



LINGUISTICS SLT BACKGROUND

Phonetics & Phonology

CONTENTS

OVERVIEW	Distinguishing phonetics & phonology
ARTICULATORY PHONETICS	Place/Manner of Articulation, the IPA chart, vowel quadrilaterals, etc.
PHONOLOGICAL RULES	Epenthesis, insertion, metathesis, assimilation, etc.
SOCIOPHONETICS	Accents, dialects, social class, gender, identity, etc.
Applications to SLT	G2P, phonological edit distance, TTS, humour, etc.
RESOURCES	Links to additional external resources to learn more

INTRODUCTION

PHONETICS

- **The “phone”** - The central unit to phonetics encompassing any distinct speech sound regardless of meaning
- **Observable** - Deals with measurable characteristics of speech
- **Production & Perception of speech**
- **Subdisciplines:** Articulatory phonetics, acoustic phonetics, auditory phonetics, etc.



TRANSCRIPTION

- **Phones:** [t, s, i:]
- **Phonemes:** /t, s, i:/
- Brackets are important!

PHONOLOGY

- **The “phoneme”** - The central unit to phonology encompassing speech sounds that are contrastive and singular meaning
- **Abstract/Theoretical** - Deals with assumed mental representations of speech sounds
- **Sound inventories and patterns within a language**
- **Subdisciplines:** Phonotactics (sound combinations), distinctive features, Optimality Theory, etc.

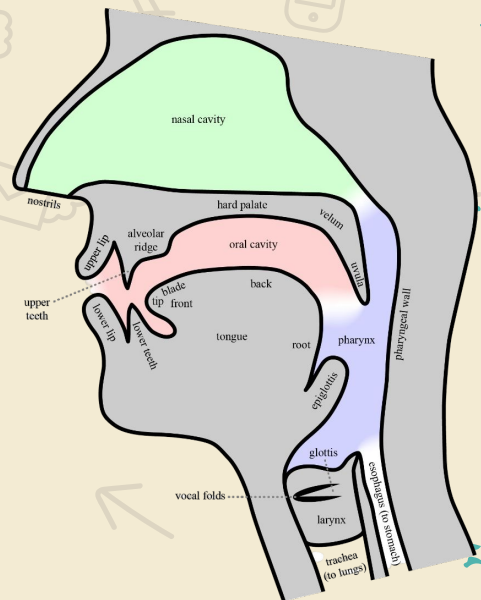
THE IPA CHART

- The International Phonetic Alphabet is a phonetic transcription standard.
- As part of this, there is an “IPA Chart” which categorises phones into place/manner of articulation (and voicing!).
- **Place** - The position of the primary articulators (e.g. tongue, lips, teeth...) during articulation
- **Manner** - The method of production for a particular sound (e.g. plosives, fricatives, affricates, nasals...)
- **Voicing** - Whether or not the vocal folds (aka., glottal folds or “vocal cords”) are vibrating

THE IPA CHART (CONSONANTS)

- **Columns:** Place of articulation (from front of mouth to the back)
- **Rows:** Manner of articulation
- Symbols to the left are unvoiced, and to the right are voiced
- Missing symbols represent sounds that are *plausible* but not known in any language
- Shaded cells are sounds judged to be impossible to articulate

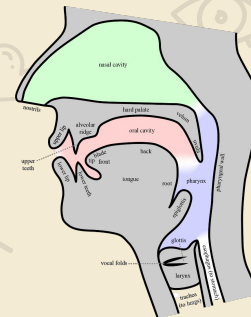
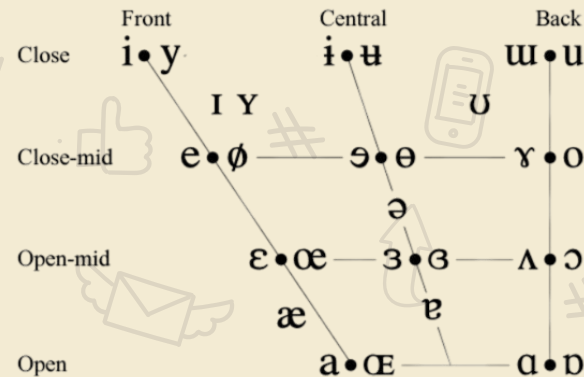
	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill	ʙ			r					ʀ		
Tap or Flap		ɸ		ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			



THE IPA CHART (VOWELS)

- **Vowel Quadrilateral** - Displays the vowels of human speech on a diagram representing the oral cavity.
- **X-Axis** - “Backness”, or how far back in the mouth the tongue is
- **Y-Axis** - Vowel height, or how high/low in the mouth the tongue is
- **Rounding** - Whether or not the lips are in a rounded position during articulation
- Symbols to the left are unrounded, and those to the right are rounded
- **Diphthongs** - Sounds formed by combining two vowels into one syllable, e.g. “kite” [kʌɪt], “boat” [bəʊt], “boy” [bɔɪ]*

*These are examples of one possible pronunciation



DISTINCTIVE FEATURES

- **Categorisation of phonological features to distinguish different phonemes and denote phonological rules**
- **Based on:** Acoustics, articulation, suprasegmentals (syllable stress etc.)
- **Subcategories:** Major Class Features, Laryngeal Features, Manner Features, Place Features, Vowel Space Features (...)
- Much more extensive than the list shown
- **Present** = “+”, **Absent** = “-”, **N/A** = “0”

Absence of turbulent
airflow during
production

Air passes through
vocal tract in a
continuous stream

	Consonantal	Sonorant	Voiced	Nasal	Continuant	Coronal
/t/	+	-	-	-	-	+
/d/	+	-	+	-	-	+
/s/	+	-	-	-	+	+
/l/	+	+	+	-	+	-
/i/	=	+	+	-	+	-

PHONOLOGICAL RULES

- **Phonology accounts for the many rules that govern how phonemes interact in speech, especially during coarticulation**
- **Coarticulation** - Changes to the speech signal due to preceding or following sounds
- Multiple rules may apply in series or in parallel following ordering rules (i.e. feeding, counterfeeding, bleeding, and counterbleeding)

Processes

- **Assimilation** - Sounds become more similar to each other by sharing features
- **Final-Obstruent Devoicing** - Voiced obstruents (e.g. plosives) at the end of a word become unvoiced
- **Epenthesis** - The addition of 1 or more sounds
- **Metathesis** - Sounds/Syllables swap position
- (...)

Notation (simplified)

- **A → B / #_**
- **A** = Original phoneme/feature matrix
- **B** = New phoneme/feature matrix
- **#** = Word boundary
- **_** = Position where new phoneme/feature matrix occurs
- **i.e. "A becomes B following a word boundary"**

Real Example (Final Obstruent Devoicing - German)

- *C[+ obstruent, +voice] → C[-voice] / __#.
- Consonants that are voiced obstruents become unvoiced consonants before a word boundary (i.e. at the end of a word)
- "Bad" /ba:t/ vs. "Bäder" /'bɛ:de/ ("bath"/"baths")

SOCIOPHONETICS

Speakers of a given language do not all sound the same

Some differences are systematic and based on identity (gender, sexuality, age) and location (i.e., accents)

American English (US)

- ~30 distinct accents

British English (UK)

- ~40 distinct accents

The UK has more accents, despite being ~40x smaller...

Additionally, English is a common additional language, leading to many further accent variations due to interaction with native language phonology and international varieties (e.g. "Indian English").



Applications to SLT

- **Grapheme-to-Phoneme conversion (G2P)**- text-to-text models for phonetic transcription of standard text
 - Fundamental to most NLP tasks handling phonetics/phonology
 - Pronunciation dictionaries for ASR/TTS (e.g. domain-specific terminology with low occurrence in training data)
- **Phonetic Edit Distance** (e.g. the PanPhon package) - Quantify the phonemic/phonetic differences between words/sounds
 - Improve autocorrect (common misspellings based on phonetic similarity)
 - Evaluation metrics (e.g. tongue twister generation)
- **Creative Language Generation/Detection**
 - Homographic puns (where spelling remains the same but pronunciation differs (e.g., “record”, “bow” etc)
 - Adhering to rhyme schemes and syllable structure for poetry/lyric generation
- **Speech Assessment**
 - Detecting phonological processes (i.e., rules) to screen people for speech therapy and categorising specific errors
 - Provide articulation instructions for language-learners (e.g. Duolingo and Babbel)
 - Healthcare applications to detecting COVID and cognitive decline
- **Accessibility of SLT**
 - Understand the nuance/variation of human speech and formally identify speaker characteristics (e.g., accents) where systems fail