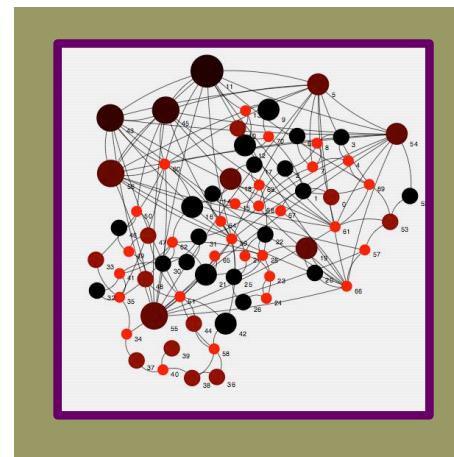
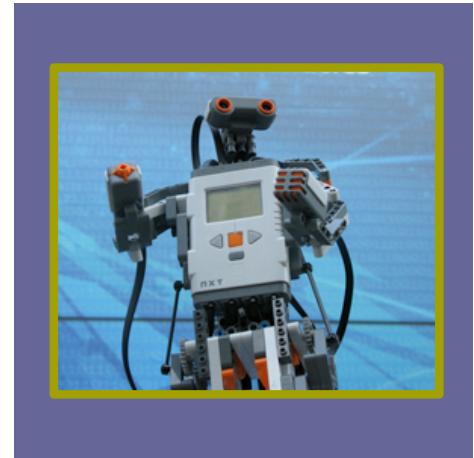




COMP40020

Human Language Technologies

Sounds, Symbols & Features
March 2019



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

Contents:

- Sounds of language
- Transcription and labelling
- Systematic organisation of sounds
- Definitions and examples

Aim:

- To give a brief introduction to some of the concepts from the domain of phonetics and phonology.

+ Phonetics

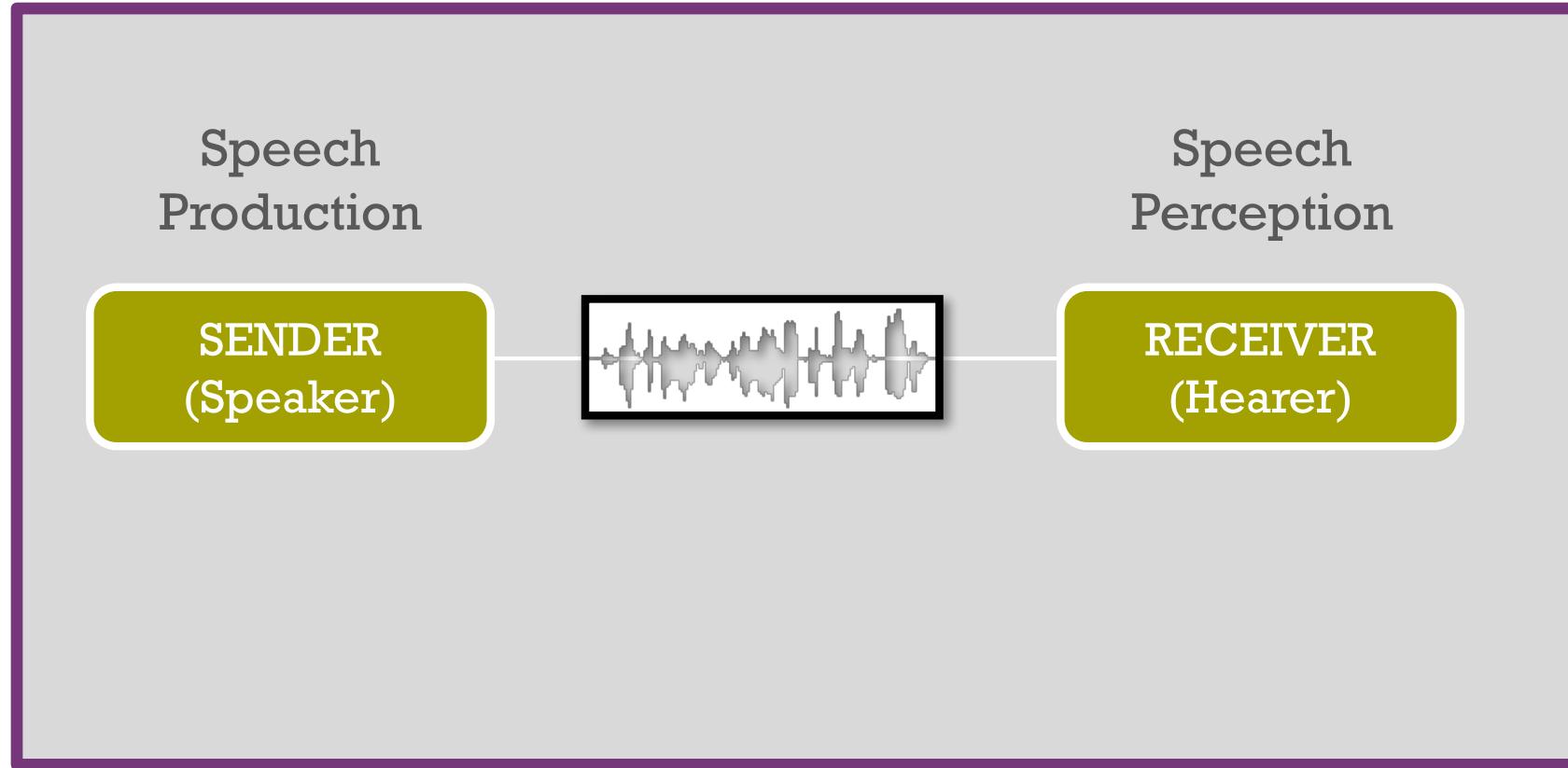
HLT11

- Phonetics is concerned with describing the speech sounds that occur in the languages of the world.
- We want to know:
 - What these sounds are
 - How they fall into patterns
 - How they change in different circumstances
- Most importantly we want to know what aspects of the sounds are necessary for conveying the meaning of what is being said.
- The first job of a phonetician is therefore trying to find out what people are doing when they are talking and when they are listening to speech. (Ladefoged (1975))



+ Observing Speech...

HLT11



+ Observing Speech...

HLT11

SENDER
(Speaker)

CHANNEL

RECEIVER
(Hearer)

Articulatory
Phonetics

Acoustic
Phonetics

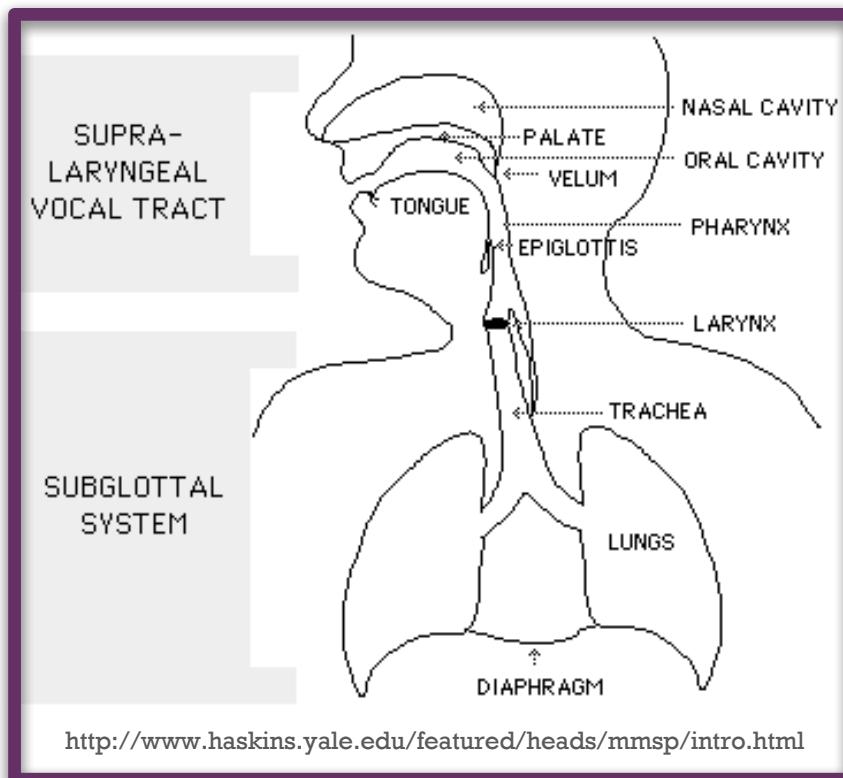
Auditory
Phonetics



+ Describing Sounds

HILT11

- The oral passages above the larynx are known as the vocal tract. The parts of the oral tract that can be used to form sounds are called **articulators**.



- air stream mechanism
- manner of articulation
- place of articulation
- degree of voicing

- Sounds produced when the vocal cords are vibrating are said to be **voiced**, as opposed to those in which the vocal cords are apart, which are said to be **voiceless**.

+ International Phonetic Alphabet

HLT11

■ The IPA Chart (Consonants)

THE INTERNATIONAL PHONETIC ALPHABET (revised to 2015)

CONSONANTS (PULMONIC)

© 2015 IPA

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		t d̪	c ʃ	k g	q G		?
Nasal	m	m̪		n		ɳ	ɲ	ŋ	N		
Trill	B			r					R		
Tap or Flap		v̪		r̪		t̪					
Fricative	ɸ β	f v̪	θ ð	s z	ʃ ʒ	s z̪	ç j̪	x y	χ ʁ	h ʕ	h̪ f̪
Lateral fricative				ɬ ɭ							
Approximant		v̪		ɹ̪		ɺ	j̪	w̪			
Lateral approximant				l̪		ɻ	ʎ	ɿ			

Symbols to the right in a cell are voiced, to the left are voiceless. Shaded areas denote articulations judged impossible.

https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf



+ International Phonetic Alphabet

HILT11

■ The IPA Chart (Diacritics)

DIACRITICS Some diacritics may be placed above a symbol with a descender, e.g. $\overset{\circ}{\text{ŋ}}$

◦ Voiceless	$\overset{\circ}{\text{n}}$ $\overset{\circ}{\text{d}}$.. Breathy voiced	$\overset{\circ}{\text{b}}$ $\overset{\circ}{\text{a}}$	▫ Dental	$\overset{\circ}{\text{t}}$ $\overset{\circ}{\text{d}}$	
◦ Voiced	$\overset{\circ}{\text{s}}$ $\overset{\circ}{\text{t}}$	~ Creaky voiced	$\overset{\circ}{\text{b}}$ $\overset{\circ}{\text{a}}$	▫ Apical	$\overset{\circ}{\text{t}}$ $\overset{\circ}{\text{d}}$	
h Aspirated	$\overset{\circ}{\text{t}}^h$ $\overset{\circ}{\text{d}}^h$	~ Linguolabial	$\overset{\circ}{\text{t}}$ $\overset{\circ}{\text{d}}$	▫ Laminal	$\overset{\circ}{\text{t}}$ $\overset{\circ}{\text{d}}$	
,	More rounded	$\overset{\circ}{\text{ɔ}}$	W Labialized	$\overset{\circ}{\text{t}}^w$ $\overset{\circ}{\text{d}}^w$	~ Nasalized	$\overset{\circ}{\text{e}}$
,	Less rounded	$\overset{\circ}{\text{o}}$	j Palatalized	$\overset{\circ}{\text{t}}^j$ $\overset{\circ}{\text{d}}^j$	n Nasal release	$\overset{\circ}{\text{d}}^n$
+	Advanced	$\overset{\circ}{\text{u}}$	Y Velarized	$\overset{\circ}{\text{t}}^y$ $\overset{\circ}{\text{d}}^y$	l Lateral release	$\overset{\circ}{\text{d}}^l$
-	Retracted	$\overset{\circ}{\text{e}}$	Γ Pharyngealized	$\overset{\circ}{\text{t}}^\Gamma$ $\overset{\circ}{\text{d}}^\Gamma$	¬ No audible release	$\overset{\circ}{\text{d}}^\neg$
..	Centralized	$\overset{\circ}{\text{ɛ}}$	~ Velarized or pharyngealized	$\overset{\circ}{\text{t}}$		
×	Mid-centralized	$\overset{\circ}{\text{ɛ}}$	↑ Raised	$\overset{\circ}{\text{e}}$ (J = voiced alveolar fricative)		
,	Syllabic	$\overset{\circ}{\text{n}}$	↓ Lowered	$\overset{\circ}{\text{e}}$ (β = voiced bilabial approximant)		
,	Non-syllabic	$\overset{\circ}{\text{e}}$	→ Advanced Tongue Root	$\overset{\circ}{\text{e}}$		
~	Rhoticity	$\overset{\circ}{\text{ə}}$ $\overset{\circ}{\text{a}}$	↔ Retracted Tongue Root	$\overset{\circ}{\text{e}}$		

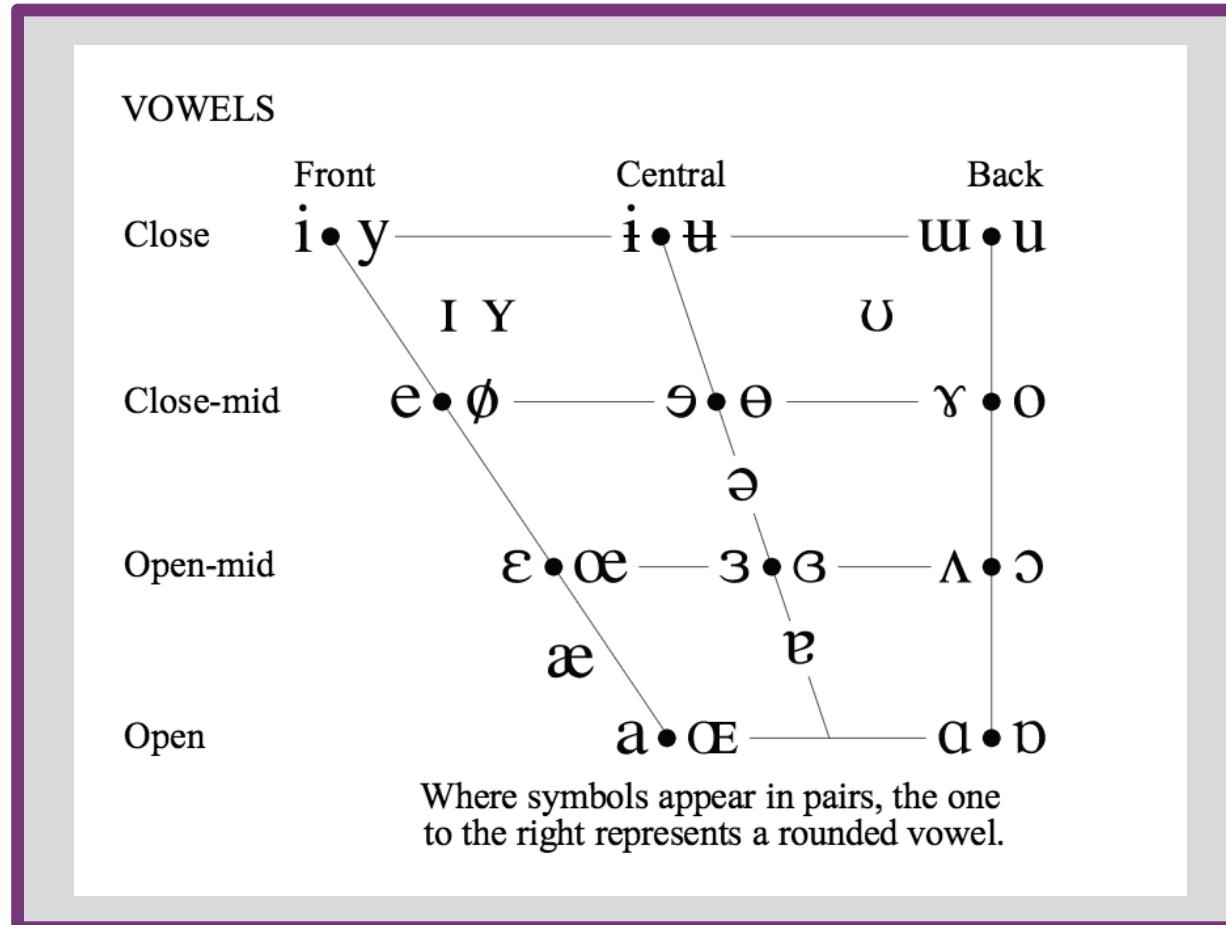
https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf



+ International Phonetic Alphabet

HLT11

■ The IPA Chart (Vowels)



https://www.internationalphoneticassociation.org/sites/default/files/IPA_Kiel_2015.pdf



+ English Phonemic Symbols

HILTII

■ www.phonemicchart.com

● - British vowels ○ - American vowels - show hints

single vowels				diphthongs			
I	i:	ʊ	u:	eɪ	ɔɪ	aɪ	
ship	sheep	book	shoot	wait	coin	like	
e	ɜː	ə	ɔː	eə	ɪə	ʊə	
left	her	teacher	door	hair	here	tourist	
æ	ʌ	ɒ	ɑː	əʊ	əʊ	/	
hat	up	on	far	show	mouth		

unvoiced consonants							
p	f	θ	t	s	ʃ	tʃ	k
pea	free	thing	tree	see	sheep	cheese	coin

voiced consonants							
b	v	ð	d	z	ʒ	dʒ	g
boat	video	this	dog	zoo	television	joke	go
m	n	ŋ	h	w	l	r	j



+ Other Alphabets

HLT11

■ Machine Readable Alphabets

SAMPA

See <https://www.phon.ucl.ac.uk/home/sampa/>
for European Languages

ARPABET

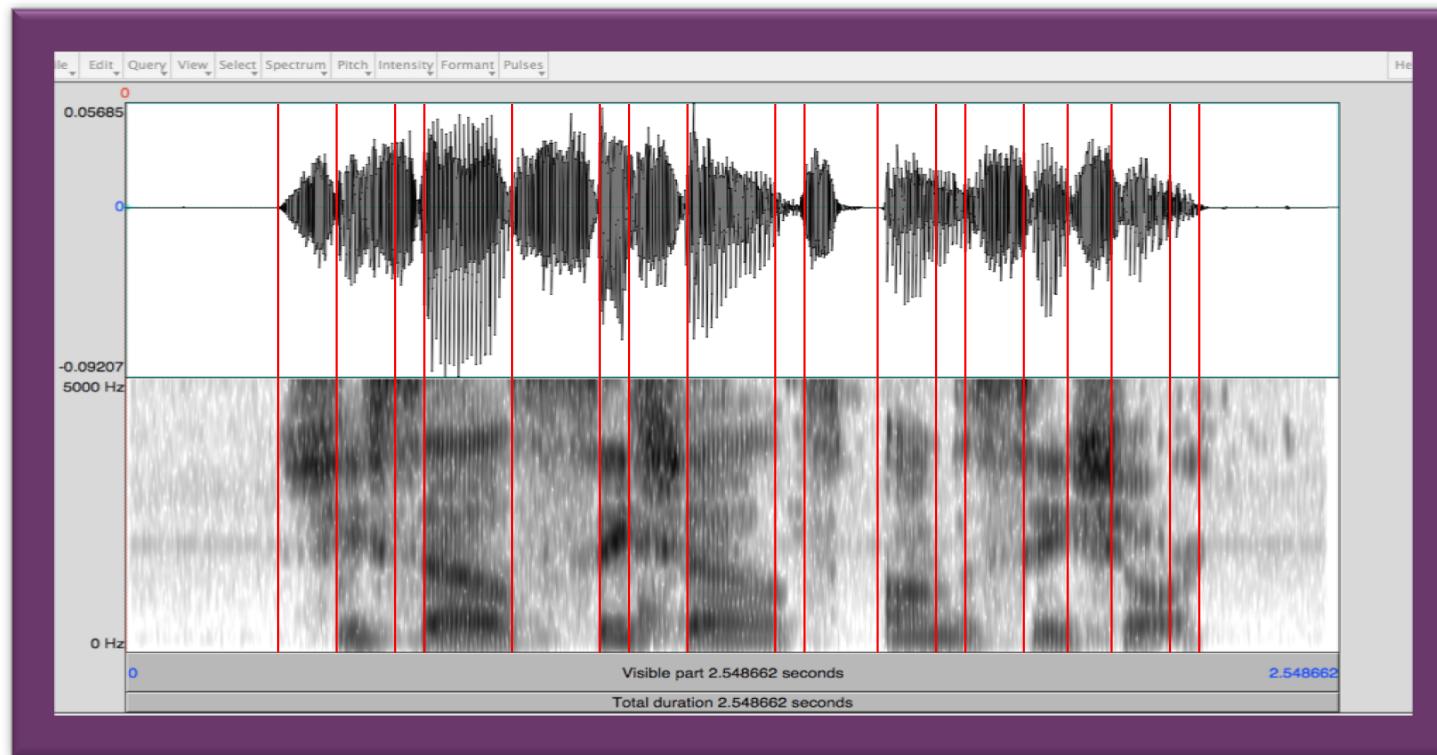
See <http://www.speech.cs.cmu.edu/cgi-bin/cmudict>
for American English



+ Sounds to Symbols

HT11

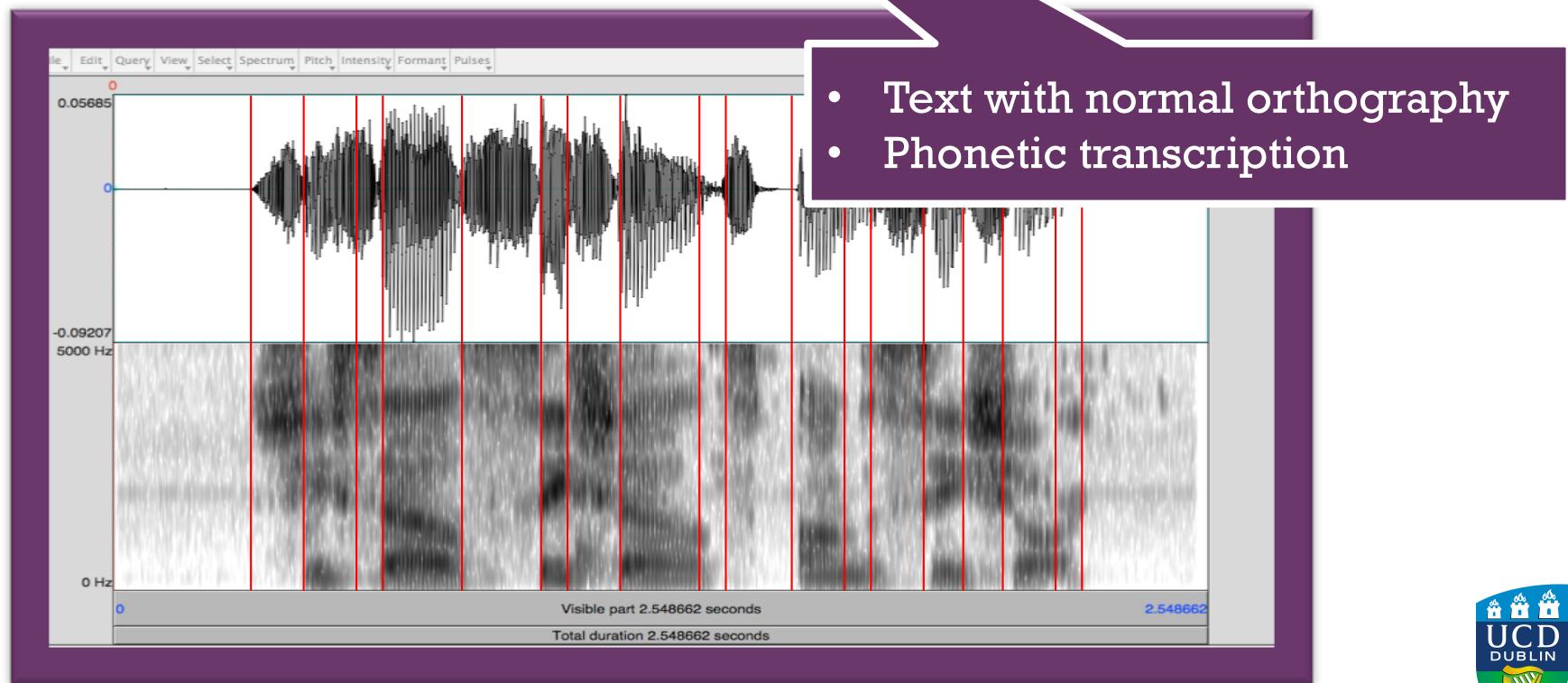
- A trained phonetician uses the speech signal, the audio file and a written text to manually annotate a speech utterance setting start and end points for each sound, or word label.



+ Sounds to Symbols

HT11

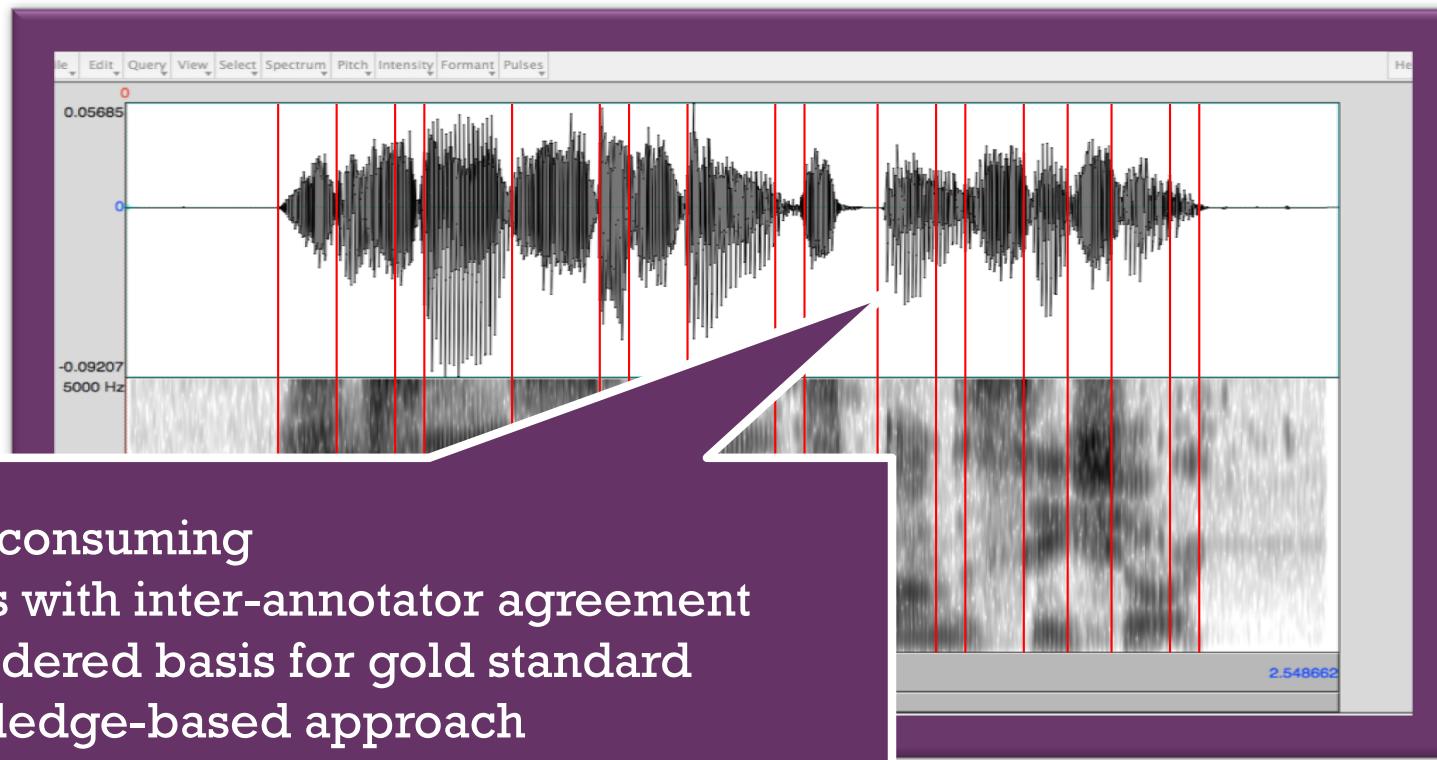
- A trained phonetician uses the speech signal, the audio file and a written text to manually annotate a speech utterance setting start and end points for each sound, or word label.



+ Sounds to Symbols

HLT11

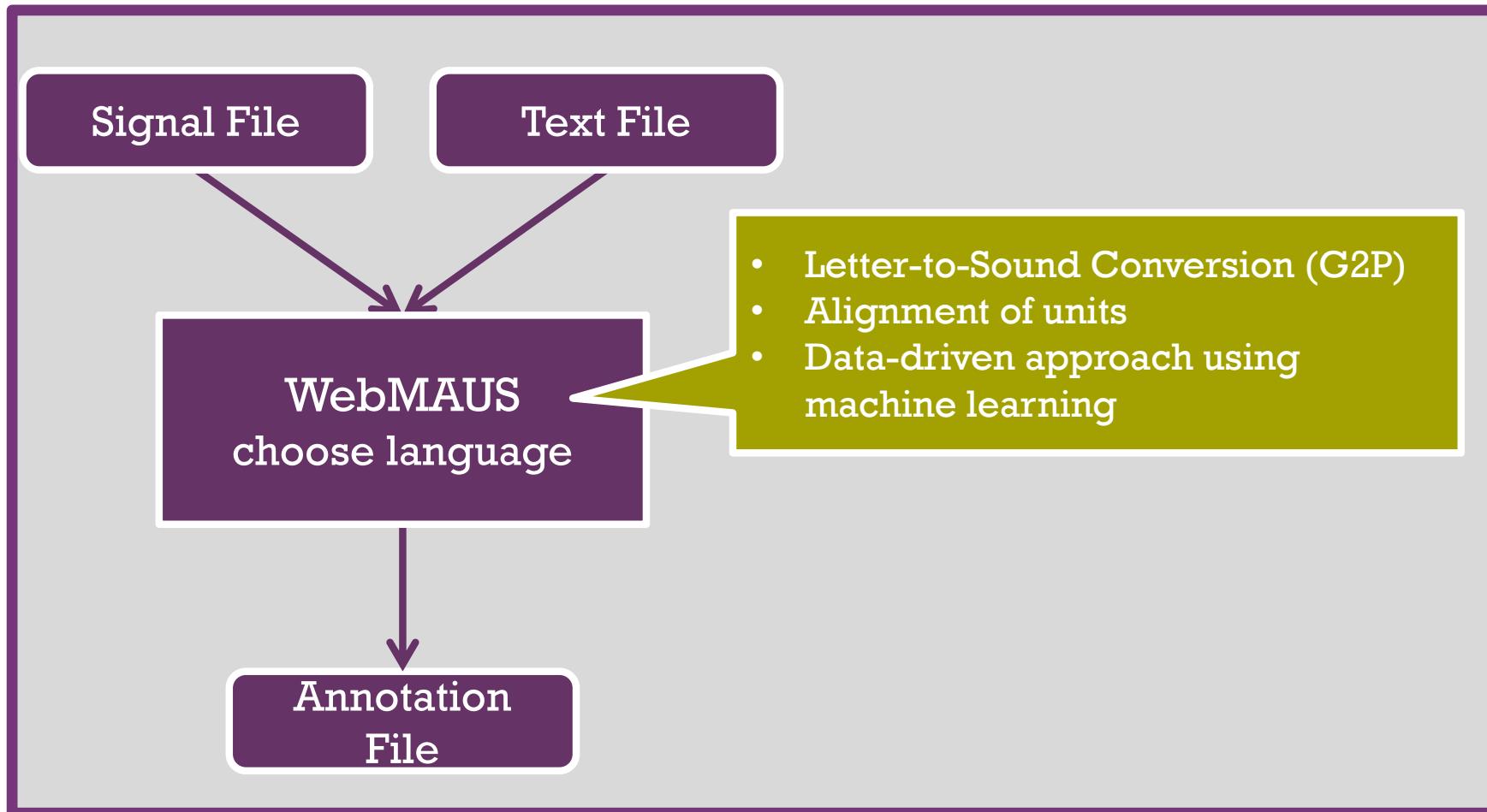
- A trained phonetician uses the speech signal, the audio file and a written text to manually annotate a speech utterance setting start and end points for each sound, or word label.



+ Sounds to Symbols

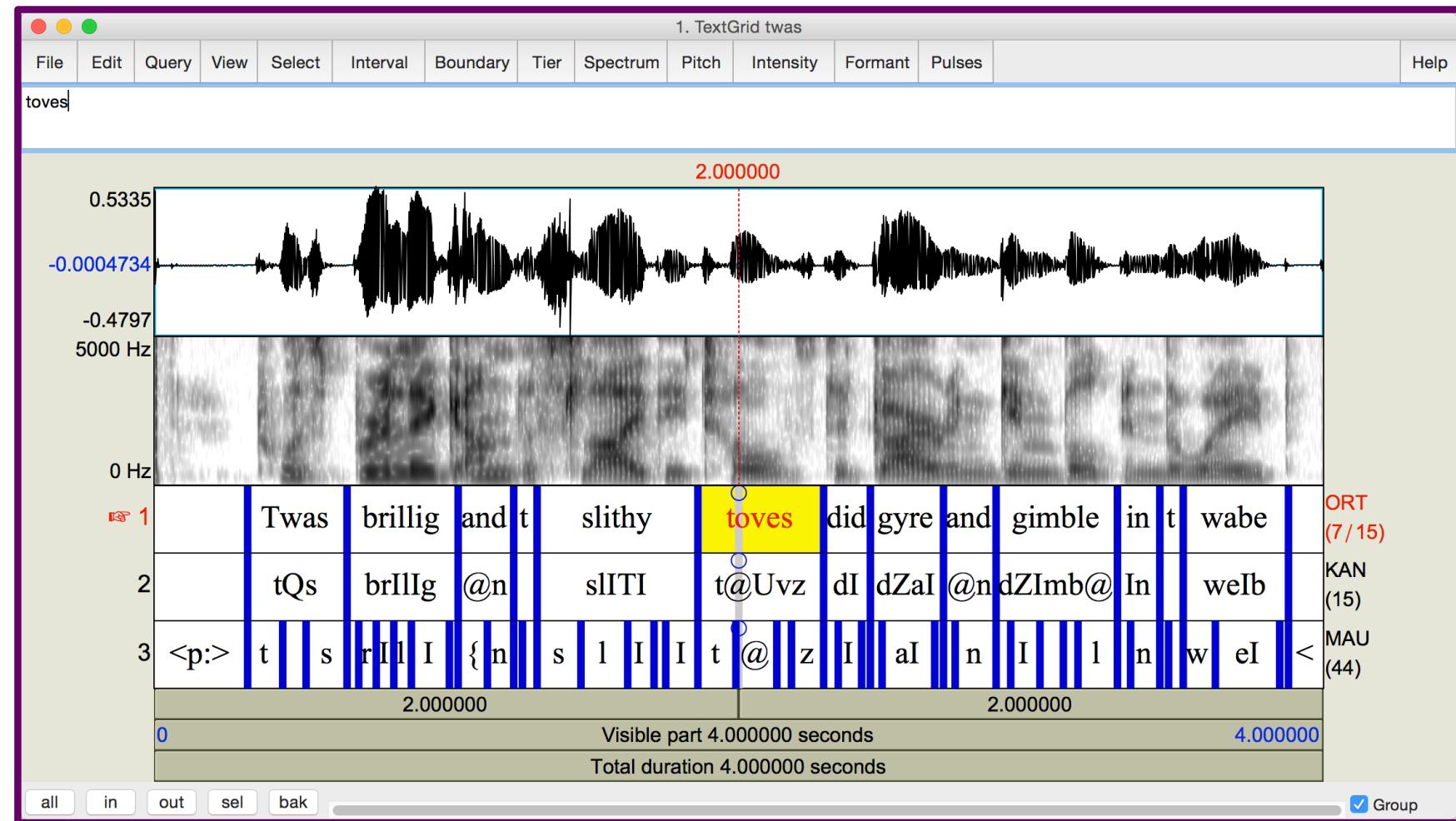
HLT11

- Alternatively, it is possible to force-align a text with an audio file using automatic alignment tools e.g. [WebMAUS](#) (see csmoodle link)



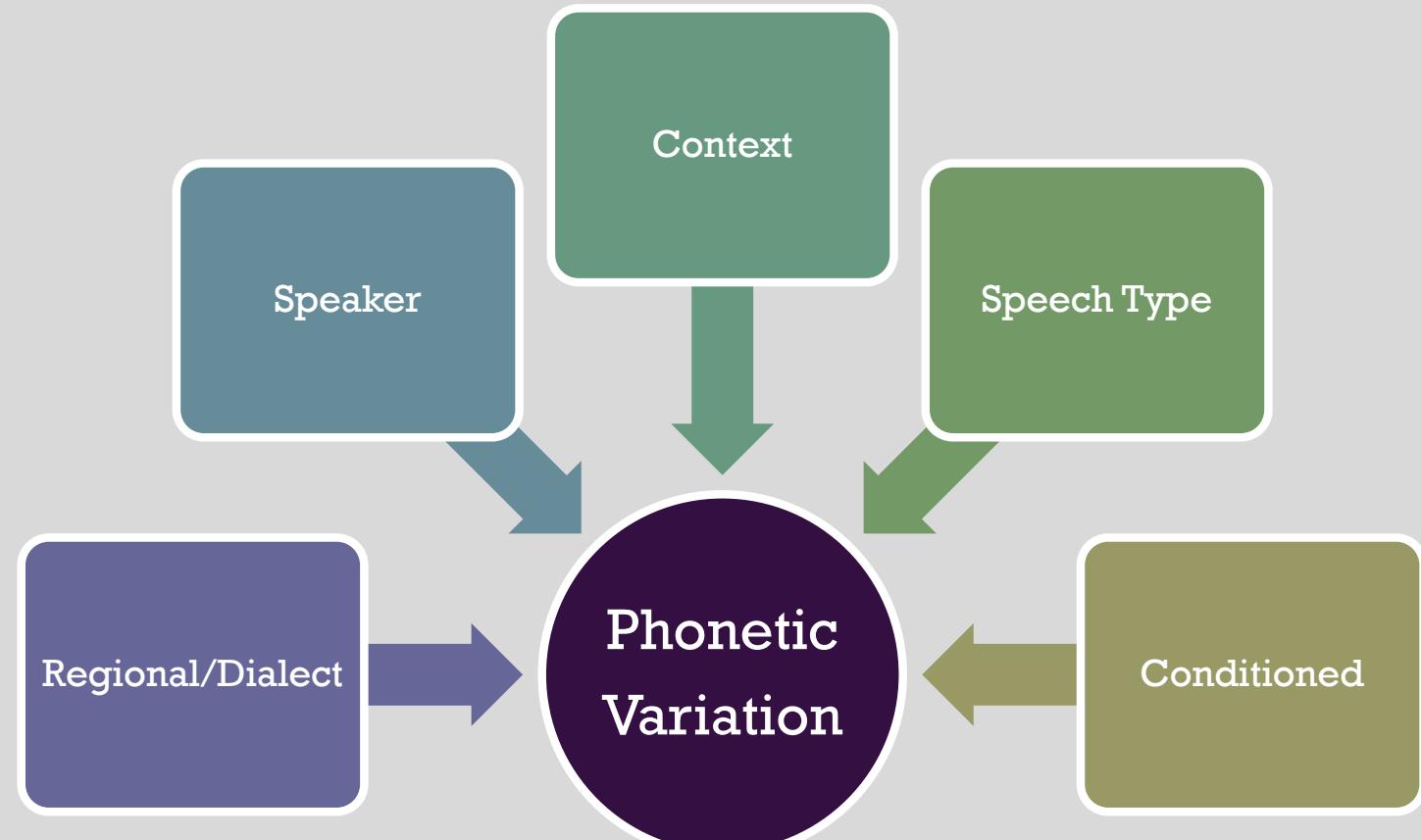
+ Sounds to Symbols

- The annotation file can then be viewed using a speech analysis tool such as [Praat](#) (see csmoodle link)



+ Sources of Variability

HLT11



+ Sources of Similarity: Cs & Vs

HLT11

scam	shop	health
comb	sixth	home
fool	photo	cannon
handy	mixed	it

- Number of sounds?
- Canonical forms?
- Syllable boundaries?



+ Sources of Similarity: Cs & Vs

HLT11

scam	shop	health
comb	sixth	home
fool	photo	cannon
handy	mixed	it

4	3	4
3	5	3
3	4	5
5	5	2

- Number of sounds?
- Canonical forms?
- Syllable boundaries?



+ Sources of Similarity: Cs & Vs

HLT11

scam	shop	health
comb	sixth	home
fool	photo	cannon
handy	mixed	it

- Number of sounds?
- Canonical forms?
- Syllable boundaries?

CCVC	CVC	CVCC
CVC	CVCCC	CVC
CVC	CVCV	CVCVC
CVCCV	CVCCC	VC



+ Sources of Similarity: Cs & Vs

HLT11

scam	shop	health
comb	sixth	home
fool	photo	cannon
handy	mixed	it

- Number of sounds?
- Canonical forms?
- Syllable boundaries?

CCVC	CVC	CVCC
CVC	CVCCC	CVC
CVC	CVCV	CV.C.VC
CVC.CV	CVCCC	VC



+ Phonemes, Phones, Allophones

HLT11

- **Phonology** is concerned with the *systematic organisation* of sounds within a language.
- **Phonemes** are sound segments that are distinctive, that contrast or distinguish words
 - /t/, /p/, /s/
- **Phones** are phonetic segments (i.e. the realisations of phonemes)
 - [t], [p], [s]
- **Allophones** are predictive (conditioned) phonetic variants of phonemes.
 - [t] in stop vs. [t^h] in top
- **Phonological rules** define how phonemes are realised as allophones in a given environment
 - /t/ → [t^h] / #__V



+ Phonemes, Phones, Allophones

HLT11

- Minimal pairs
 - cat:bat, mad:sad, bead:deed
- Complementary distribution
 - in mutually exclusive environments e.g. German: [ç] vs. [x]
 - but must be phonetically similar i.e. not English [h] vs. [ŋ]
- Free variation
 - [t] and [?] at the end of some English words
- Phonemes are organised systematically within a particular language i.e. there is an underlying grammar which defines the phonemes of the language and the way in which they can combine.



+ Systematic Organisation

HLT11

■ *blant*

- considered to be well-formed by a native speaker of English
→ idiosyncratic (accidental) gaps in the lexicon

■ *bnanlt*

- considered to be ill-formed by a native speaker of English
→ systematic gaps in the lexicon

- Statements of permissible combinations of phonemes in a language are known as its **phonotactics** (constraints on well-formedness of phonological structures)



+ Jabberwocky

HILT 1

'Twas brillig, and the slithy toves
Did gyre and gimble in the wabe:
All mimsy were the borogroves,
And the mome raths outgrabe.

"Beware the Jabberwock, my son!
The jaws that bite, the claws that catch!
Beware the Jubjub bird, and shun
The frumious Bandersnatch!"

He took his vorpal sword in hand:
Long time the manxome foe he sought
So rested he by the Tumtum tree,
And stood awhile in thought.

And, as in uffish thought he stood,
The Jabberwock, with eyes of flame,
Came whiffling through the tulgey wood,
And burbled as it came!

One, two! One, two! And through and through
The vorpal blade went snicker-snack!
He left it dead, and with its head
He went galumphing back.

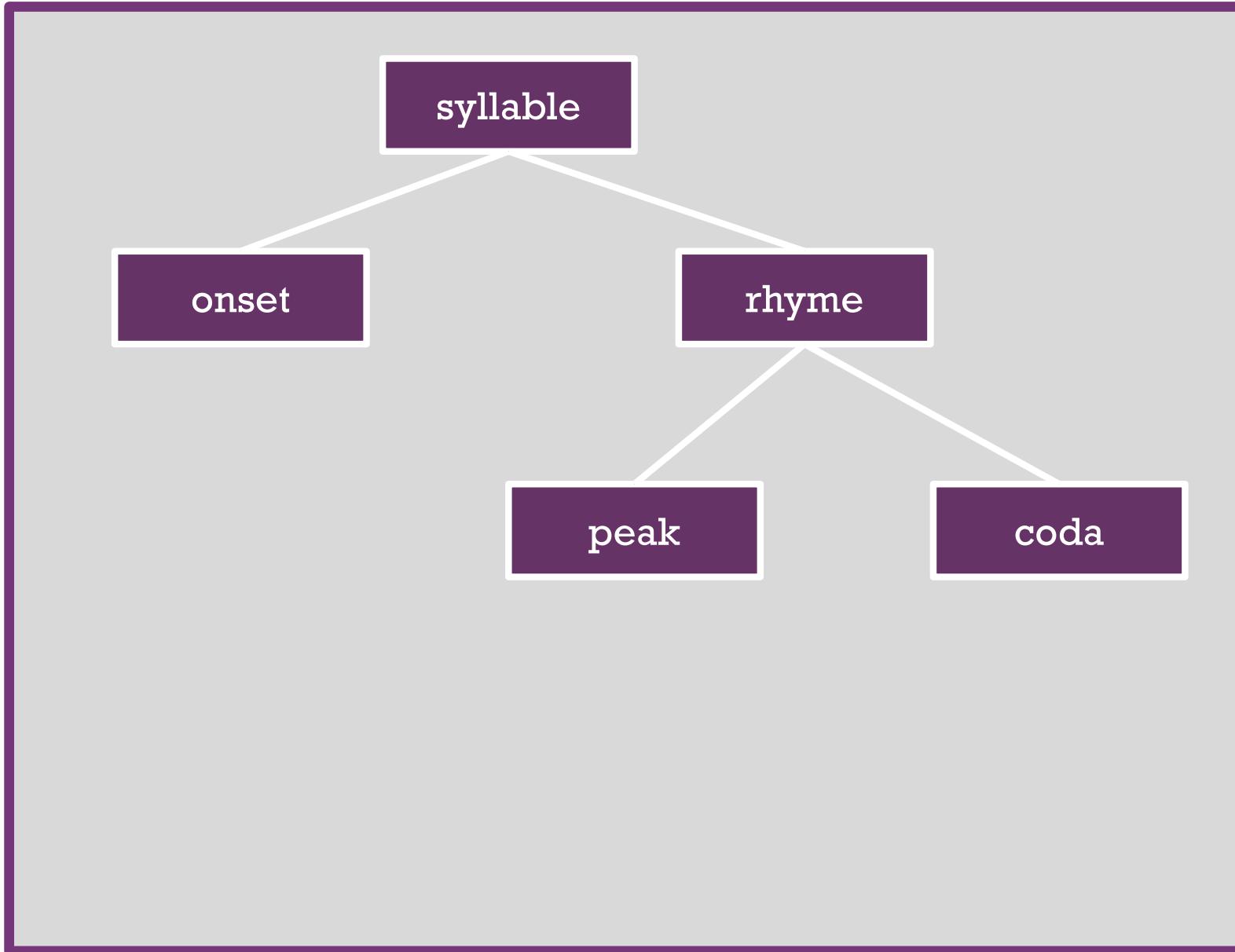
"And hast though slain the Jabberwock?
Come to my arms, my beamish boy!
O frabjous day! Callooh! Callay!"
He chortled in his joy.

Lewis Carroll (1871)



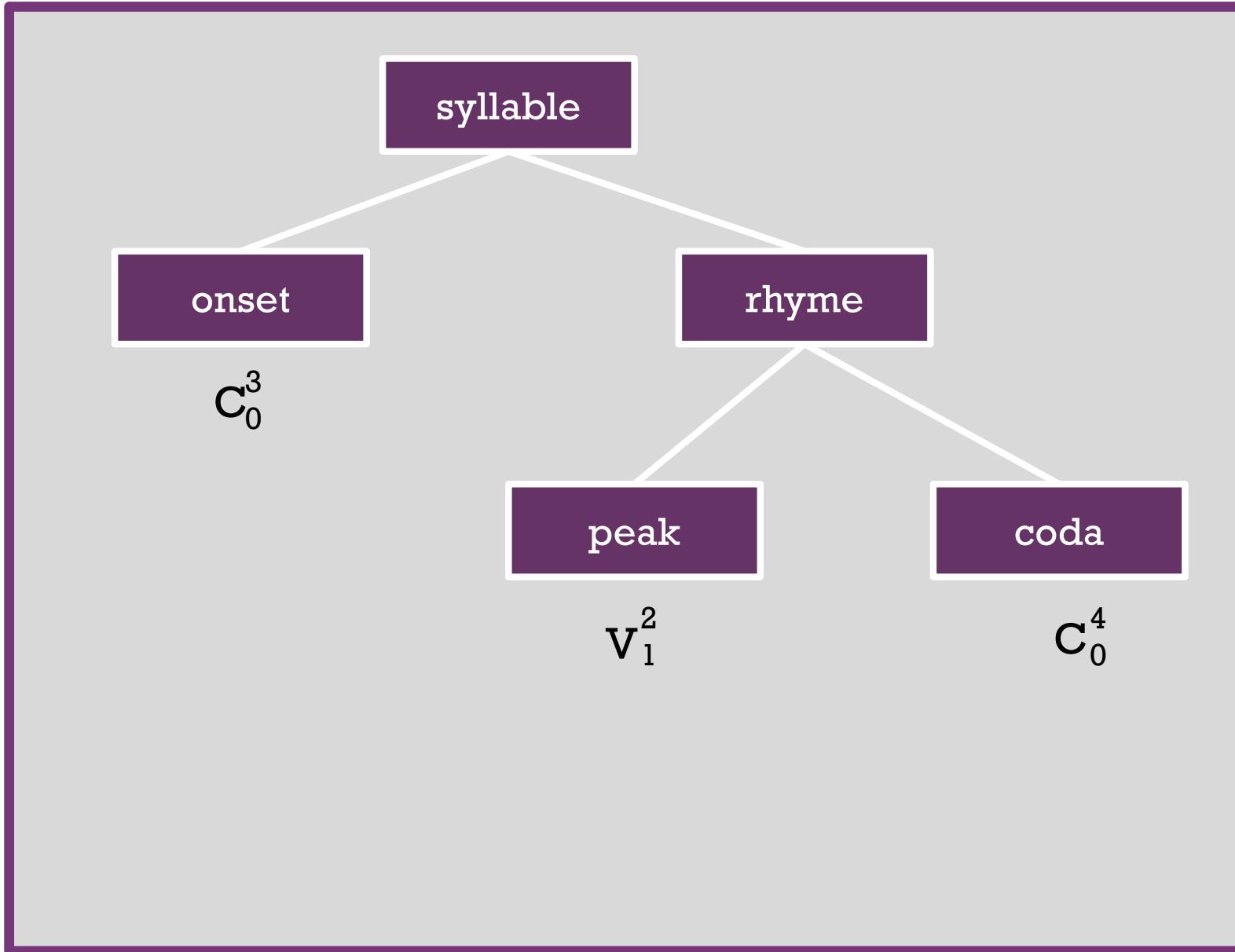
+ Systematic Organisation

HLT11



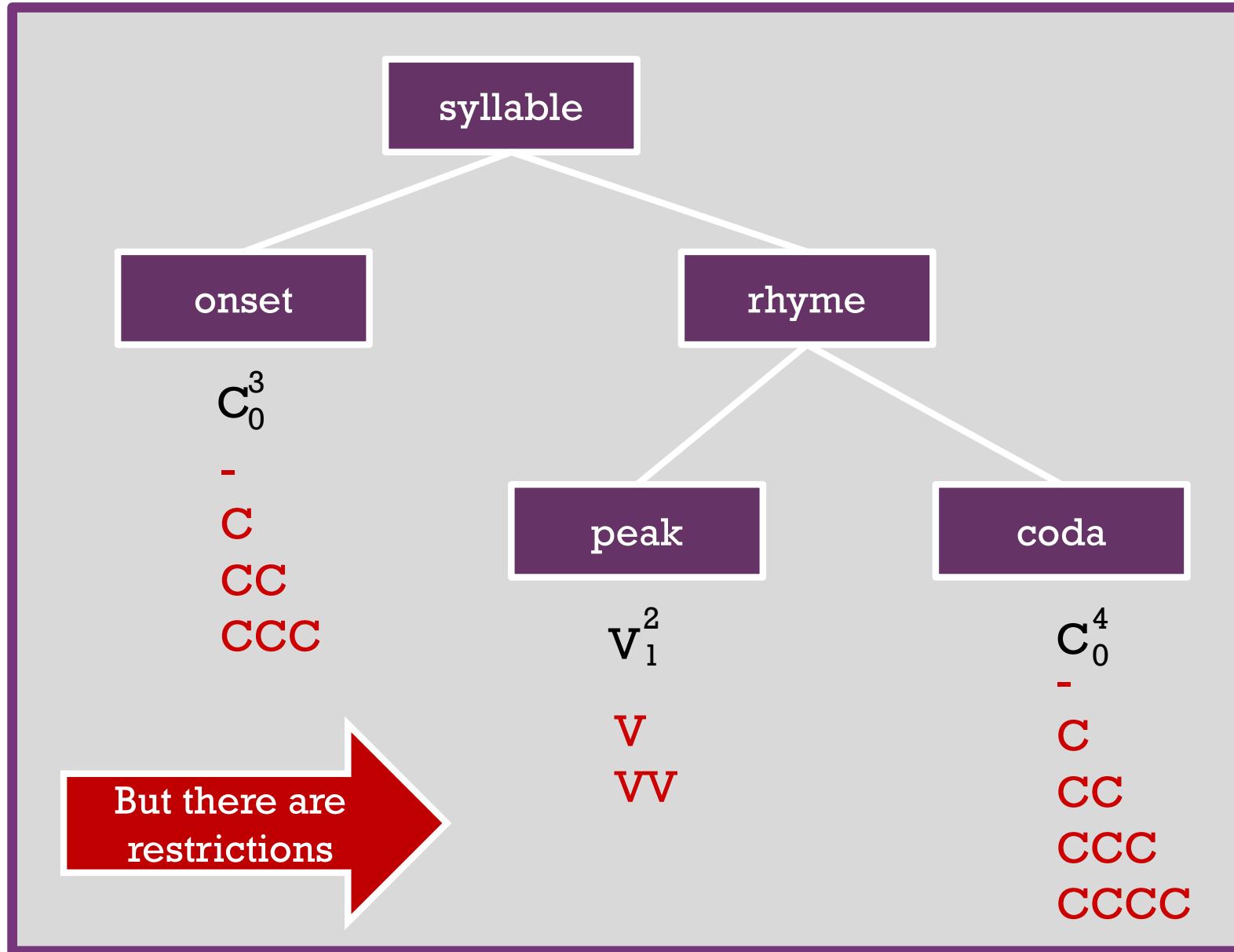
+ Systematic Organisation

HLT11



+ Systematic Organisation

HLT11



+ Systematic Organisation

HLT11

Some combinations in English onset:

	l	r
p	+	+
b	+	+
t		+
d		+
k	+	+
g	+	+

Some combinations in German coda:

	p	t	k	f	s	ʃ	m	n	ŋ	l	ʁ	x	ç
l	+	+	+	+	+	+	+	+	+				+
ʁ	+	+	+	+	+	+	+	+	+	+			+



+ Symbols to Features

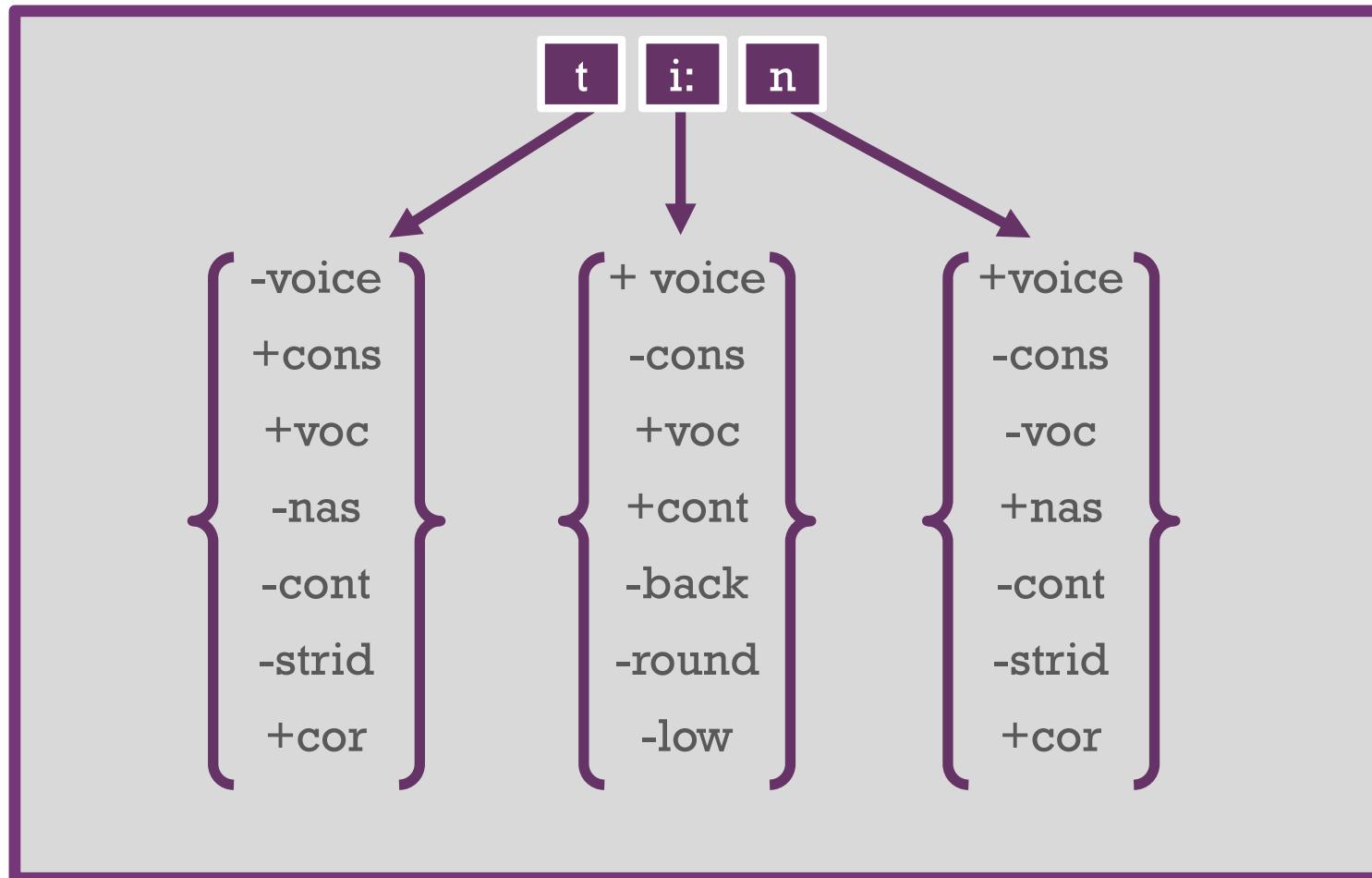
HLT11

- Phonemes have common properties and can be grouped into phonological classes according to the properties they possess e.g.
 - /p/, /t/, /k/ are all voiceless plosives
 - /m/, /n/, /ŋ/ are all nasals
- Phonological classes can be expressed in different ways using e.g.
 - Unary features
 - Binary features
 - Multi-valued features
- Often a distinction is made between a **linear** (segmental) approach and **multi-linear** (autosegmental or non-linear) approach.



+ Linear Approach

The word *teen* can be regarded as a linear sequence of phonemes (or segments):

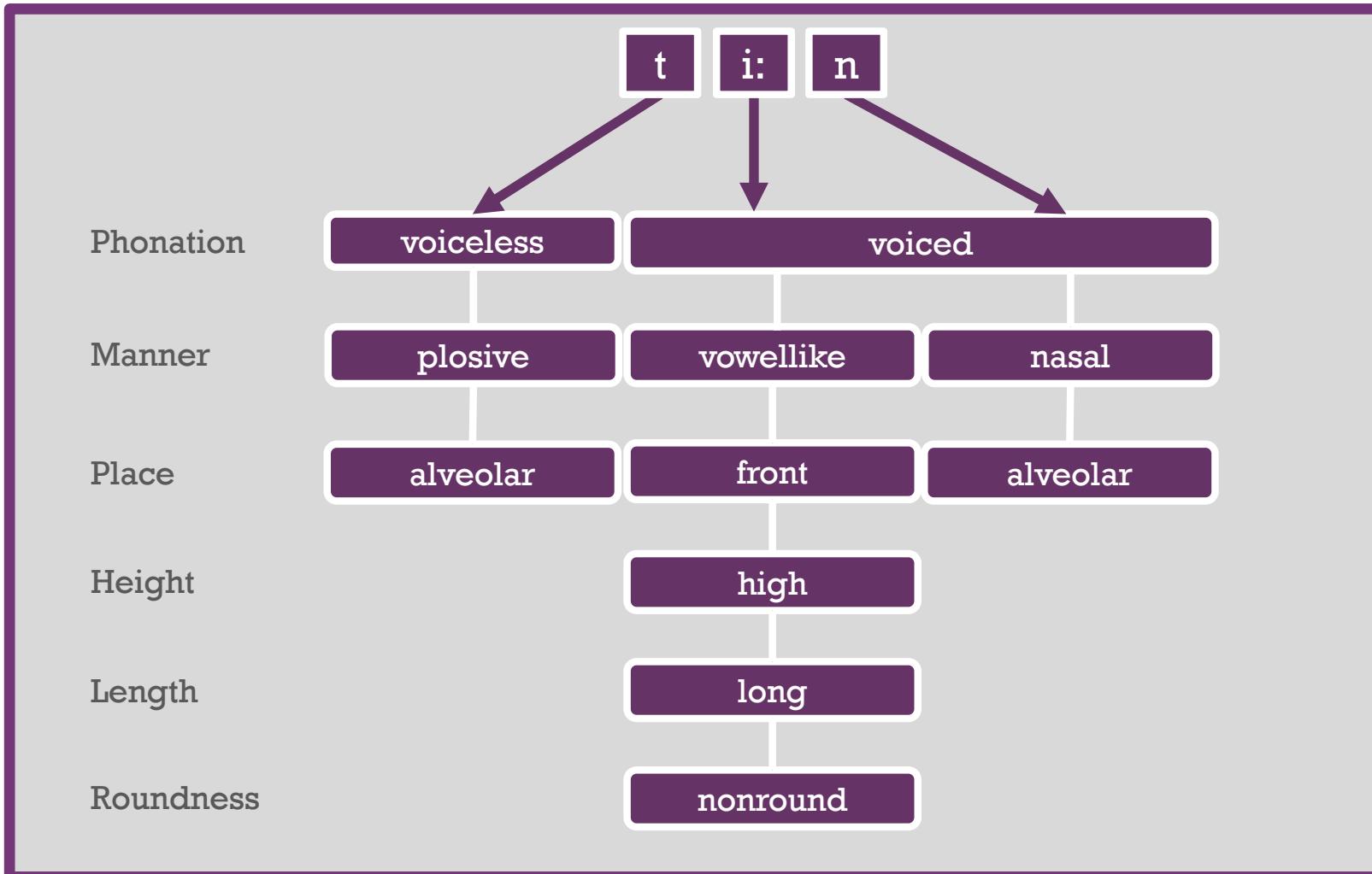


Segmental Representation with Binary Features

+ Multi-linear Approach

HLT11

The word *teen* can also be regarded as a multi-linear representation of overlapping autonomous feature segments:



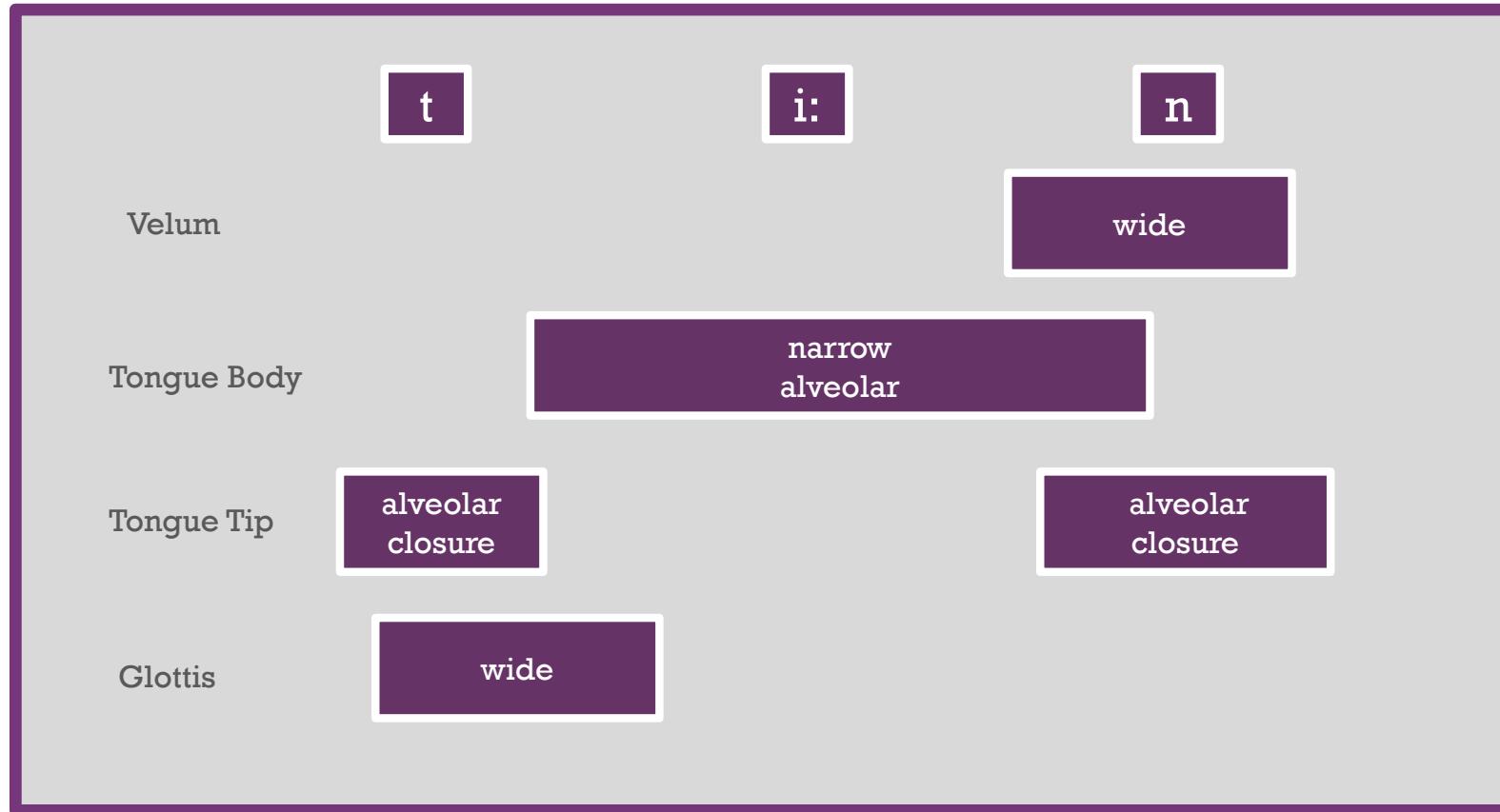
“Autosegmental” Representation with Multi-Valued Features



+ Multi-linear Approach

HILT11

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

