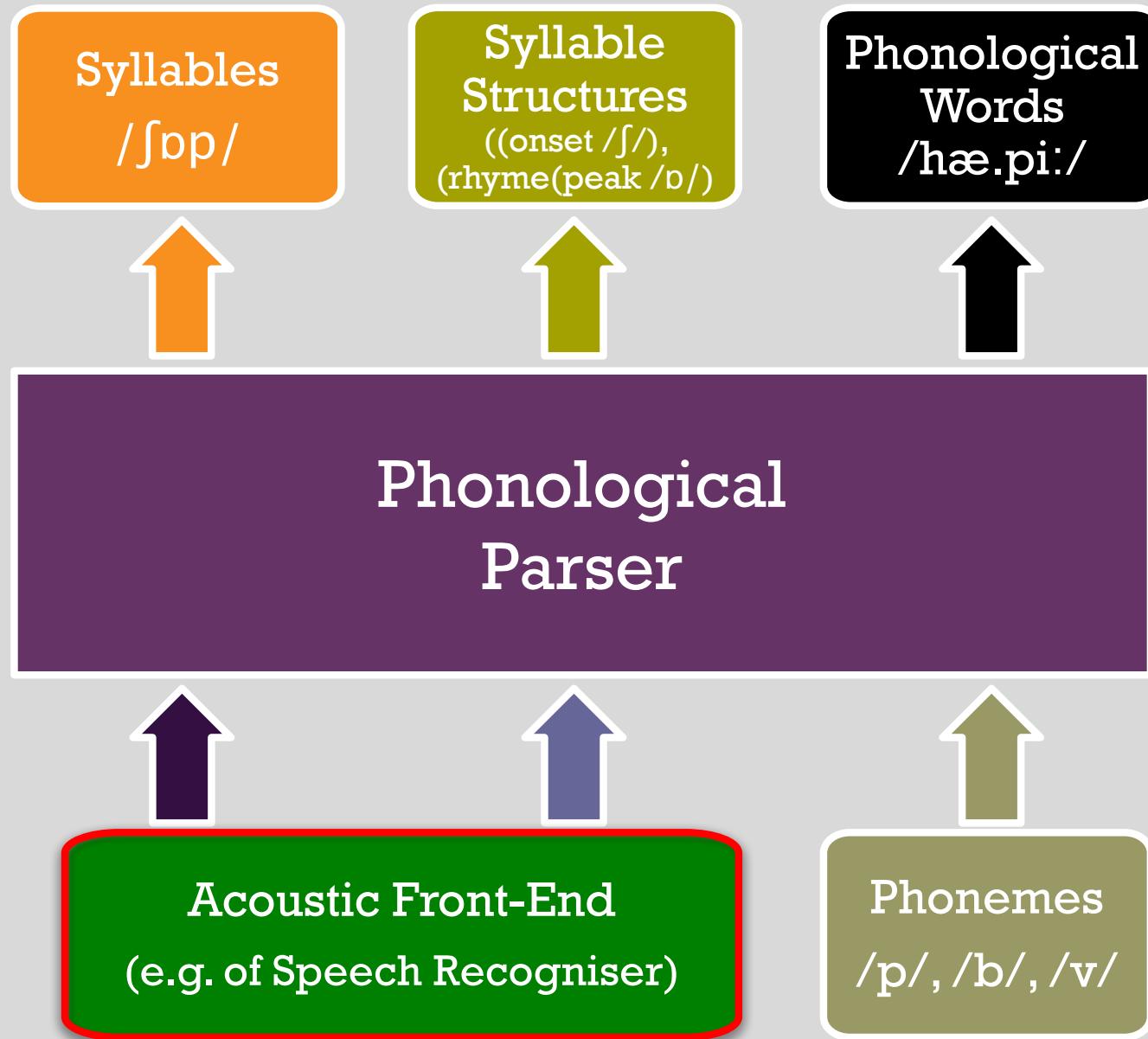
+ Boundary Placement \approx Parsing

HLT12



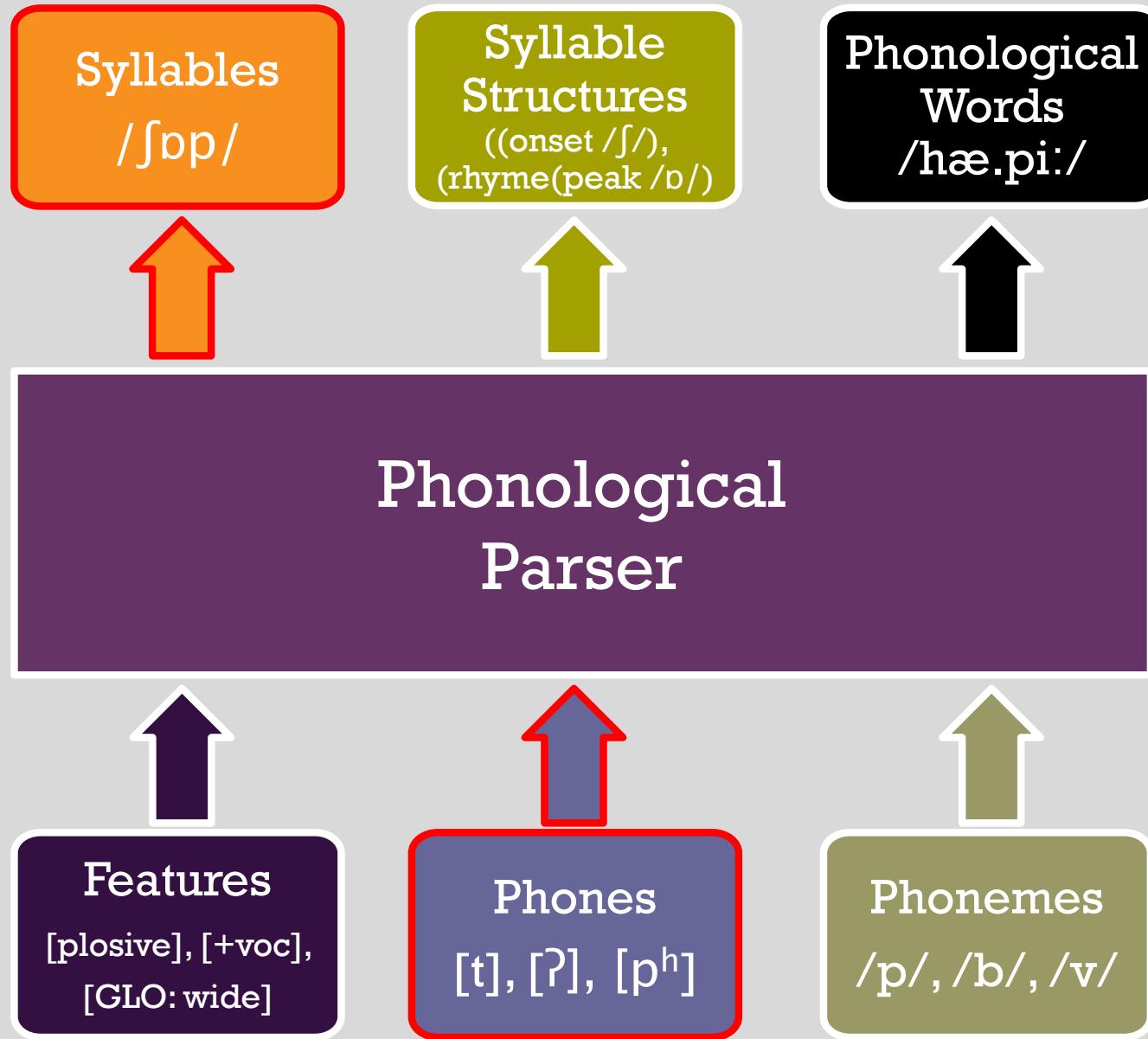
+ Phonological Parsing

HLT12



+ Phonological Parsing

HLT12



+ Phonological & Allophonic Rules

- Based on Church (1987):

HLT12

d I j ə h I r I ? t^h ə t^h a: m



+ Phonological & Allophonic Rules

HLT12

- Based on Church (1987):



- /h/ is always syllable initial ($/h/ \rightarrow [h] / \$ _ V$)

+ Phonological & Allophonic Rules

HLT12

- Based on Church (1987):



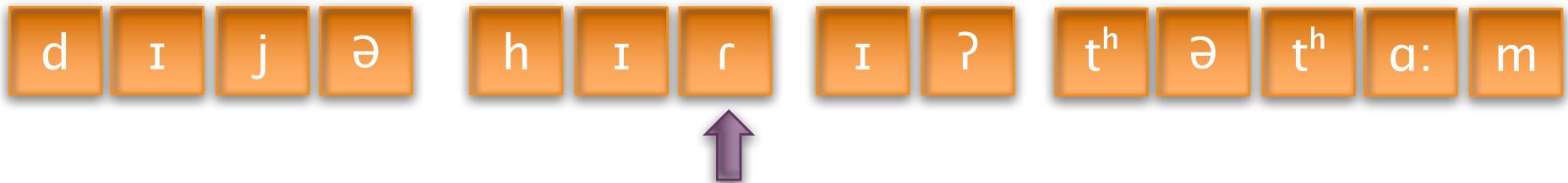
- [?] is always syllable final (/t/ → [?] / ____ \$)



+ Phonological & Allophonic Rules

HLT12

- Based on Church (1987):



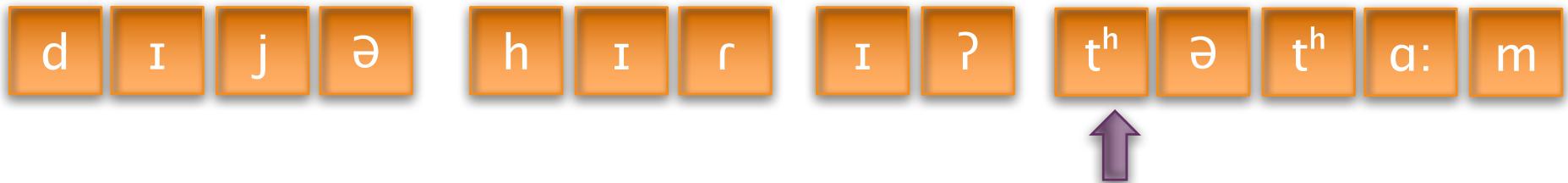
- [r] is always syllable final (/t/ → [r] / ___\$V)



+ Phonological & Allophonic Rules

HLT12

- Based on Church (1987):

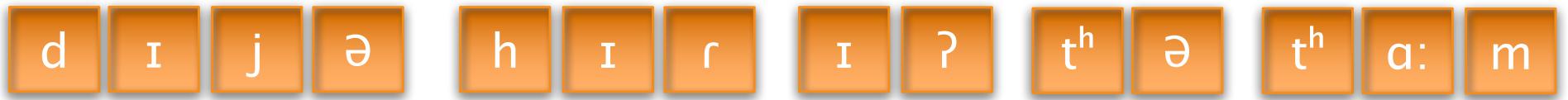


- [t^h] is always syllable initial (/t/ → [t^h] / \$__V)

+ Phonological & Allophonic Rules

HLT12

- Based on Church (1987):



- [t^h] is always syllable initial (/t/ → [t^h] / \$__V)

+ Decoding to Words

Parsed Transcription

d i j θ



Decoding

did you

h I r



hit

I ?



it

t^h θ



to

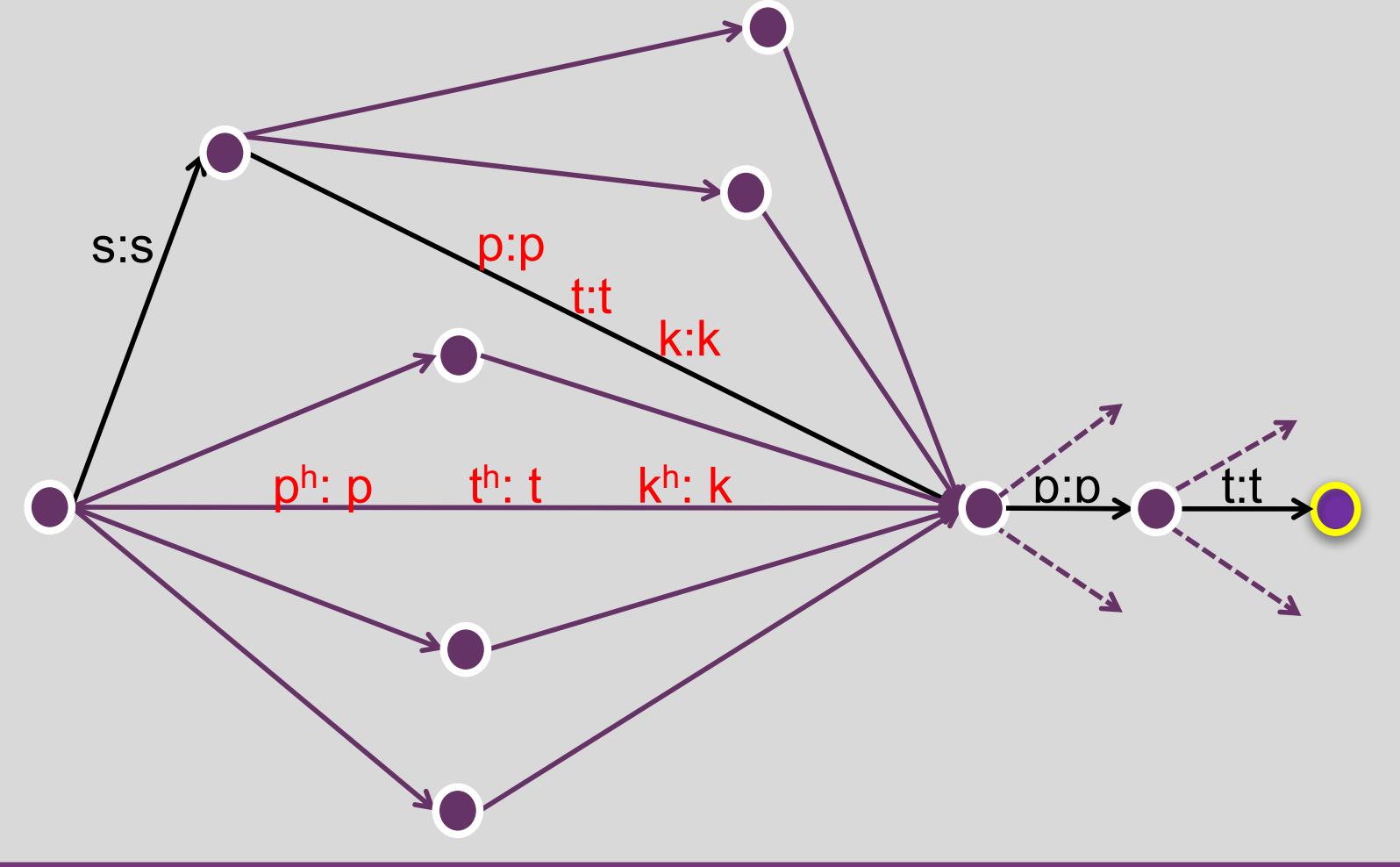
t^h a: m



Tom

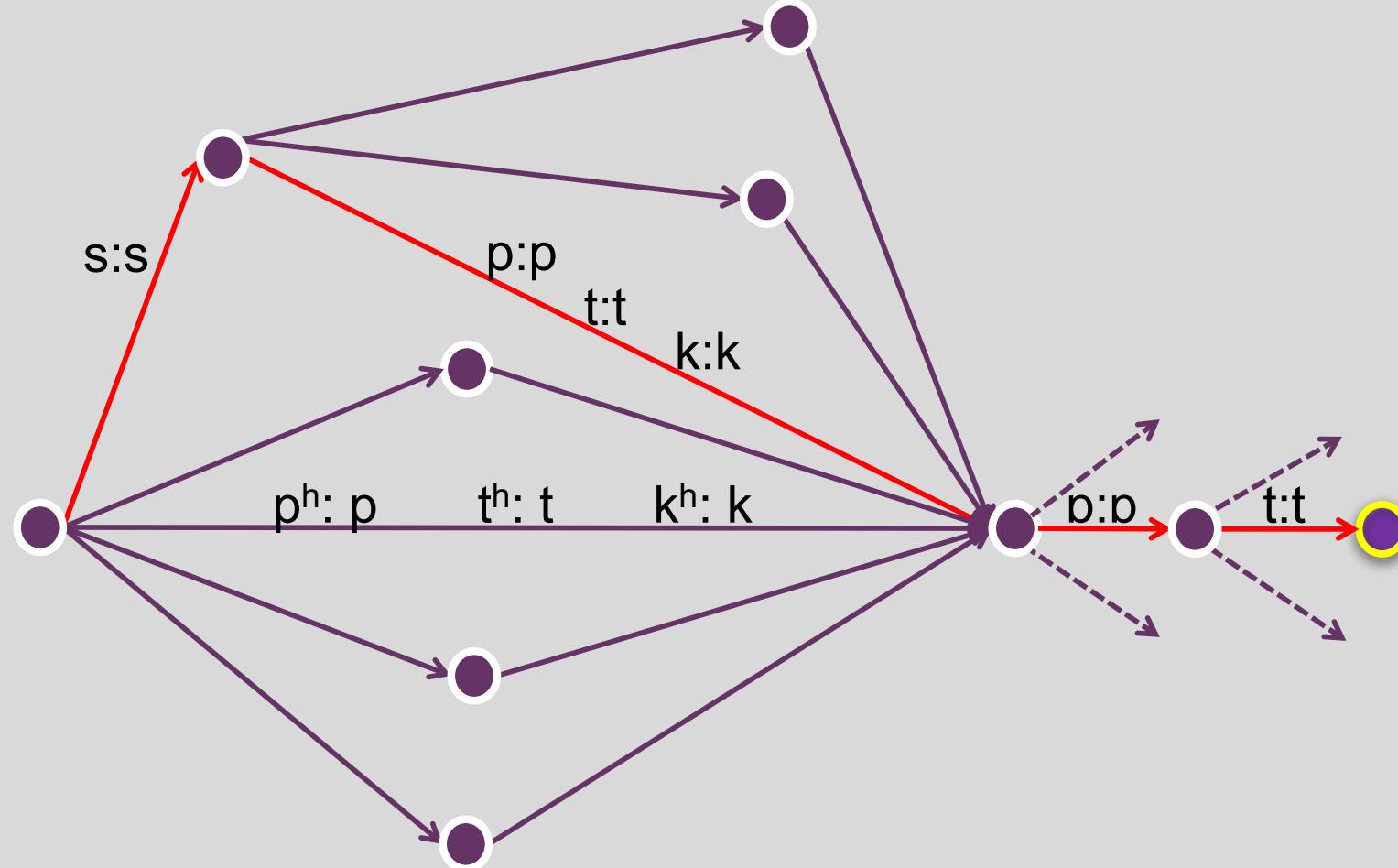
+ Allophonic-to-Phonemic Transducer

HLT12



+ Allophonic-to-Phonemic Transducer

HLT12



s p b t

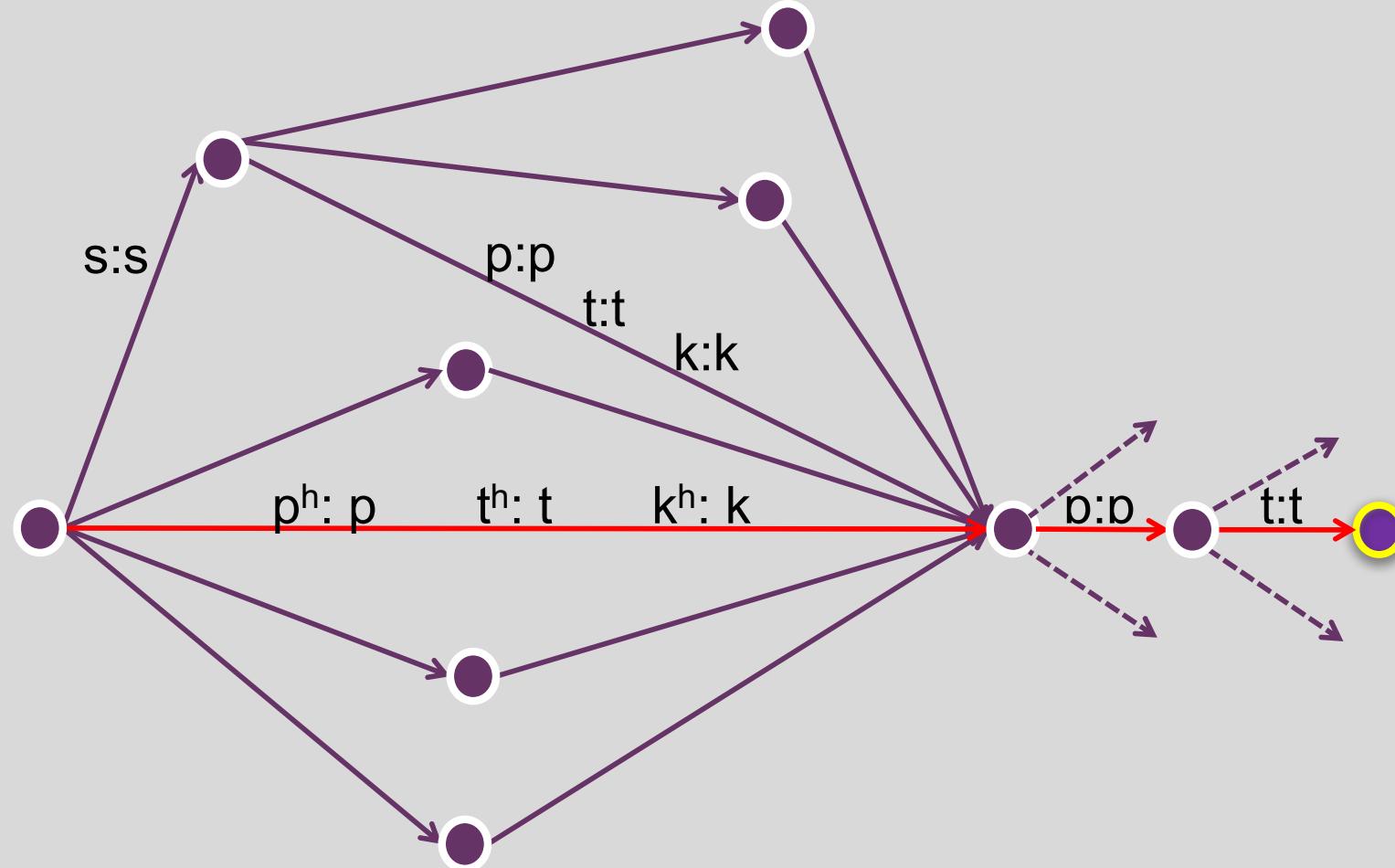


s p b t



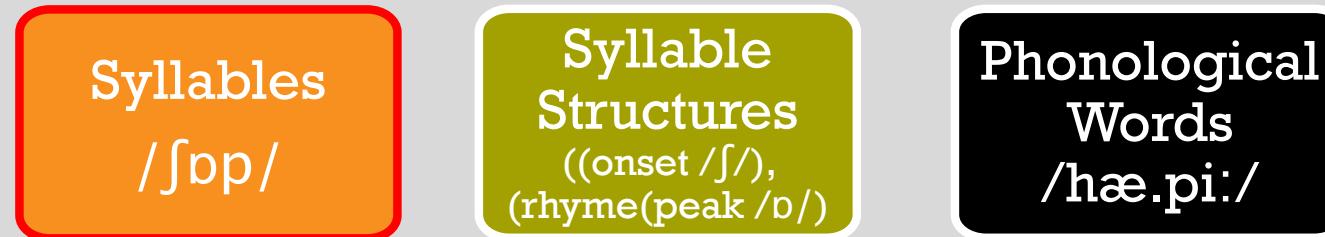
+Allophonic-to-Phonemic Transducer

HLT12



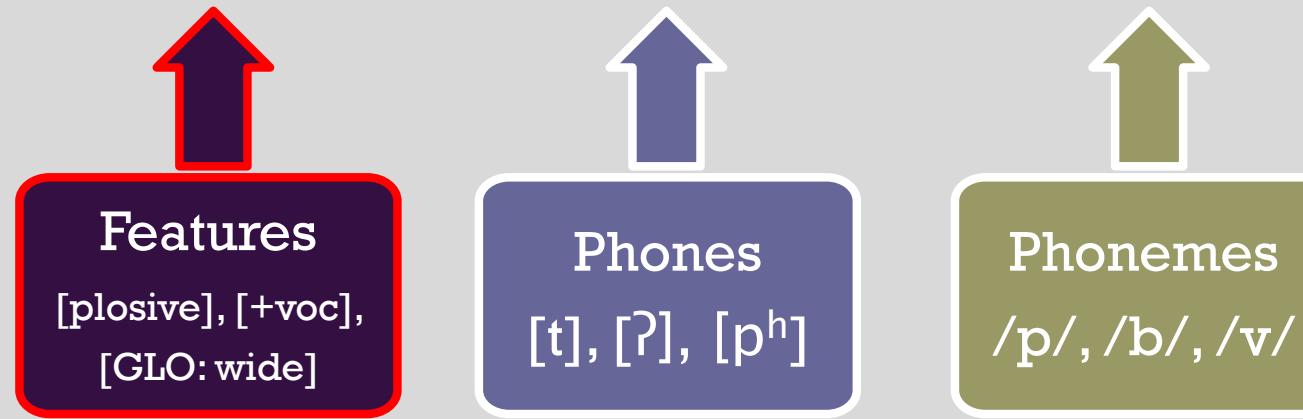
+ Phonological Parsing

HLT12



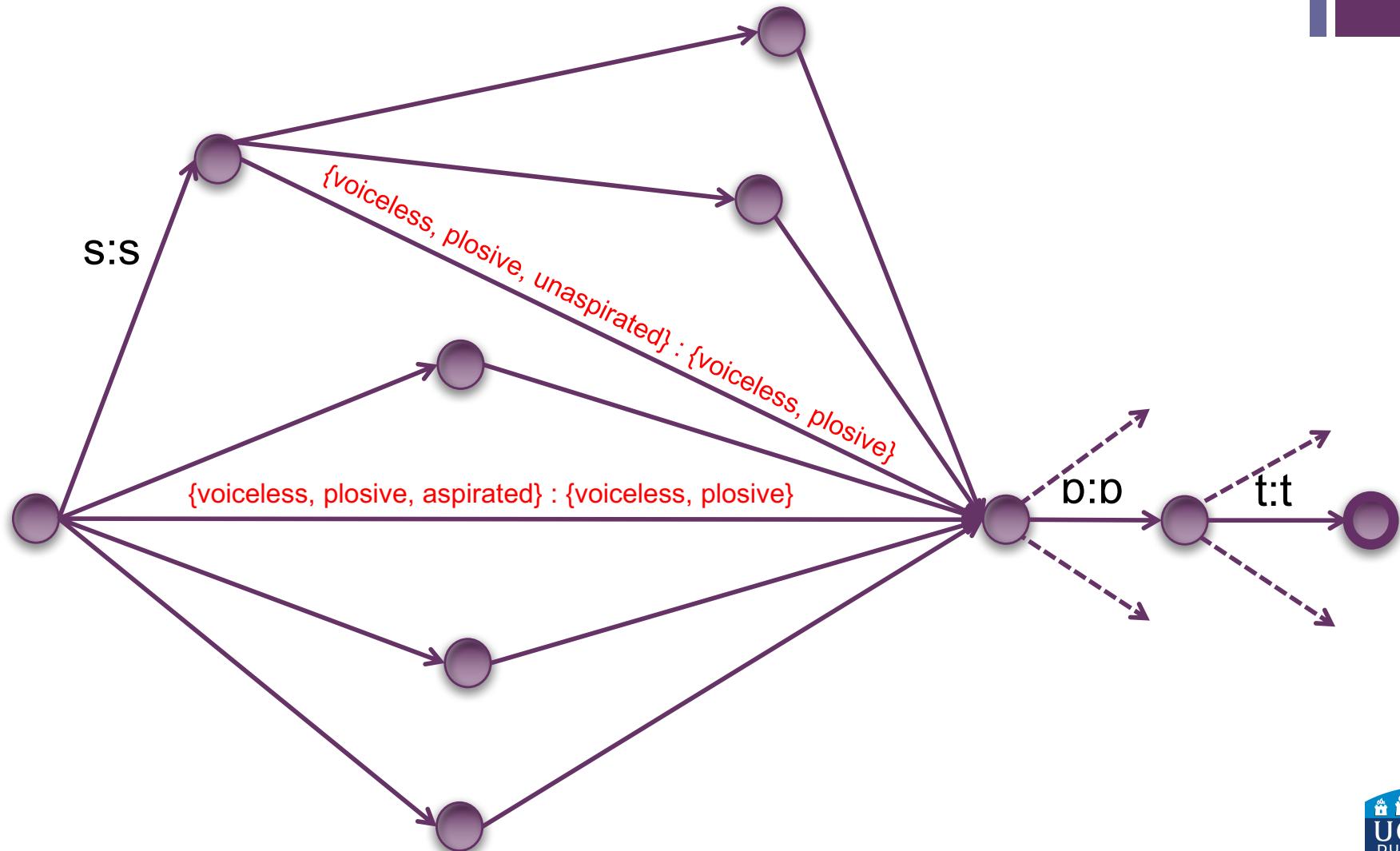
Phonological
Parser

Phonological
“Grammar”



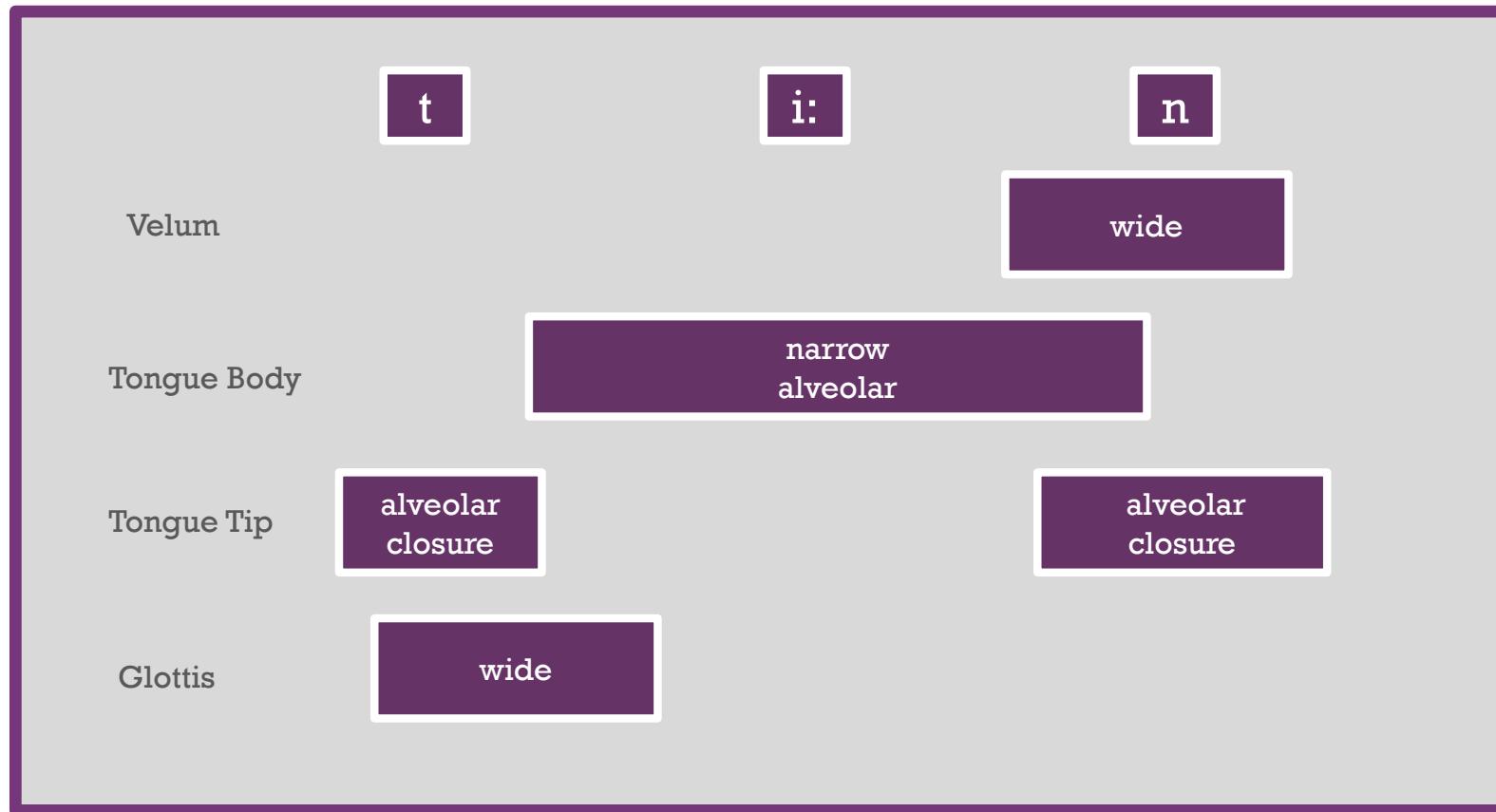
+ Feature Phonotactic Transducer

HLT12



+ Recall: Multi-linear Approach

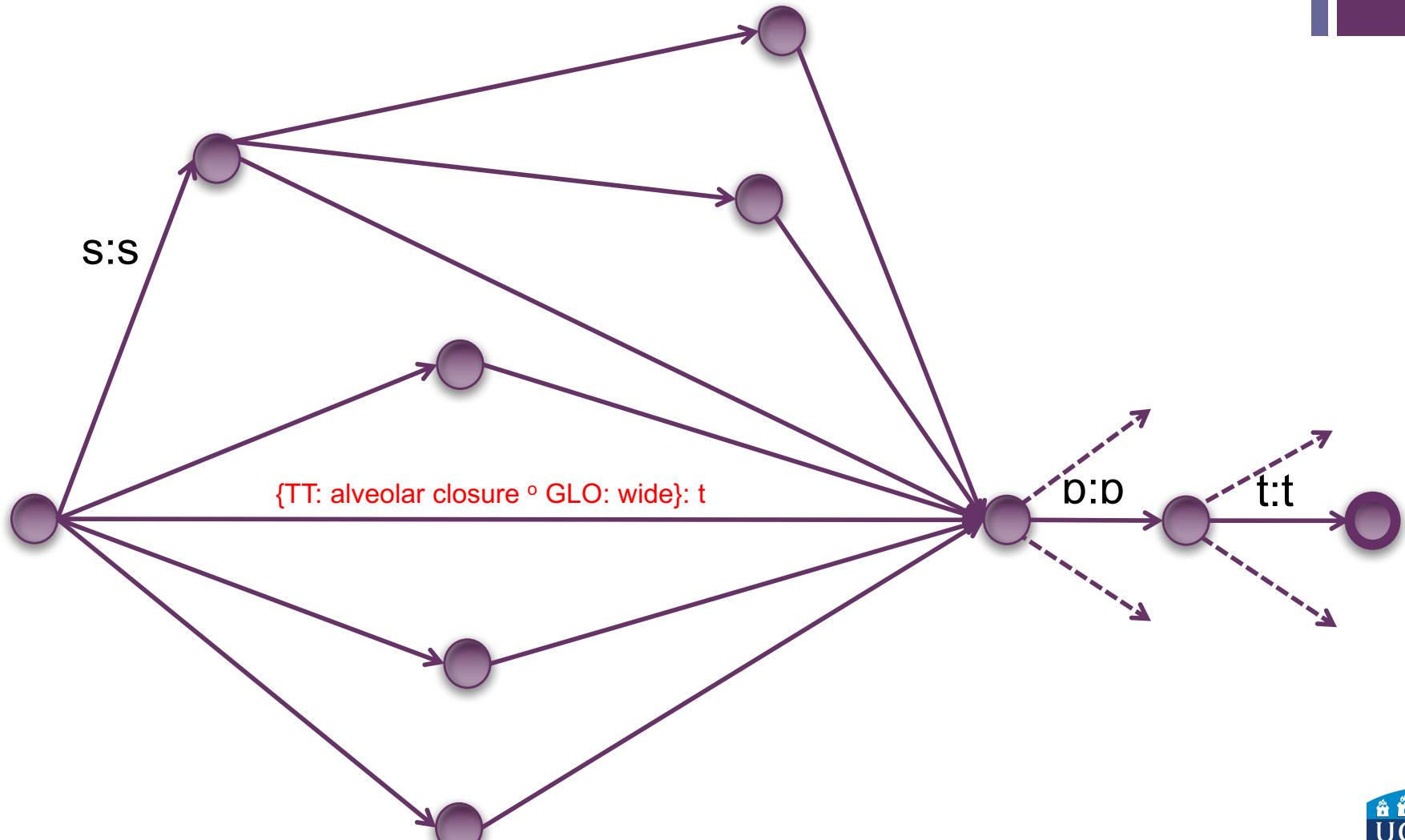
The word *teen* can also be regarded as a multi-linear representation of overlapping articulatory gestures also accounting for duration:



Gestural Score as used in Articulatory Speech Synthesis

+ Multi-Linear Phonotactic Transducer

HLT12



+ Finite-State Toolkits

HLT12

■ Xerox Finite State Toolkit

- <http://web.stanford.edu/~laurik/fsmbook/home.html>

■ OpenFST

- <http://www.openfst.org/>

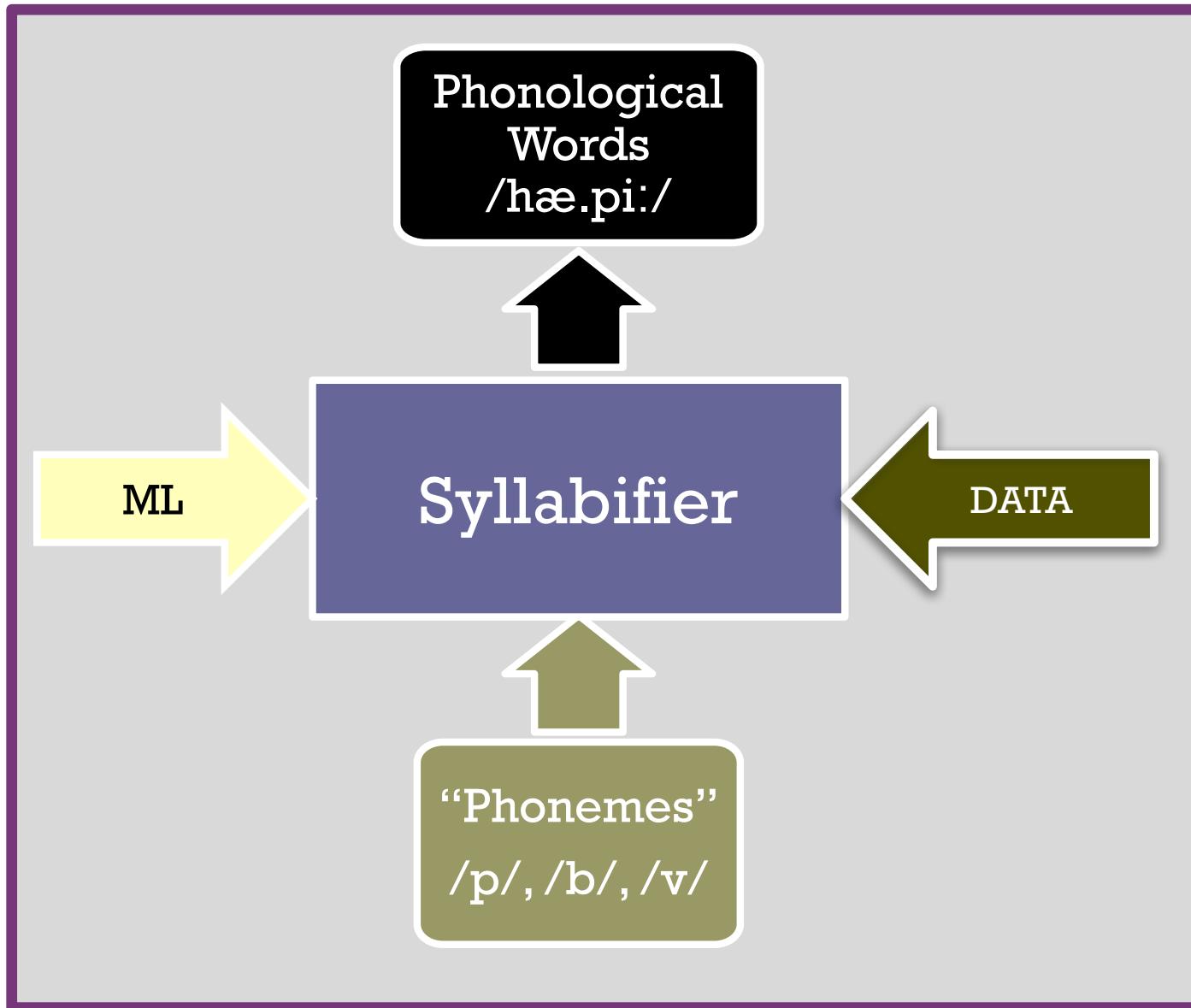
■ HFST

- <https://kitwiki.csc.fi/twiki/bin/view/KitWiki/HfstHome>

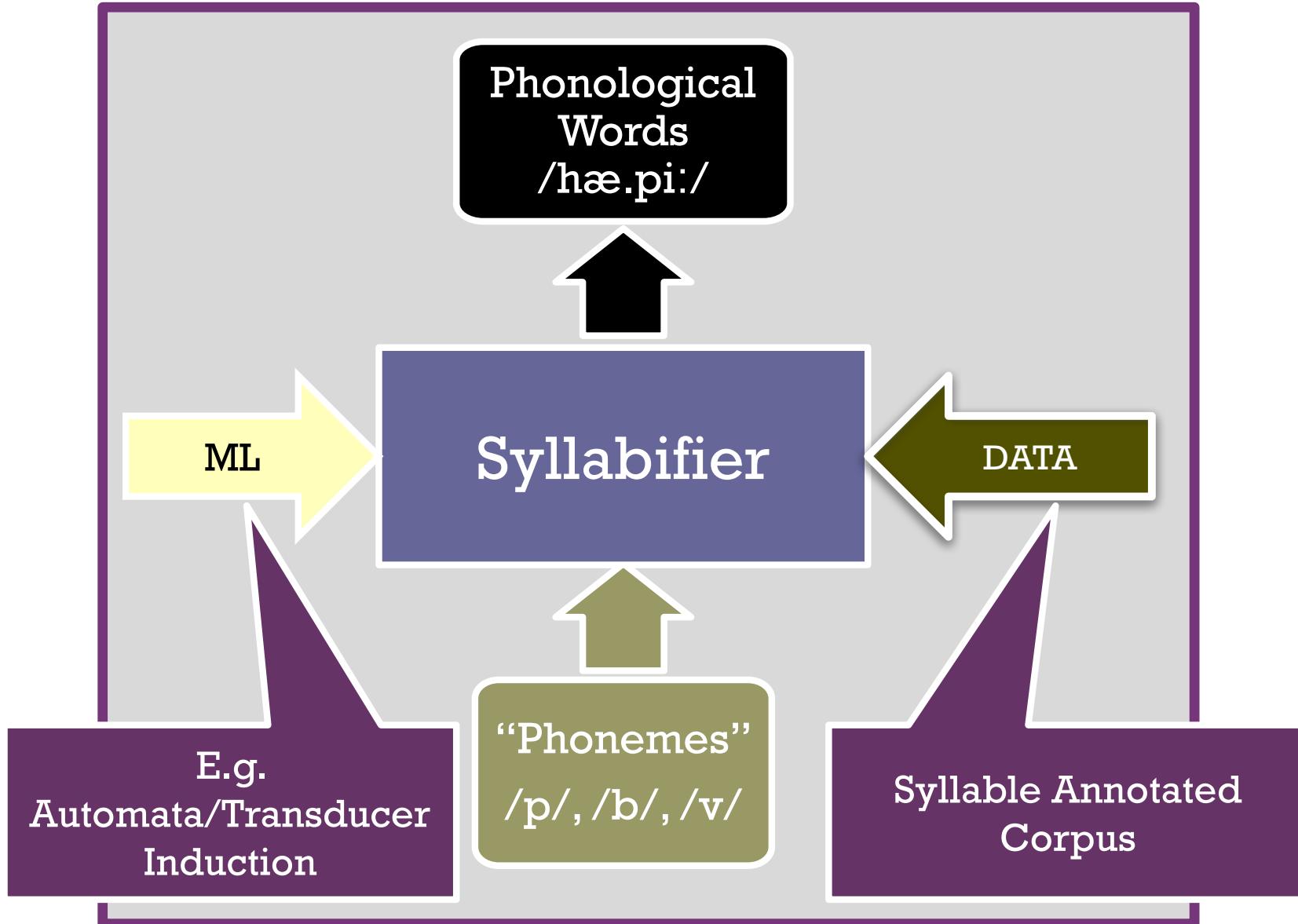


+ Syllabification

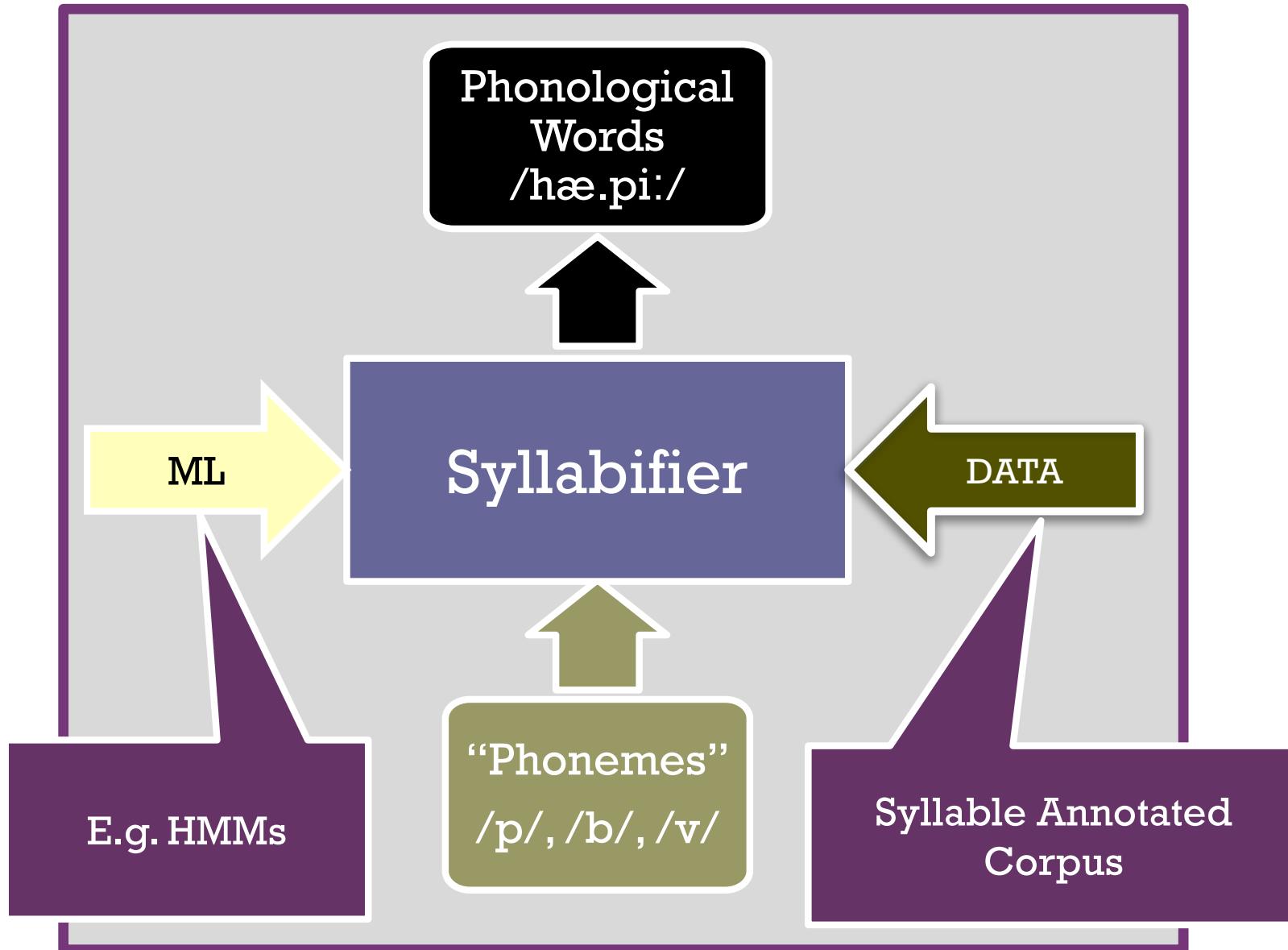
HLT12



+ Syllabification

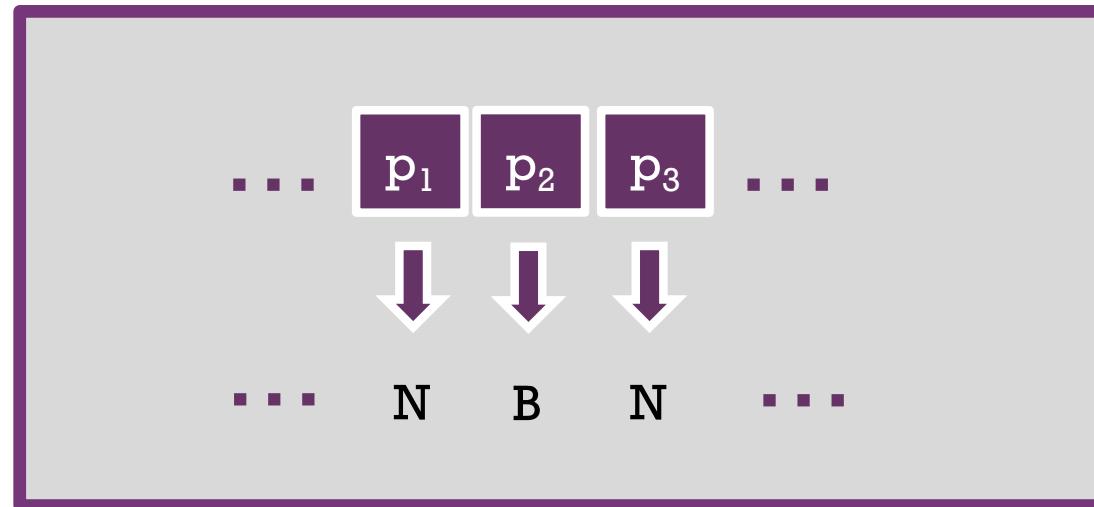


+ Syllabification



+ Syllabification as Tagging...

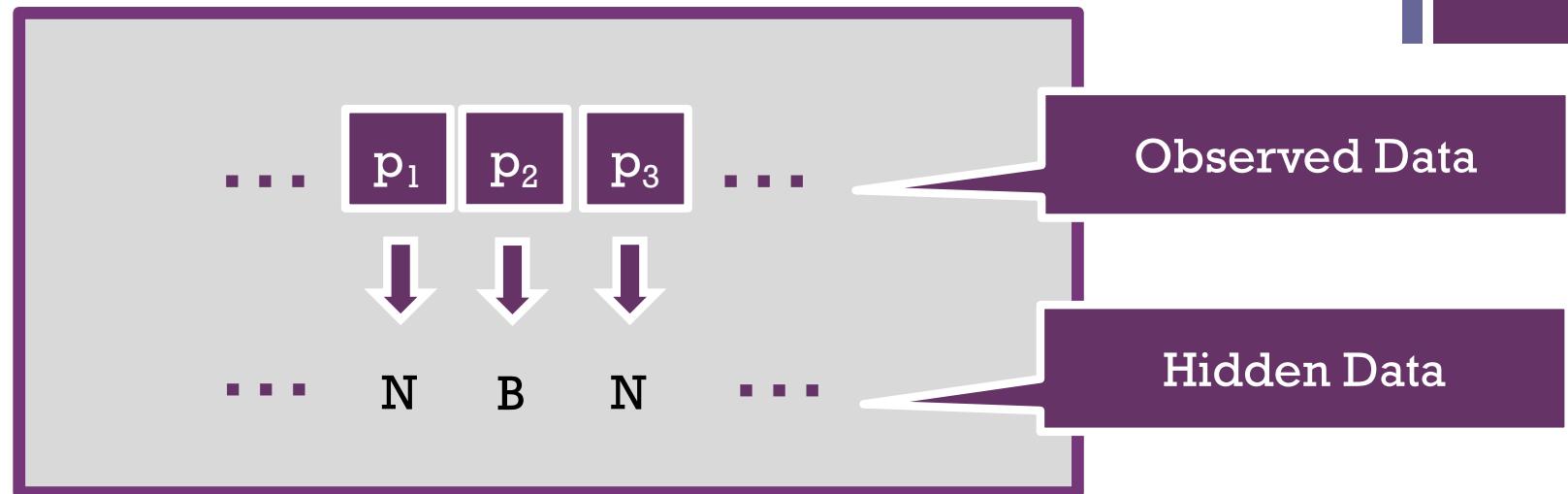
■ Boundary (B) vs No Boundary (N)



+ Syllabification as Tagging...

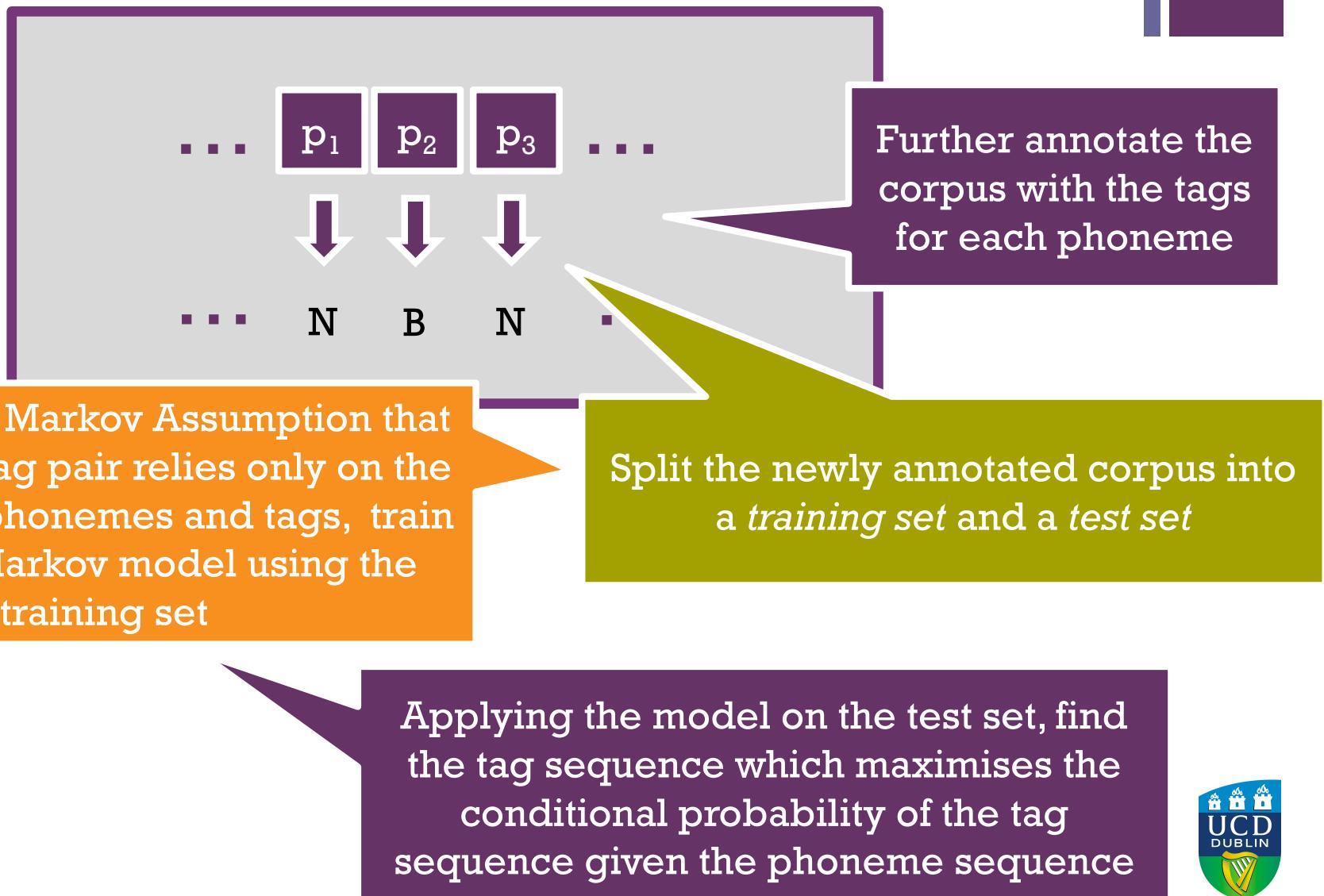
HLT12

■ Boundary (B) vs No Boundary (N)



+ Syllabification as Tagging...

■ Boundary (B) vs No Boundary (N)



+ Syllabification as Tagging...

■ Boundary (B) vs No Boundary (N)

