Phonetics Analytics in DROP

v5.21 18 October 2021

Phonetics – Introduction

Overview

1. Definition of Phonetics and Phoneticians: *Phonetics* is a branch of linguistics that studies how humans perceive and product sounds, or in the case of sign languages, the equivalent aspects of sign (O'Grady (2005), Wikipedia (2021)). Phoneticians – linguists who specialize in phonetics – study the physical properties of speech.
2. Disciplines that Comprise Phonetics Study: The field of phonetics is traditionally divided into three sub-disciplines based on the research questions involved such as how humans plan and execute movements to produce speech – articulatory phonetics; how different movements affect the properties of the resulting sound – acoustic phonetics; or how humans convert sound waves to linguistic movements – auditory phonetics.
3. Minimal Linguistic vs Phonological Unit: Traditionally, the minimal linguistic unit of phonetics has been the phone – a speech sound in a language – which differs from the minimal phonological unit of phoneme; the phoneme is an abstract categorization of phones.
4. Speech Perception/Production in Languages: Phonetics broadly deals with two aspects of human speech; a) the way humans make sounds, and b) perception – the way speech is understood. The communication modality of a language describes the method by which a language produces and perceives speech.
5. Languages with Oral-Aural Modalities: Languages with oral-aural modalities such as English produce speech orally; i.e., using the mouth, and perceive speech aurally – using the ears.
6. Languages with Manual Visual Modalities: Sign languages, such Auslan and ASL, have a manual-visual modality, producing speech manually – using the hands – and perceiving speech visually – using the eyes.

7. Languages with Tactile Signing Modality: ASL and some other sign languages in addition have a manual-manual dialect for use in tactile signing in deafblind speakers where the signs are produced by the hands and perceived by the hands as well.
8. Non-linguistic to Speech Translation: Language production consists of several independent processes which transform a non-linguistic message into a spoken or signed linguistic signal.
9. Lexical Selection - Choosing Word Items: After identifying a message to be linguistically encoded, a speaker must select the individual words – known as lexical items – to represent that message in a process called lexical selection.
10. Assignment of Words to Phonemes: During phonological encoding, the mental representation of the words is assigned their phonological content as a sequence of phonemes to be produced. These phonemes are specified for articulatory features which denote particular goals such as closed lips or the tongue in a particular position.
11. Phonemes as Muscle Command Sequence: These phonemes are then coordinated into a sequence of muscle commands that can be sent to muscles, and when these commands are executed properly the intended sounds are produced.
12. Airstream Disruption/Modification using Articulators: These movements disrupt and modify an airstream which results in a sound wave. The modification is done by articulators, with different places and manners of articulation producing different results.
13. Places and Manner of Articulation: For example, the words *task* and *sack* both begin with alveolar sounds in English, but differ in how far the tongue is from the alveolar ridge. The difference has large effects on the airstream and thus the sound that is produced. Similarly, the direction and the source of the airstream can affect the sound.
14. Pulmonic/Glottal/Lingual Airstream Modification: The most common airstream mechanism is pulmonic – using the lungs – but the glottis and the lungs can also be used to produce airstreams.
15. Decoding Signals into Linguistic Units: Language perception is the process by which a linguistic signal is decoded and understood by the listener. In order to perceive speech, the continuous acoustic signal must be converted into discrete linguistic units such as phonemes, morphemes, and words.

16. Detection of the Linguistic Categories: In order to correctly identify and categorize sounds, listeners prioritize certain aspects of the signal that can reliably distinguish between linguistic categories.
17. Visual Information Augmenting Acoustic Cues: While certain cues are prioritized over others, many aspects of the signal can contribute to perception. For example, though oral languages prioritize acoustic information, the McGurk effect shows that the visual information is used to distinguish ambiguous information when acoustic ones are unreliable.
18. Articulatory Phonetics: Modern phonetics has three main branches. The first, articulatory phonetics, studies the way sounds are made with the articulators.
19. Acoustic Phonetics: This studies the acoustic results of different articulations.
20. Auditory Phonetics: This studies the way listeners perceive and understand linguistic signals.

Production

1. Sequential Steps of Speech Production: Language production consists of several inter-dependent processes which transform a non-linguistic message into a spoken or signed linguistic signal. Linguists debate whether the process of language production occurs in a series of stages, i.e., serial processing, or whether production processes occur in parallel.
2. Linguistic Encoding through Lexical Selection: After identifying the message to be linguistically encoded, a speaker must select the individual words – known as lexical items – to represent that message in a process called lexical selection.
3. Word's Lemma - Semantic/Grammatic Determination: The words are selected based on their meaning, which in linguistics is called semantic information. Lexical selection activates the word's lemma, which contains both semantic and grammatic information about the word (Dell and O'Seaghdha (1992)). Again, linguists debate whether these stages can interactor whether they occur serially – for e.g., compare Motley, Camden, and Baars (1982) with Dell and O'Seaghdha (1992). For ease of description, the language production process is this chapter is described as a series of independent stages, though recent evidence shows that this

is inaccurate (Sedivy (2019)). Jaeger, Furth, and Hilliard (2012) contain further description of the interactive activation models.

4. Transferring Lexical Words to Phonemes: After an utterance has been planned – or after part of an utterance has been planned (Gleitman, January, Nappa, Trueswell (2007) provide evidence for production before a message has been completely planned), it then goes through phonological encoding. In this stage of language production, the mental representation of the words is assigned their phonological content as a sequence of phonemes.
5. Transformation to Muscle Movement Commands: The phonemes are specified for articulatory features which denote particular goals such as closed lips or tongue in a particular location. These phonemes are then coordinated into a sequence of muscle commands that can be sent to the muscles, and when these commands are executed properly the intended sound are produced (Boersma (1998)).
6. Message To Sound Transformation Summary: Thus, the process of production from message to sound can be summarized as the following sequence (Boersma (1998), Sedivy (2019)):
    a. Message Planning
    b. Lemma Selection
    c. Retrieval and Assignment of Phonological Word Forms
    d. Articulatory Specification
    e. Muscle Commands
    f. Articulation
    g. Speech Sounds

Place of Articulation

1. Phonetic Definition of a Consonant: Sounds that are made by a full or a partial constriction of the vowel tract are called consonants. Consonants are produced in the vocal tract, usually the mouth, and the location of this constriction affects the resulting sound.

2. Tongue Position Impact on Sound: Because of the close connection between the position of the tongue and the resulting sound, the place of articulation is an important concept in many sub-disciplines of phonetics.
3. Constriction Articulator and its Location: Sounds are partly categorized by the location of the constriction as well as the part of the body doing the constricting. For example, in English the words *fought* and *thought* are a minimal pair differing only in the organ making the constriction rather than the location of the constriction.
4. Example: Labiodental vs. Linguolabial Articulation: The *f* in *fought* is a labiodental articulation made with the bottom lip against the teeth. The *th* in *thought* is a linguolabial articulation made by the tongue against the teeth. Constrictions made by the lips are called labials while those made with the tongue are called linguals.
5. Places of Articulation on Tongue: Constrictions made in the tongue can be made in several parts of the vocal tract, broadly classified into coronal, dorsal, and radical places of articulation. Coronal articulations are made with the front of the tongue, dorsal articulations are made with the body of the tongue, and radical articulations are made in the pharynx (Ladefoged and Johnson (2011)).
6. Fine Grained Places of Articulation: These divisions are not sufficient for distinguishing and describing all speech sounds (Ladefoged and Johnson (2011)). For example, in English sounds [*s*] and [ʃ] are both coronal, but they are produced are different places in the mouth. To account for this, more detailed places of articulation are needed based upon the area of the mouth in which the constriction occurs (Ladefoged and Maddieson (1996)).

Labial

1. Three Types of Labial Articulators: Articulators involving lips can be made in three different ways: with both lips (bilabial), with one lip and the teeth (labiodental), and with the teeth and the upper lip (linguolabial) (Ladefoged and Maddieson (1996)). Depending on the definition

used, some or all kinds of articulations can be categorized into the class of labial articulations.

2. Lip Movements in Bilabial Consonants: Bilabial consonants are made with both lips. In producing these sounds the lower lip moves farthest to meet the upper lip, which also moves down slightly (Maddieson (1993)), though in some cases the force from the sir moving from the aperture – the opening between the lips – may cause the lips to separate faster than they can come together (Fujimura (1961)).
3. Incomplete Closures using Bilabial Articulators: Unlike most other articulations, both articulators are made from soft issue, and so bilabial stops are more likely to be produced with incomplete closures than articulations involving hard surfaces like teeth or palate.
4. Active Movement of the Upper Articulator: Bilabial stops are also unusual in that an articulator in the upper section of the vocal tract actively moves downwards, as the upper lip shows some active downward movement (Ladefoged and Maddieson (1996)).
5. Lip Movements in Linguolabial Consonants: Linguolabial consonants are made with the blade of the tongue approaching and contacting the upper lip. As in bilabial articulations, the upper lip moves slightly towards the mor active articulator.
6. IPA Symbols for Linguolabial Consonants: Articulations in this group do not have their own symbols in the International Phonetic Alphabet, rather, they are formed by combining an apical symbol with a diacritic implicitly placing them in the coronal category (Ladefoged and Maddieson (1996), International Phonetic Association (2015)). They exist in a number of languages indigenous to Vanuatu such as Tangoa.
7. Lip Movements in Labial Consonants: Labiodental consonants are made by the lower lip rising to the upper teeth. Labiodental consonants are most often fricatives while labiodental nasals are also typically common (Ladefoged and Maddieson (1996)).
8. Occurrence of True Labiodental Plosives: There is a debate as to whether true labiodental plosives occur in an any natural languages, though a number of languages are reported to have labiodental plosives including Zulu (Doke (1926)), Tonga (Guthrie (1948)), and Shubi (Ladefoged and Maddieson (1996)).

Coronal

1. Tongue Positions of Coronal Consonants: Coronal consonants are made with the tip or the blade of the tongue and, because of the agility of the front of the tongue, represent a variety not only in place but also in the posture of the tongue.
2. Coronal Consonants Places of Articulation: The coronal places of articulation represent the areas of the mouth where the tongue contacts of makes a constriction, and include dental, alveolar, and post-alveolar locations.
3. Apical vs Laminal Tongue Posture: Tongue postures using the tip of the tongue can be apical if using the top of the tongue tip, laminal if made with the blade of the tongue, or sub-apical if the tongue tip is curled back and the bottom of the tongue is used.
4. Widespread Coronal Manner of Articulation: Coronals are unique as a group in that every manner of articulation is attested (Ladefoged and Maddieson (1996), International Phonetic Association (2015)). Australian languages are well-known for a large number of coronal contrasts exhibited within and across languages in the region (Ladefoged and Maddieson (1996)).
5. Production/Classification of Dental Consonants: Dental consonants are made with the tip or the blade of the tongue and the upper teeth. They are divided into two groups based upon the part of the tongue used to produce them: apical dental consonants are produced with the tongue touching the teeth; interdental consonants are produced with the blade of the tongue as the tip of the tongue sticks out in front of the teeth. No language is known to use both contrastively though they can exist allophonically.
6. Production of the Alveolar Consonants: Alveolar consonants are made with the tip or the blade of the tongue at the alveolar ridge just behind the teeth and can similarly be apical or laminal (Ladefoged and Maddieson (1996)).
7. Variations among Dental/Alveolar Consonants: Cross-linguistically, dental consonants and alveolar consonants are frequently contrasted leading to a number of generalizations of cross-linguistic patterns. The different places of articulation also tend to be contrasted in the part of the tongue used to produce them: most languages with dental stops have laminal dentals, while languages with apical stops usually have apical dentals.

8. Laminality Based Contrast in Languages: Languages rarely have two consonants in the same place with a contrast in laminality, though Taa (!*Xóõ*) is a counter-example to this pattern (Ladefoged and Maddieson (1996)).
9. Mutually Exclusive Dental/Alveolar Stops: If a language has only one of a dental or an alveolar stop, it will usually be laminal if it is a dental stop, and the stop will usually be apical if it is an alveolar stop, though for example Temne and Bulgarian (Scatton (1984)) do not follow this pattern (Ladefoged and Maddieson (1996)).
10. Languages with Dental/Alveolar Stops: If a language has both an apical and a laminal stop, the laminal stop is more likely to be affricated as in Isoko, though Dahalo shows the opposite pattern with alveolar stops being more affricated (Ladefoged and Maddieson (1996)).
11. Varying Definitions of Retroflex Consonants: Retroflex consonants have several different definitions depending on whether the position of the tongue or the position of the roof is given prominence.
12. Retroflex Consonants Articulatory Roof Positions: In general, they represent a group of articulations in which the tip of the tongue is curled up to some degree. In this way retroflex articulations can occur in several different locations on the roof of the mouth including alveolar, post-alveolar, and palatal regions.
13. Retroflex from Underside of Tongue: If the underside of the tongue tip makes contact with the roof of the mouth, it is sub-apical though apical post-alveolar sounds are also described as retroflex (Ladefoged and Maddieson (1996)).
14. Presence of Sub-apical Retroflex Sounds: Typical examples of sub-apical retroflex stops found in Dravidian languages, and in some languages indigenous to the Southwest United States. The contrastive difference between dental and alveolar stops is the slight retroflexion on the alveolar stops (Ladefoged and Maddieson (1996)). Acoustically, retroflexion tends to affect the higher formants.
15. Articulations behind the Alveolar Ridge: Articulations taking place just behind the alveolar ridge, known as post-alveolar consonants, have been referred to using a number of different terms.
16. Post-alveolar and Laminal Articulations: Apical post-alveolar consonants are often called retroflex, while laminal articulations are sometimes called post-alveolar; in the Australianist literature, these laminal stops are often described as *palatal* though they are produced further

forward than the palate region typically described as palatal (Ladefoged and Maddieson (1996)).

17. Imprecise Nature of Alveolar Stops: Because of individual anatomical variation, the precise articulation of palate-alveolar stops – and coronals in general – can vary within a speech community (Ladefoged and Maddieson (1996)).

Dorsal

1. Three Types of Dorsal Consonants: Dorsal consonants are those consonants made by the tongue body rather than the tip or the blade, and are typically produced at the palate, the velum, or the uvula.
2. Palatal Consonant and Dorsal Contrasts: Palatal consonants are made using the tongue body against the hard palate on the roof of the mouth. They are frequently contrasted with velar or uvular consonants, though it is rare for a language to contrast all three simultaneously, with Jaqaru as a possible example of a three-way contrast (Ladefoged and Maddieson (1996)).
3. Velar Consonants and their Occurrence: Velar consonants are made using the tongue body against the velum. They are incredibly common cross-linguistically; almost all languages have a velar stop.
4. Coarticulation among Velars and Vowels: Because both velars and vowels are made using the tongue body, they are highly affected by coarticulation with vowels and can be produced as far forward as the hard palate and as far back as the uvula. These variations are divided into front, central, or back velars in parallel with the vowel space (Ladefoged and Maddieson (1996)). They are hard to distinguish phonetically from palatal consonants, though they are produced slightly behind the area of prototypical palatal consonants (Keating and Lahiri (1993)).
5. Properties of the Uvular Consonants: Uvular consonants are made by the tongue body approaching or contacting the uvula. They are rare, occurring in an estimated 19 percent of world’s languages, and large regions of America and Africa have no language with uvular

consonants. In languages with uvular consonants, stops are most frequent followed by continuants, including nasals (Maddieson (2013)).

Pharyngeal and Laryngeal

1. Definition of Laryngeal/Pharyngeal Consonants: Consonants made by the constriction of the throat are pharyngeal, and those that are made by a constriction in the throat are laryngeal. Laryngeals are made using the vocal folds as the larynx is far too down the throat to reach the tongue. Pharyngeals, however, are close enough to the mouth that parts of the tongue can reach them.
2. Consonants of the Radical Category: Radical consonants either use the root of the tongue or the epiglottis during production and are produced very far back in the vocal tract (Ladefoged and Maddieson (1996)).
3. Pharyngeal Consonants - Production and Properties: Pharyngeal consonants are produced by retracting the root of the tongue far enough to almost touch the wall of the pharynx. Due to production difficulties only fricatives and approximants are produced this way (Ladefoged and Maddieson (1996), Lodge (2009)).
4. Epiglottal Consonants - Production and Types: Epiglottal consonants are made with the epiglottis and the back wall of the pharynx. Epiglottal consonants have been recorded in Dahalo (Ladefoged and Maddieson (1996)). Voiced epiglottal consonants are not deemed possible due to the cavity between the epiglottis and the glottis being too small to permit voicing (Ladefoged and Maddieson (1996)).
5. Glottal Consonants Production and Constraints: Glottal consonants are those produced using the vocal folds in the larynx. Because the vocal folds are the source of phonation and are below the oral-nasal vocal tract, a number of glottal consonants, such as a voice glottal stop, are impossible.

6. Types of Glottal Consonants Possible: Three glottal stops are possible – a voiceless glottal stop and two glottal fricatives, and all are attested in natural languages (International Phonetic Association (2015)).
7. Production and Purpose of Glottal Stops: Glottal stops, produced by closing the vocal folds, are notably common in world's languages (Ladefoged and Maddieson (1996)). While many languages use them to demarcate phrase boundaries, some languages such as Huatla Mazatec have them as contrastive phonemes.
8. Ways of Realizing Glottal Stops: Additionally, glottal stops can be realized as laryngealization of the following vowel in a language (Ladefoged and Maddieson (1996)). Glottal stops, especially between vowels, do not usually form a complete closure. True glottal stops normally occur only when they are geminated (Ladefoged and Maddieson (1996)).

The Larynx

1. Larynx and Vocal Folds/Cords: The larynx, commonly known as the "voice box", is a cartilaginous structure in the trachea responsible for phonation. The vocal folds/cords are held together so that they vibrate, or held apart so that they do not.
2. Vocal folds Position and Tension: The positions of the vocal folds are achieved by the movement of the arytenoid cartilages (Ladefoged and Johnson (2011)). The intrinsic laryngeal muscles are responsible for moving the arytenoid cartilages as well as modulating the tension of the vocal folds (Seikel, Drumright, and King (2016)).
3. Separation of the Vocal Folds: If the vocal folds are not close or tense enough, they will vibrate sporadically or not at all. If they vibrate sporadically it will result in creaky or breathy voice, depending on the degree; if they don't vibrate at all, the result will be voicelessness.
4. Pressure Differential for an Airflow: In addition to correctly positioning the vocal folds, there must be air flowing across them or they will not vibrate. The difference in pressure across the glottis required for voicing is estimated at $1-2\ cm\ H_2O - 98.0665 - 196.033\ Pascals$

(Ohala (1997)). The pressure differential can fall below the levels required for phonation either because of an increase in pressure above the glottis – super-glottal pressure – or a decrease in pressure below the glottis – subglottal pressure.

5. Control of the Subglottal Pressure: The sub-glottal pressure is maintained by the respiratory muscles.
6. Control of the Supra-glottal Pressure: Supra-glottal pressure, with no restrictions or articulations, is equal to the atmospheric pressure. However, because articulations – especially consonants – represent constrictions of the airflow, the pressure in the cavity behind those constrictions can increase resulting in a higher supra-glottal pressure (Chomsky and Halle (1968)).

Lexical Access

1. Two-stage Theory of Lexical Access: According to the lexical access model, two different stages of cognition are employed; thus, this concept is known as the two-stage theory of lexical access.
2. First Stage - Lexical Selection: The first stage, lexical selection, provides information about the lexical items required to construct the functional level representation. These items are retrieved according to their specific semantic and syntactic properties, but phonological forms are not yet made available at this stage.
3. Second Stage - Retrieval of Wordforms: The second stage, retrieval of wordforms, provides information required for building position level representation (Altmann (2002)).

Articulatory Models

1. Coordinate System Basis for Articulation: When producing speech, articulators move through and contact particular locations in space resulting in changes to the acoustic signal. Some models of speech production take this as a basis for modeling articulation in a coordinate system that may be internal to the body, i.e., intrinsic, or external, i.e., extrinsic.
2. Conception behind Intrinsic Coordinate System: Intrinsic coordinate systems model the movements of the articulators as positions and angles of joints in the body. Intrinsic coordinate models of the jaw often use two to three degrees of freedom, representing translation and rotation.
3. Drawbacks of the Intrinsic Coordinate Systems: These face issues modeling the tongue, which, unlike joints of jaw or arms, is a muscular hydrostat, like an elephant trunk, which lacks joins (Lofqvist (2010)). Because of the different physiological structures, movement paths of the jaw are relatively straight lines during speech and mastication, while movement of the tongue follow curves (Munhall, Ostry, and Flanagan (1991)).
4. Rationale behind Extrinsic Coordinate Systems: Straight-line movements have been used to argue articulations as planned in extrinsic rather than intrinsic space, though extrinsic coordinate systems also include acoustic coordinate spaces, not just physical coordinate spaces (Lofqvist (2010)).
5. Extrinsic Coordinate Space – Inverse Problem: Movements that assume that movements are planned in an extrinsic space run into an inverse problem of explaining the muscle and joint locations which produce the observed path or acoustic signal. The arm, for example, has 7 degrees of freedom and 22 muscles, so multiple different joint and muscle configurations can lead to the same final position.
6. Non-unique Muscle Movement Mappings: For models of planning in acoustic space, the same one-to-many problem applies as well, with no unique mappings from the physical or the acoustic targets to the muscle movements required to achieve them. Concerns about the inverse problem may be exaggerated, however, as speech is a highly learned skill using neurological structures that evolved for the purpose (Lofqvist (2010)).
7. The Target Equilibrium-Point Model: The equilibrium-point model proposes a resolution to the inverse problem by arguing that movement targets can be represented as the position of the muscle mass pairs acting on a joint (Feldman (1966)). Importantly, muscles are modeled as springs, and the target is the equilibrium point for the modeled spring-mass system.

8. Advantages of the Equilibrium-Point Model: By using springs, the equilibrium-point model can easily account for compensation and response when movements are disrupted. They are considered a coordinate model because they assume that these muscle positions are represented as points in space – equilibrium points – where the spring-like action of the muscles converges (Bizzi, Hogan, Mussa-Ivaldi, and Giszter (1992), Lofqvist (2010)).
9. The *Minimal Unit* Gestural Model: Gestural approaches to speech production propose that articulations be represented as movement patterns rather than particular coordinates to hit. The minimal unit is a gesture that represent a group of "fundamentally equivalent articulatory movement patterns that are actively controlled with reference to a given speech-relevant goal, e.g., a bilabial closure" (Salzman and Munhall (1989)).
10. Task-Specific Groupings of Muscles: These groups represent coordinated structures or synergies which view movements not as individual muscle movements but as task-dependent groupings of muscle which work together as a single unit (Mattingly (1990), Lofqvist (2010)).
11. Reduction in Degrees of Freedom: This reduces the degrees of freedom in articulation planning, a problem especially in intrinsic coordinate models, which allows for any movement that achieves the speech goal, rather than encoding the particular movements in the abstract representation.
12. Coarticulations under the Gestural Model: Coarticulation is well-described by gestural models, as articulations at faster speech rates can be explained as composites of independent gestures at slower speech rates (Lofqvist (2010)).

Acoustics

1. Importance of Place/Manner of Articulation: Speech sounds are created by the modification of the airstream which results in a sound wave. The modification is done by articulators with different places and manners of articulation producing different acoustic results. Since it is

not just the vocal tract but also the position of the tongue that can affect the resulting sound, manner of articulation is important for describing the speech sound.

2. English Example: *Tack* vs *Sack*: The words *tack* and *sack* both begin with the alveolar sounds in English, but differ in how far the tongue is from the alveolar ridge. This difference has a large effect on the airstream and therefore the sound that is produced. Similarly, the direction and the source of the airstream can affect the sound.

Voicing and Phonation Types

1. Speech Sounds – Voiced vs. Voiceless: A major distinction between speech sounds is whether they are voiced. Sounds are voiced when then vocal folds begin to vibrate in the process of phonation.
2. Speech Sounds with/without Phonation: Many sounds can be produced with or without phonation, though physical constraints make phonation difficult or impossible for some articulations.
3. Sound Source for Voiced Articulations: When articulations are voiced, the main source of noise is the periodic vibration of the vocal folds.
4. Other Non-phonation Acoustic Sources: Articulations by voiceless plosives have no acoustic source and are noticeable by their silence, but other voiceless sounds like fricatives create their own acoustic source regardless of phonation.
5. Acoustics of the Phonation Sources: Phonation is controlled by the muscles of the larynx, and languages make use of more acoustic detail than binary voicing. During phonation, the vocal folds vibrate at a certain rate. This vibration results in a periodic waveform that comprises the fundamental frequency and its harmonics.
6. Control of Fundamental Phonation Frequency: The fundamental frequency of the acoustic wave can be controlled by adjusting the muscles of the larynx, and listeners perceive this fundamental frequency as the pitch.

7. Pitch Manipulation in Language Communication: Tonal languages use pitch manipulation to convey lexical information, and many languages use pitch to mark prosodic or pragmatic information.
8. Determinants of the Vocal Fold Vibration: For vocal folds to vibrate, they must be in proper position and there must be air flowing through the glottis (Ohala (1997)).
9. Glottal States for Phonation Ranges: Phonation types are modeled on a continuum of glottal states from completely open, i.e., voiceless, to completely closed, i.e., glottal stop. The optimal position for vibration, and the phonation type most used in speech, the modal voice, exists in the middle of these two extremes.
10. Causes of Breathy/Creaky Voice: If the glottis is slightly wider, breathy voice occurs, while bringing the folds closer together results in creaky voice (Gordon and Ladefoged (2001)).
11. Characteristics of Typical Speech - Modal Voice: The normal phonation pattern used in typical speech is the modal voice, where the vocal folds are held close together with moderate tension. The vocal folds vibrate as a single unit periodically and efficiently with a full glottal closure and no aspiration (Gobl and Ni Chisaide (2010)).
12. Voiceless Phones and Glottal Stop: If the vocal folds are pulled farther apart, they do not vibrate and so produce voiceless phones. If they are held firmly together, they produce a glottal stop (Gordon and Ladefoged (2001)).
13. Production of Breathy/Whispery Voices: If the vocal folds are held slightly further apart than in modal voicing, they produce phonation types like breathy voice – or murmur – and whispery voice. The tension across the vocal ligaments – the vocal chords – is less than modal voicing allowing for air to flow more freely.
14. Characteristics of Breathy/Whispery Voices: Both breathy voice and whispery voice exist on a continuum loosely characterized as going from the more periodic waveform of the breathy voice to the more noisy waveform of the whispery voice. Acoustically, both tend to dampen the first formant with the whispery voice showing more extreme deviations (Gobl and Ni Chisaide (2010)).
15. Production of Creaky Voice: Holding the vocal folds more tightly together results in a creaky voice. The tension across the vocal folds is less than in modal voice, but they are held together tightly resulting in only the ligaments of vocal folds vibrating. The pulses are highly irregular, with low pitch and frequency amplitude (Gobl and Ni Chisaide (2010)).

16. Voicing/Voicing Distinction across World's Languages: Some languages do not maintain a voicing distinction for some consonants – Hawaiian, for example, does not contrast voiced and voiceless plosives – but all languages use voicing to some degree.
17. Phonemic Voicing Contrast for Vowels: For example, no language is known to have phonemic contrast for vowels – exceptions are languages like Japanese, where vowels are produced as voiceless in certain contexts.
18. Phonemic Contrast across Glottal Positions: Other positions of the glottis, such as breathy and creaky voice, are used in a number of languages, like Jalapa Mazatec, to contrast phonemes while in other languages, such as English, they exist allophonically.
19. Identification of Voicing in Segments: There are several ways to determine if a segment is voiced or not, the simplest being to feel the larynx during the speech and note when vibrations are felt. More precise measurements can be obtained through acoustic analysis of a spectrogram or a spectral slice.
20. Spectrographic Analysis of Voiced Segments: In spectrographic analysis, voiced segments show a voicing bar, a region of high acoustic energy, in the low frequencies of voice segments (Dawson and Phelan (2016)).
21. Uncovering Signals of Spectral Slice: In examining a spectral splice, i.e., the acoustic spectrum at a given point in time, the model of the vowel pronounced reverses the filtering of the mouth producing the spectrum of the glottis. A computational model of the unfiltered glottal signal is then fitted to the inverse filtered acoustic signal to determine the characteristics of the glottis (Gobl and Ni Chisaide (2010)).
22. Visual Analysis using Special Instruments: Visual analysis is also available using specialized medical equipment such as ultrasound and endoscopy (Dawson and Phelan (2016)).

Vowels

1. Vowel Production – Unrestricted Vocal Tract: Vowels are broadly characterized by the area of the mouth in which they are produced, but because they are produced without a

constriction in the vocal tract, their precise description relies in measuring acoustic correlates of the tongue position.

2. Cavity Resonance Impacted by Tongue Position: The location of the tongue during vowel production changes the frequency at which the cavity resonates, and it is these resonances – known as formants – which are measured and used to characterize vowels.
3. Definition of the Vowel Height: Vowel height traditionally refers to the highest point of the tongue during articulation (Ladefoged and Maddieson (1996)).
4. Classification of the Height Parameter: The height parameter is divided into four primary levels – high/close, close-mid, open-mid, and low/open. Vowels whose height are in the middle are referred to as mid.
5. Opened-Close/Closed-Open Vowels: Slightly opened-close vowels and slightly closed-open vowels are referred to as near-close and near-open vowels, respectively. The lowest vowels are not articulated with a lowered tongue, but also by lowering the jaw (Lodge (2009)).
6. IPA Vowels:

	Front	**Central**	**Back**
Close			
Near-close			
Close-mid			
Mid			
Open-mid			
Near-open			
Open			

7. Superfluousness of the 7 IPA Vowels: While the IPA implies that there are 7 levels of vowel height, it is unlikely that a given language can minimally contrast all 7 levels. Chomsky and Halle (1968) suggest that there are only three levels, although four levels of vowel height

seem to be needed to describe Danish and it is possible that some languages may even need five (Ladefoged and Maddieson (1996)).

8. Classification of the Backness Parameter: Vowel backness is divided into three levels: front, central, and back. Languages usually do not contrast more than two levels of vowel backness. Languages claimed to have a three-way backness distinction include Nimboran and Norwegian (Ladefoged and Maddieson (1996)).
9. Characteristics of the Lip Position: In most languages, the lip position during vowel production can be classified as either rounded or unrounded/spread, although other types of lip positions, such as compression and protrusion, have been described.
10. Lip Position/Height/Backness Correlation: Lip position is correlated with height and backness: front and low vowels tend to be rounded whereas back and high vowels are usually unrounded (Ladefoged and Maddieson (1996)). Paired vowels on the IPA chart have the spread vowel on the left and the rounded vowel on the right (Lodge (2009)).
11. Additional Generic Vowel Characterization Features: Together with the universal vowel features described above, some languages have additional features such as nasality, length, and different types of phonation such as voiceless and creaky.
12. Specialized Tongue Gesture Descriptor Parameters: Sometimes more specialized tongue gestures, such as rhoticity, advanced tongue root, pharyngealization, stridency, and frication are required to describe a certain vowel (Ladefoged and Maddieson (1996)).

Manner of Articulation

1. Articulator Modification of Vocal Tract: Knowing the place of articulation is not enough to fully describe a consonant, the way in which the stricture happens is equally important. Manner of articulation describes exactly how the articulator modifies, narrows, or closes off the vocal tract (Ladefoged and Maddieson (1996)).
2. Manner of Articulation for Plosives: Stops – also referred to as plosives – are consonants where the airstream is completely obstructed. Pressure builds up in the mouth during the

stricture, which is then released as a small burst of sound when the articulators move apart. The velum is raised so that the air cannot flow through the nasal cavity.

3. Production of a Nasal Stop: If the velum is lowered and allows for air to flow through the nose, the result is a nasal stop. However, phoneticians always refer to nasal stops as just *nasals* (Ladefoged and Maddieson (1996), Ladefoged and Johnson (2011)).
4. Manner of Articulation for Fricatives: Fricatives are consonants where the airstream is made turbulent by partially, but not completely, obstructing part of the vocal tract. Sibilants are special type of fricatives where the turbulent airstream is directed towards the teeth (Ladefoged and Maddieson (1996)), creating a high-pitched hissing sound (Ladefoged and Johnson (2011)).
5. Manner of Articulation for Affricates: Affricates are a sequence of steps followed by a fricative in the same place (Ladefoged and Johnson (2011)).
6. Manner of Articulation for Approximants: In an approximant, the articulators come close together, but not to such an extent that allows a turbulent airstream (Ladefoged and Johnson (2011)).
7. Manner of Articulation for Laterals: Laterals are consonants in which the airstream is obstructed along the center of the vocal tract, allowing the airstream to flow freely on one or both sides (Ladefoged and Johnson (2011)). Laterals are also defined as consonants in which the tongue is contracted in such a way that the airstream is greater around the sides than over the center of the tongue (Ladefoged and Maddieson (1996)).
8. Manner of Articulation for Trills: Trills are consonants in which the tongue or the lips are set in motion by the airstream (Ladefoged and Johnson (2011)). The stricture is formed in such a way that the airstream causes a repeating pattern of opening and closing of the soft articulator(s) (Ladefoged and Maddieson (1996)). Apical trills typically consist of two or three periods of vibration.
9. Taps/Flaps Manner of Articulation: Taps and flaps are single, rapid, usually apical gestures where the tongue is thrown against the roof of the mouth, comparable to a very rapid stop (Ladefoged and Johnson (2011)). These terms are sometimes used interchangeably, but some phoneticians make a distinction (Ladefoged and Maddieson (1996)). In a tap, the tongue contacts the roof in a single motion, whereas in a flap the tongue moves tangentially to the roof of the mouth, striking it in passing.

10. Mechanism behind Glottalic Airstream Articulation: During a glottalic airstream mechanism, the glottis is closed, trapping a body of air. This allows for the remaining air in the vocl tract to be moved separately.
11. Ejective/Implosive Manner of Articulation: An upward movement of the glottis will move the air out, resulting in an ejective consonant. Alternatively, the glottis can lower, sucking air into the mouth, which results in an implosive consonant (Ladefoged and Johnson (2011)).
12. Clicks/Velaric Airstream Articulation Mechanism: Clicks are stops in which the tongue movement causes the airstream to be sucked in the mouth, this is referred to as a velaric airstream (Ladefoged and Maddieson (1996)). During the click the air becomes rarefied between two articulatory closures, producing a loud *click* sound when the anterior is released.
13. Influx and Efflux Click Consonants: The release of the anterior closure is referred to as the click influx. The release of the posterior closure, which can be velar or uvular, is the click efflux.
14. Click Usage in African Languages: Clicks are used in several African language families, such as Khoisan and Bantu languages (Ladefoged and Maddieson (1996)).

Pulmonary and Subglottal System

1. Lung Pressure and Pulmonic Egress: The lungs drive nearly all speech production, and their importance in phonetics is due to their creation od pressure in pulmonic sounds. The most common kind of sounds is the pulmonic egress, where air is exhaled from the lings (Ladefoged and Johnson (2011)).
2. Pulmonic Ingress and its Occurrence: The opposite is possible, although no language is known to have pulmonic ingressive sounds as phonemes. Many languages such as Swedish use them for paralinguistic articulations such as affirmations, though this is the case in a number of geographically diverse languages (Eklund (2008)).

3. Pulmonic Air Draw vs Vital Capacity: Both ingressive and egressive sounds rely on holding the vocal folds at a particular posture and using the lungs to draw the air across the vocal folds so that they either vibrate – voiced – or do not vibrate – voiceless (Ladefoged and Johnson (2011)). Pulmonic articulations are restricted by the volume of air exhaled in a given respiratory cycle, known as the vital capacity.
4. Maintenance of Super/Sub Glottal Pressure: The lungs are used to maintain two kinds of pressure in order to produce and modify phonation. To produce phonation at all, the lings must maintain a pressure of $3 - 5\ cm\ H_2O$ higher than the pressure above the glottis.
5. Pressure Differential for Suprasegmentals: However, small and fast adjustment are made to the subglottal pressure to modify speech for suprasegmental features like stress. A number of thoracic muscles are used to make these adjustments.
6. Vital Capacity Pressure Differential Retention: Because the lungs and the thorax stretch during inhalation, the elastic forces of the lungs alone can produce differentials for lung volumes above 50 percent of the vital capacity (Seikel, Drumright, and King (2016)).
7. Vital Capacity Pressure Maintenance Mechanism: Above 50 percent of the vital capacity, the respiratory muscles are used to "check" the elastic forces of the thorax to maintain a stable pressure differential. Below that volume, they are used to increase the sub-glottal pressure by actively exhaling air.
8. Accommodating Linguistic and Metabolic Needs: During speech, the respiratory is modified to accommodate both the linguistic and the biological needs. Exhalation, usually about 60 percent of the respiratory cycle at rest, is increased to about 90 percent of the respiratory cycle. Because metabolic needs are relatively stable, the total air moved in most cases of speech remain about the same as quiet tidal breathing (Seikel, Drumright, and King (2016)).
9. Age/Loudness Vital Capacity Impact: Increase in speech intensity of 18 dB – a loud conversation – has relatively little impact on the volume of air moved. Because their respiratory systems are not as developed as adults, children tend to use a larger proportion of their vital capacity compared to adults, with more deeper inhales (Seikel, Drumright, and King (2016)).

Source-Filter Theory

1. Source-Filter Model of Speech: The source-filter model of speech is a theory of speech production which explains the link between vocal tract posture and acoustic consequences (Johnson (2011)). The noise source in many cases is the larynx during the process of voicing, though other noise sources can be modeled the same way.
2. Factors Impacting the Generated Acoustic Patterns: The shape of the supraglottal vocal tract as the filter, and different configurations of the articulators result in different acoustic patterns. The changes are predictable.
3. Acoustics of Vocal Tract Modeling: The vocal tract can be modeled as a sequence of tunes, closed at one end, with varying diameters, and by using equations for acoustic resonance, the acoustic effect of an articulatory posture can be derived (Johnson (2011)).
4. Acoustics Produced by the Vocal Folds: The process of inverse filtering uses this principle to analyze the source spectrum produced by the vocal folds during the voicing. By taking the inverse of the predicted filter, the acoustic effects of the supraglottal vocal tract can be undone giving the acoustic spectrum produced by the vocal folds (Johnson (2011)). This allows for the quantitative study of various phonation types.

Perception

1. Language Perception – Decoding Linguistic Signal: Language perception is the process by which the linguistic signal is decoded and understood by the listener.
2. Decomposition into Phonemes/Morphemes/Words: In order to perceive speech, the continuous acoustic signal must be converted into discrete linguistic units such as phonemes, morphemes, and words (Sedivy (2019)).
3. Prioritization/Enhancement of Acoustic Cues: In order to correctly identify and categorize sounds, listeners prioritize certain aspects of the signal that can reliably distinguish between

linguistic categories. While certain cues can be prioritized over others, many aspects of the signal can contribute to the perception. For example, though oral languages prioritize acoustic information, the McGurk effect shows that the visual information is used to distinguish ambiguous information where acoustic cues are unreliable (Sedivy (2019)).

4. Acoustic Signal/Category Perception Mapping: While listeners can use a variety of information to segment speech signal, the relationship between acoustic signal and category perception is not a perfect mapping. Because of coarticulation, noisy environment, and individual differences, there is a high degree of acoustic variability within categories (Sedivy (2019)).
5. Perceptual Invariance - Definition and Motivation: Known as the problem of *perceptual invariance*, listeners are able to reliably perceive categories despite the variability in acoustic information. In order to accommodate this, listeners rapidly accommodate new speakers and will shift their boundaries between the categories to match the acoustic distinctions their conversational partner is making (Sedivy (2019)).

Perception – Audition

1. Air Pressure to Sound Transform: Audition, the process of hearing sounds, is the first stage of perceiving speech. Articulators cause systematic change in air pressure which travel as sound waves to the listener's ears.
2. Ear Drum to Cochlear Bones: The sound waves then hit the listener's eardrum causing it to vibrate. The vibration of the eardrum is transmitted by the ossicles – three small bones of the middle ear – to the cochlea (Johnson (2003)).
3. Tonotopic Design of Basilar Membrane: The cochlea is a spiral-shaped, fluid-filled tube divided lengthwise by the organ of Corti which contains the basilar membrane. The basilar membrane increases in thickness as it travels through the cochlea causing different

frequencies to resonate at different locations. This tonotopic design allows for the ear to analyze sound in a manner similar to a Fourier transform (Johnson (2003)).

4. Acoustic to Neuronal Signal Conversion: The differential vibration of the basilar causes the hair cells within the organ of Corti to move. This cases depolarization of the hair cell and ultimately a conversion of the acoustic signal into a neuronal signal (Schachter, Gilbert, and Wegner (2011)).
5. Production of the Action Potentials: While the hair cells do not produce action potentials themselves, they release neurotransmitter at synapses with the fibers of the auditory nerve, which does produce action potentials. In this way, patterns of oscillations on the basilar membrane are converted to spatiotemporal firings which transmit information about the sound to the brainstem (Yost (2003)).

Prosody

1. Cross-Speech Auditory Properties/Degrees: Besides consonants and vowels, phonetics also describes properties of speech that are localized to segments but to greater units of speech such as syllables and phrases. Prosody includes auditory characteristics such as pitch, speech rate, duration, and loudness.
2. Prosody Property Correlates across Languages: Languages use these properties in different degrees to implement stress, pitch accents, and intonation – for example, stress in English and Spanish is correlated with changes in pitch and duration, whereas stress in Welsh is more consistently corelated with pitch than duration, and stress is Thai is only correlated with duration (Cutler (2005)).

Theories of Speech Perception

1. Motivation behind the Motor Theory: Early theories of speech perception such as the motor theory attempted to solve the problem of perceptual invariance by arguing that speech production and perception are closely linked.
2. Strong Form of Motor Theory: In its strongest form, motor theory argues that speech perception requires the listener to access the articulatory representation of sound (Sedivy (2019)); in order to properly categorize a sound, a listener reverse engineers the articulation which would produce the sound and by identifying the gestures is able to retrieve the linguistic category (Galantucci, Fowler, and Turvey (2006)).
3. Weak Form of Motor Theory: While findings such as the McGurk effect and case studies from patients with neurological injuries have provided support for the motor theory, further experiments have not supported the string form of the motor theory, though there is some support for weaker forms of motor theory which claim a non-deterministic relationship between production and perception (Galantucci, Fowler, and Turvey (2006), Skipper, Devlin, and Lametti (2017), Sedivy (2019)).
4. Successor Theories of Speech Perception: Successor theories of speech perception place the focus on acoustic cues to sound categories and can be grouped into two broad categories: abstractionist theories and episodic theories (Goldinger (1996)).
5. Idea behind the Abstractionist Theories: In abstractionist theories, speech perception involves the identification of an idealized lexical object based on a signal reduced to its necessary components and normalizing the signal to counteract speaker variability.
6. Motivation behind the Episodic Thesis: Episodic theories such as the exemplar model argue that speech perception involves accessing detailed memories, i.e., episodic memories, of previously heard tokens. The problem of perceptual invariance is explained by the episodic theories as an issue of familiarity; normalization is a by-product of exposure to more variable distributions rather than a discrete process as abstractionist theories claim (Goldinger (1996)).

Sub-disciplines – Acoustic Phonetics

Acoustic phonetics deals with the acoustic properties of speech sounds. The sensation of sound is caused by pressure fluctuations which cause the eardrum to move. The ear transforms this movement into neuronal signals that the brain registers as sound. Acoustic waveforms are records that measure these pressure fluctuations (Johnson (2011)).

Sub-discipline – Articulatory Phonetics

Articulatory phonetics deals with the ways in which speech sounds are made.

Sub-disciplines – Auditory Phonetics

1. Transforming Acoustics for Human Perception: Auditory phonetics studies show how human perceive speech sounds. Due to the fact that the anatomy of the auditory system distorts the speech signals, humans do not experience speech sounds as perfect acoustic records. For example, the auditory impressions of volume, measured in decibels dB, do not linearly match the difference in sound pressure (Johnson (2011)).
2. Acoustics vs. Listener Mismatch Characterization: The mismatch between acoustic analysis and what the listener hears is especially noticeable in speech sounds that have a lot of high-

frequency energy, such as certain fricatives. To reconcile this mismatch, functional models of auditory system have been developed (Johnson (2011)).

Describing Sounds

1. Ways of Specifying Speech Phones: Human languages use many different sounds and in order to compare them linguists must be able to describe sounds in a way that is language independent. Speech sounds can be described in a number of different ways.
2. Gross Categorization - Consonants and Vowels: Most commonly, speech sounds are referred to by the mouth movements needed to produce them. Consonants and vowels are two categories that phoneticians define by the movements in a speech sound. More fine-grained descriptors are parameters such as place of articulation.
3. Fine-grained Consonant/Vowel Chart: Place of articulation, manner of articulation, and voicing are used to describe consonants and are the main divisions of the International Phonetic Alphabet consonant chart. Vowels are described by their height, backness, and rounding.
4. Sign Language Specification Parameters: Sign language I described by a similar but a distinct set of parameters to describe signs: location, movement, hand-shape, palm-orientation, and non-manual features.
5. Describing Sounds in Oral Languages: In addition to articulatory descriptions, sounds in oral languages can be described using their acoustics. Because the acoustics are a consequence of the articulation, both methods of description are sufficient to distinguish sounds with the choice between the systems dependent on the phonetics being investigated.
6. Consonants - Obstruction of the Vocal Tract: Consonants are speech sounds that are articulated with a complete or a partial closure of the vocal tract. They are generally produced by a modification of the airstream exhaled from the lungs.

7. Respiratory Organs Manipulating the Airflow: The respiratory organs used to create and modify the airflow are divided into three regions: the vocal tract – supra-laryngeal, the larynx, and the subglottal system. The airstream can either be egressive – out of the vocal tract, or ingressive – into the vocal tract.
8. Pulmonic, Glottalic, and Click Consonants: In pulmonic, the airstream is produced by the lungs in the sub-glottal system and passes through the larynx and the vocal tract. Glottalic sounds use an airstream created by the movements of the larynx without airflow from the lungs. Click consonants are articulated through the rarefaction of the air using the tongue, followed by the release of the forward closure of the tongue.
9. Vowels - No Vocal Tract Obstruction: Vowels are syllabic speech sounds which are pronounced without any obstruction in the vocal tract (Ladefoged and Maddieson (1996)). Unlike consonants, which usually have definite places of articulation, vowels are defined in relation to a set of reference vowels called cardinal vowels. Three properties are needed to define vowels: tongue height, tongue backness, and lip roundedness.
10. Distinction between Monophthongs and Diphthongs: Vowels that are articulated with a stable quality are called monophthongs; a combination of two separate vowels in the same syllable is a diphthong (Gussenhoven and Jacobs (2017)).
11. IPA Chart Vowel Representation Scheme: In the IPA, the vowels are represented on a trapezoid shape representing the human mouth; the vertical axis represents the mouth from floor to roof and the horizontal axis represents the front-back dimension (Lodge (2009)).

Transcription

1. IPA Symbols for Oral Phones: Phonetic transcription is a system for transcribing phones that occur in a language, whether oral or sign. The most widely known system of phonetic transcription – the International Phonetic Alphabet IPA – provides a standardized set of symbols for oral phones (International Phonetic Association (1999), O’Grady (2005)).

2. Purpose and Usage of IPA: The standardized nature of the IPA enables its users to transcribe accurately and consistently the phones of several languages, dialects, and idiolects (Ladefoged and Maddieson (1996), International Phonetic Association (1999), O'Grady (2005)). The IPA is a useful tool not only for the study of phonetics, but also for language teaching, professional acting, and speech pathology (Ladefoged and Johnson (2011)).
3. Standardized Symbols for Sign Languages: While no sign language has a standardized writing system, linguists have developed their own notation systems that describes handshape, location, and movement. The Hamburg Notation System – HamNoSys – is similar to IPA in that it allows varying levels of detail.
4. Comparing HamNoSys, KOMVA, and Stokoe: Some notations such as KOMVA and the Stokoe system were designed for use in dictionaries; they also make use of alphabetic letters in the local language for handshapes whereas HamNoSys represents handshapes directly.
5. SignWriting Easy-to-Learn Language: SignWriting aims to be an easy-to-learn writing system for sign languages, although it has not been officially adopted by any deaf community yet (Baker, van den Bogaerde, Pfau, and Schermer (2016)).

Sign Languages

1. Visual Perception of Sign Languages: Unlike spoken languages, words in sign languages are perceived with the eyes instead of the ears. Signs are articulated with hands, upper body, and head. The main articulators are the hands and the arms.
2. Proximal and Distal Sign Movements: Relative parts of the arm are described with the terms proximal and distal. Proximal refers to a part closer to the torso whereas a distal part is further away from it. For example, a wrist movement is distal compared to an elbow movement. Due to requiring less energy, distal movements are easier to produce.

3. Criteria Restricting Formation of Sign: Various factors – such as muscle flexibility or being considered taboo – restricted what can be considered a sign (Baker, van den Bogaerde, Pfau, and Schermer (2016)).
4. Signs Articulated Close to Face: Native signers do not look at their conversation partner's hands. Instead, their gaze is fixated on the face. Because peripheral vision is not as focused as the center of the visual field, signs articulated near the face allow for more subtle differences in finger movements and articulation to be perceived (Baker, van den Bogaerde, Pfau, and Schermer (2016)).
5. Signs Produced with Two Hands: Unlike spoken languages, sign languages have two identical articulators – the hands. Signers may use whichever hand they prefer with no disruption in communication.
6. First Universal Constraint - Symmetry Condition: Due to universal neurological limitations, two-handed signs generally have the same kind of articulation in both hands; this is referred to as the Symmetry Condition (Baker, van den Bogaerde, Pfau, and Schermer (2016)).
7. Second Universal Constraint - Dominance Condition: The second universal constraint I the Dominance Condition, which holds that when two handshapes are involved, one hand will remain stationary and have a more limited set of handshapes compared to the dominant, moving hand (Baker, van den Bogaerde, Pfau, and Schermer (2016)).
8. Sign Influenced Coarticulation - Assimilation/Deletion: Additionally, it is common for one hand in a two-handed sign to be dropped during informal conversations, a process referred to as weak drop (Baker, van den Bogaerde, Pfau, and Schermer (2016)). Just like words in spoken languages, coarticulation may cause signs to influence each other's form. Examples include handshape of neighboring signs becoming similar to each other, i.e., assimilation, or weak drop, i.e., an instance of deletion.

References

- Altmann, G. (2002): *Psycholinguistics: Critical Concepts in Psychology* **Routledge** London, UK
- Baker, A., B. van den Bogaerde, R. Pfau, and T. Schermer (2016): *The Linguistics of Sign Languages* **John Benjamins Publishing Company** Amsterdam, Netherlands
- Bizzi, E., N. Hogan, F. Mussa-Ivaldi, and S. Giszter (1992): Does the Nervous System use Equilibrium-Point Control to Guide Single and Multiple Joint Movements? *Behavioral and Brain Sciences* **15 (4)** 603-613
- Boersma, P. (1998): *Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives* **Holland Academy Graphics** The Hague
- Chomsky, N., and M. Halle (1968): *Sound Pattern of English* **Harper and Row**
- Cutler, A. (2005): Lexical Stress, in: *The Handbook of Speech Perception (Editors: Pisoni, D. B., and R. Remez)* **Blackwell**
- Dawson, H., and M. Phelan (2016): *Language Files: Materials for an Introduction to Linguistics 12th Edition* **The Ohio State University Press** Columbus, OH
- Dell, G. and P. O'Seaghdha (1992): Stage of the Lexical Access in Language Production *Cognition* **4 (1-3)** 287-314
- Doke, C. M. (1926): *The Phonetics of the Zulu Language* **Witwatersrand University Press** Johannesburg, South Africa
- Eklund, R. (2008): Pulmonic Ingressive Phonation: Diachronic and Synchronic Characteristics, Distribution, and Function in Animal and Human Sound Production and in Human Speech *Journal of the International Phonetic Association* **35 (3)** 235-324
- Feldman, A. G. (1966): Functional Tuning of the Nervous System with Control of Movement or Maintenance of Steady Posture, III: Mechanographic Analysis of the Execution by Man of the Simplest Motor Task *Biophysics* **11** 565-578
- Fujimura, O. (1961): Bilabial Stop and Nasal Consonants: A Motion Picture Study and its Acoustical Implications *Journal of Speech and Hearing Research* **4 (3)** 233-247
- Galantucci, C. Fowler, and M. Turvey (2006): The Motor Theory of Speech Perception Reviewed *Psychonomic Bulletin and Review* **13 (3)** 361-377.
- Gleitman, L., D. January, R. Nappa, and J. Trueswell (2007): On the Give and Take between Event Apprehension and Utterance Formulation *Journal of Memory and Language* **57 (4)** 544-569

- Gobl, C., and A. Ni Chisaide (2010): Voice source Variation and its Communicative Functions, in: *The Handbook of Phonetic Sciences 2nd Edition (Editors: Hardcastle, W. J., J. Laver, and F. E. Gibbon)* **Wiley**
- Goldinger, S. (1996): Words and Voices: Episodic Trace Spoken Word Identification and Recognition Memory *Journal of Experimental Psychology: Learning, Memory, and Cognition* **22 (5)** 1166-1183
- Gordon, M., and P. Ladefoged (2001): Phonation Types: A Cross-linguistic Overview *Journal of Phonetics* **29 (4)** 383-406
- Gussenhoven, C., and H. Jacobs (2017): *Understanding Phonology 4th Edition* **Rutledge** New York, NY
- Guthrie, M. (1948): *The Classification of Bantu Languages* **Oxford University Press** London, UK
- International Phonetic Association (1999): *Handbook of the International Phonetic Association* **Cambridge University Press** Cambridge, UK
- International Phonetic Association (2015): *International Phonetic Alphabet* **International Phonetic Association** London, UK
- Jaeger, F., K. Furth, and C. Hilliard (2012): Phonological Overlap affects Lexical Selection During Sentence Production *Journal of Experimental Psychology: Learning, Memory, and Cognition* **38 (5)** 1439-1449
- Johnson, K. (2011): *Acoustic and Auditory Phonetics 3rd Edition* **Wiley-Blackwell**
- Keating, P., and A. Lahiri (1993): Fronted Velars, Palatalized Velars, and Palatals *Phonetic* **50 (2)** 73-101
- Ladefoged, P., and I. Maddieson (1996): *The Sounds of World's Languages* **Blackwell** Oxford, UK
- Ladefoged, P., and K. Johnson (2011): *A Course in Phonetics 6th Edition* **Wadsworth**
- Lodge, K. (2009): *A Critical Introduction to Phonetics* **Continuum International Publishing Group** New York, NY
- Lofqvist, A. (2010): Theories and Models of Speech Production, in: *The Handbook of Phonetic Sciences 2nd Edition (Editors: Hardcastle, W. J., J. Laver, and F. E. Gibbon)* **Wiley**

- Maddieson, I. (1993): Investigating Ewe Articulations with Electromagnetic Articulography *Forschungberichte des Instituts fur Phonetik und Sprachtliche Kommunikation der Universitat Munchen* **31** 181-214
- Maddieson (2013): Uvular Consonants, in: *The World Atlas of Language Structures Online (editors: Dryer, M. S., and M. Haspelmath)* **Max Planck Institute for Evolutionary Anthropology** Leipzig, Germany
- Mattingly, I. (1990): The Global Character of Phonetic Gestures *Journal of Phonetics* **18 (3)** 445-452
- Motley, M., C. Camden, and B. Baars (1982): Covert Formulation and Editing of Anomalies of Speech Production: Evidence from Experimentally Elicited Slips of the Tongue *Journal of Verbal Learning and Verbal Behavior* **21 (5)** 578-594
- Munhall, K., D. Ostry, and J. Flanagan (1991): Coordinate Spaces in Speech Planning *Journal of Phonetics* **19 (3-4)** 293-307
- O'Grady, W. (2005): *Contemporary Linguistics: An Introduction 5th Edition* **Bedford/St. Martin's**
- Ohala, J. (1997): Aerodynamics of Phonology *Proceedings of the Seoul International Conference on Linguistics* **92**
- Saltzman, E., and K. Munhall (1989): Dynamical Approach to Gestural Patterning in Speech Production *Ecological Psychology* **1 (4)** 333-382
- Scatton, E. (1984): *A Reference Grammar of Modern Bulgarian* **Slavica**
- Schachter, D., D. Gilbert, and D. Wegner (2011): Sensation and Perception, in: *Psychology (Editors: Linsmeiser, C.)* **Worth Publishers**
- Sedivy, J. (2019): *Languages in Mind: An Introduction to Psycholinguistics 2nd Edition* **Oxford University Press**
- Seikel, J. A., D. Drumright, and D. King (2016): *Anatomy and Physiology for Speech, Language, and Hearing 5th Edition* **Cengage**
- Skipper, J., J. Devlin, and D. Lametti (2017): The Hearing Ear is always Close to the Speaking Tongue: Review of the Role of the Motor System in Speech Perception *Brain and Language* **164** 77-105
- Yost, W. (2003): Audition, in: *Handbook of Psychology: Experimental Psychology (Editors: Healy, A. F., and R. W. Proctor)* **John Wiley and Sons**

Acoustic Phonetics

1. Coverage Scope of Acoustic Phonetics: *Acoustic Phonetics* is a subfield of phonetics, which deals with the acoustic aspects of speech sounds. Acoustic phonetics investigates the time domain features such as the mean squared amplitude of a waveform, its duration, its fundamental frequency, or frequency domain features such as the frequency spectrum, or even combined spatiotemporal features and the relationship of these properties to other branches of phonetics, e.g., articulatory or auditory phonetics, and to linguistic concepts such as phonemes, phrases, or utterances (Wikipedia (2021)).
2. ILPR - Approximation of the Voice Signal: Integrated Linear Prediction Residuals – ILPR – was an effective feature proposed which closely approximates the voice source signal (Ananthapadmanabha (1995)). This proved to be very effective in the accurate estimation of the epochs or the glottal closure instant (Prathosh, Ananthapadmanabha, and Ramakrishnan (2013)).
3. Speaker Information Supplementing Voice Signal: Ramakrishnan, Abhiram, and Mahadeva Prasanna (2015) showed that the discrete cosine transform coefficients of the ILPR contains speaker information that supplements the mel frequency cepstral coefficients.
4. Characterizing Stop Consonants - Plosion Index: Plosion index is another scalar, time-domain feature that was introduced by Ananthapadmanabha, Prathosh, and Ramakrishnan (2014) for characterizing closure-burst transition of stop consonants.

References

- Ananthapadmanabha, T. V. (1995): Acoustic Factors determining Perceived Voice Quality, in: *Vocal Fold Physiology – Voice Quality Control (Editors: Fujimura, O., and M. Hirano)* **Singular Publishing Group** San Diego, CA
- Ananthapadmanabha, T. V., A. P. Prathosh, and A. G. Ramakrishnan (2014): Detection of the Closure-burst Transition of Stops and Affricates in Continuous Speech using the Plosion Index *Journal of the Acoustical Society of America* **135 (1)** 460-471
- Prathosh, A. P., T. V. Ananthapadmanabha, and A. G. Ramakrishnan (2013): Epoch Extraction Based on Integrated Linear Prediction Residual using Plosion Index *IEEE Transactions on Audio, Speech, and Language Processing* **21 (12)** 2471-2480
- Ramakrishnan, A. G., B. Abhiram, and S. R. Mahadeva Prasanna (2015): Voice Source Characterization using Pitch Synchronous Discrete Cosine Transform for Speaker Identification *Journal of the Acoustical Society of America* **137 (6)** EL469-475
- Wikipedia (2021): Acoustic Phonetics

Articulatory Phonetics

Overview

1. Focus of Articulatory Phonetics Subfield: The field of *articulatory phonetics* is a subfield of phonetics that studies *articulation* and the way humans produce speech. Articulatory phoneticians explain how humans produce speech sounds via the interaction of different physiological studies. Generally, articulatory phonetics is concerned with the transformation of aerodynamic energy into acoustic energy (Wikipedia (2021)).
2. Aerodynamic Energy off Vocal Tract: Aerodynamic energy refers to the airflow through the vocal tract. Its potential form is air pressure; its kinetic form is the actual dynamic airflow.
3. Pressure Differential Causing Acoustic Energy: Acoustic energy is variation in air pressure that can be represented as sound waves, which are then perceived by the human auditory system as sound. Note that although sound is air pressure variations, the variations must be at a high enough rate to be perceived as sound. If the variation is too low, it will be inaudible.
4. Shaping the Flow of Airstream: Sound is produced by expelling air from the lungs. However, to vary the sound quality in a way useful for speaking, the two speech organs normally move towards each other to contact each other to create an obstruction that shapes the air in a particular fashion.
5. Place vs Manner of Articulation: The point of maximum obstruction is called the *place of articulation*, and the way the obstruction forms and releases is the *manner of articulation*.
6. Example - Pronouncing the Bilabial Plosive: For example, when making the *p* sound, the lips come together tightly, blocking the air momentarily and causing a build-up of the air pressure. The lips then release suddenly causing a burst of sound. The place of articulation of this sound is therefore called *bilabial*, and the manner is called *stop*, also known as a *plosive*.

Components

1. Vocal Tract Components Generating the Sound: The vocal tract can be viewed as an aerodynamic-bio mechanic model that includes three main components:
    a. Air cavities
    b. Pistons
    c. Air Valves
2. Components of the Air Cavity: Air cavities are containers of air molecules of specific volumes and masses. The main air cavities present in the articulatory system are the supraglottal and the subglottal cavities. They are so named because the glottis, the openable space between the vocal folds internal to the larynx, separates the two cavities.
3. Entities in the Supraglottal Cavity: The supraglottal cavity or the orinasal cavity is divided into an oral sub-cavity – the cavity from the glottis to the lips excluding the nasal cavity – and a nasal sub-cavity, the cavity from the velopharyngeal port, which can be closed by raising the velum.
4. Entities in the Subglottal Cavity: The subglottal cavity consists of the trachea and the lungs.
5. The External, Atmospheric Air Cavity: The atmosphere external to the articulatory stem may also be considered an air cavity whose potential connecting points with respect to the body are the nostrils and the lips.
6. Pistons - The Air Pressure Initiators: Pistons are initiators. The term *initiator* refers to the fact that they are used to initiate a change in the volumes of the air cavities, and, by Boyle's Law, the corresponding air pressure of the cavity.
7. Initiation of the Airstream Mechanism: The term *initiation* refers to the change. Since changes in air pressure between connected cavities leads to airflow between the cavities, the initiation is also referred to as an *airstream mechanism*.
8. Pistons of the Articulatory System: The three structures present in the articulatory system are the larynx, the tongue body, and the physiological structures used to manipulate the lung volume – in particular, the floor and the walls of the chest. The lung pistons are used to initiate a pulmonic airstream – this is found in all languages.

9. <u>Initiation of the Glottalic Airstream Mechanism</u>: The larynx is used to initiate the glottalic airstream mechanism by changing the volume of the supraglottal and the subglottal cavities via the vertical movement of the larynx, with a closed glottis. Ejectives and implosives are made via this airstream mechanism.
10. <u>Initiation of the Velaric Airstream Mechanism</u>: The tongue body creates a velaric airstream by changing the pressure within the oral cavity: the tongue body changes the mouth sub-cavity. Click consonants use the velaric airstream mechanism.
11. <u>Control of the Piston Action</u>: Pistons are controlled by various muscles.
12. <u>Valves – Cross-Cavity Airflow Regulator</u>: Valves regulate airflow between cavities. Airflow occurs when an air valve is open and there is a pressure difference between the connecting cavities. When the air valve is closed, there is no airflow.
13. <u>Air Valves of the Articulatory System</u>: The air valves are the vocal folds – the glottis – which regulate between the supra-glottal and the sub-glottal cavities, the velopharyngeal folds which regulate between the oral and the nasal cavities, the tongue which regulates between the oral cavity and the atmosphere, and the lips, which also regulate between the oral cavity and the atmosphere.
14. <u>Control of the Air Valves</u>: Like the pistons, the air valves are also controlled by various muscles.

Initiation

To produce any kind of sound, there must be a movement of air. To produce sounds people can interpret as spoken words, the movement of air must pass through the vocal cords, up through the throat, and into the mouth or the nose to then leave the body. Different sounds are formed by different positions of the mouth – or as linguists call it, *the oral cavity*, to distinguish it from the nasal cavity.

Vowels

1. Airstream Passage during Vowel Production: Vowels are produced by the passage of the air through the larynx and the vocal tract. Most vowels are voiced, i.e., the vocal folds are vibrating.
2. Open Vocal Tract/No Obstruction: Except in some marginal cases, the vocal tract is open, so the airstream is able to escape without generating fricative noise.
3. Control of Vowel Quality Variation: Variation in vowel quality is produced using the articulatory structures described below.

Articulators – Glottis

1. Vibration Produced by the Glottis: The glottis is the opening between the vocal folds located in the larynx. Its position creates different vibration patterns to distinguish voiced and voiceless sounds (Laver (1994)).
2. Pitch Control using Vocal Folds: In addition, the pitch of the vowel is altered by changing the frequency of vibration of the vocal folds. In some languages there are contrasts among vowels with different phonation types (Ladefoged and Maddieson (1996)).

Articulators – Pharynx

1. Pharynx Location and Pharyngealized Vowels: The pharynx is the region of the vocal tract below the velum and above the larynx. Vowels may be pharyngealized – also *epiglottalized*,

sphincteric, and *strident* – by means of a retraction of the tongue root (Ladefoged and Maddieson (1996)).

2. Vowels Produced using ATR Feature: Vowels may be articulated with Advanced Tongue Root ATR (Laver (1994)). There is a discussion as to whether the ATR vowel feature is different from the Tense/Lax distinction in vowels (Ladefoged and Maddieson (1996)).

Velum

1. Nasals Produced by Raising the Velum: The velum – or the soft palate – controls the air flow through the nasal cavity. Nasals and nasalized sounds are produced by lowering the velum and allowing air to escape through the nose.
2. Nasalization in Vowel Production: Vowels are normally with the soft palate raised so that no air escapes through the nose. However, vowels may also be nasalized by lowering the soft palate. Many languages use nasalization constructively (Ladefoged and Maddieson (1996)).

Tongue

1. Tongue Control of Vowel Articulation: The tongue is a highly flexible organ that is capable of being moved in many different ways. For vowel articulation, the principal variations are the vowel height and the dimensions of the backness and the frontness (Ladefoged and Maddieson (1996)).
2. Production of Rhotic/Rhotacized Vowels: A less common variation in vowel quality can be produced by a change in the shape of the front tongue, resulting in a rhotic or a rhotacized vowel (Ladefoged and Maddieson (1996)).

Lips

The lips play a major role in vowel articulation. It is generally believed that two major variables are in effect: lip-rounding or labialization, and lip protrusion.

Airflow

1. Boyle's Law for Articulatory Cavities: For all practical purposes, temperature can be treated as constant in the articulatory system. Thus, Boyle's law can be written as the following two equations:

$$P_1V_1 = P_2V_2$$

$$\frac{V_1}{V_1 + \Delta V} = \frac{P_1 + \Delta P}{P_1}$$

2. Application to the Subglottal Cavity: As applied to the description of the subglottal cavity, when the lung pistons contract the lungs, the volume of the subglottal cavity decreases while the subglottal air pressure increases. Conversely, if the lungs are expanded, the air pressure decreases.
3. Mouth Open + Vocal Folds Closed: A situation can be considered where:
   a. The vocal fold valve is closed separating the supraglottal cavity from the subglottal cavity,

b. The mouth is open and, therefore, the supraglottal air pressure is equal to the atmospheric pressure, and
c. The lungs are contracted resulting in a subglottal pressure that has increased to a pressure greater than the atmospheric pressure.

4. Opening the Vocal Cord Valve: If the vocal fold valve is subsequently opened, the previously two separate cavities become one unified cavity although the cavities will still be aerodynamically isolated because the glottalic valve between them is really small and constrictive.
5. Pressure Inequalities over Glottal Cavities: Pascal's Law states that the pressure within a system must equalize throughout the system. When the subglottal pressure is greater than the supraglottal pressure, there is pressure inequality in the unified cavity.
6. Airflow across Cavities until Pressure Equalizes: Since pressure is a force applied to a surface area by definition and a force is the product of mass and acceleration according to Newton's second Law of Motion, the pressure inequality will be resolved by having a part of the mass of air molecules found in the subglottal cavity move to the supraglottal cavity. This movement of mass is airflow. The airflow will continue until a pressure equilibrium is reached.
7. Glottalic Airstream Mechanism - Closed Cavity: Similarly, in an ejective consonant with a glottalic airstream mechanism, the lips or the tongue, i.e., the buccal or the lingual valve, are initially closed and the closed glottis – the laryngeal position – is raised decreasing the oral cavity volume behind the valve closure and increasing the pressure compared to the volume and the pressure at a resting state.
8. Glottalic Airstream Mechanism - Open Cavity: When the closed valve is opened, airflow will result from the cavity behind the initial closure outward until intra-oral pressure is equal to the atmospheric pressure.
9. Pressure Equalization by Airflow Movement: That is, air will flow from a cavity of higher pressure to a cavity of lower pressure until the equilibrium point; the pressure as potential energy is, thus, converted into airflow as kinetic energy.

Sound Sources

1. Periodic and Aperiodic Sound Sources: Sound sources refer to the conversion of aerodynamic energy into acoustic energy. There are two main types of sound sources in the articulatory system: periodic – or more precisely, semi-periodic, and aperiodic.
2. Origin of Periodic Sound Source: A periodic sound source is a vocal fold vibration produced at the glottis found in vowels and consonants. A less common periodic sound source is the vibration of the oral articulator like the tongue found in alveolar trills.
3. Origin of Aperiodic Sound Source: Aperiodic sound sources are the turbulent noises of fricative consonants and the short-noise burst of plosive releases produced in the oral cavity.
4. Voicing/Unvoicing in Vocal Cords: Voicing is a common period sound source in spoken languages and is related to how closely vocal cords are placed together. In English, there are only two possibilities, *voiced* and *unvoiced*.
5. Vocal Cord Voicing for Vowels: Voicing is caused by vocal cords being held close together, so that air passing through them makes them vibrate. All normally spoken vowels are voiced, as are all sonorants except *h*, as well as some of the remaining sounds (*b, d, g, v, z, zh, j,* and the *th* found in *this*).
6. Origin of Vocal Cord Voicelessness: All the rest are voiceless sounds, with the vocal cords held for enough apart that there is no vibration; however, there is still a certain amount of audible friction, as in the sound *h.*
7. Prominence of Voiceless Speech Sounds: Voiceless sounds are not very prominent unless there is some turbulence, as in stops, fricatives, and affricates; this is why sonorants in general only occur voiced. The exception is during whispering, when all sounds are voiceless.

Periodic Sources

1. Non-vocal Fold Vibration: 20-40 Hz
2. Vocal Fold Vibration:
    a. Lower Limit => 70-80 Hz modal, i.e., bass; 30-40 Hz creaky
    b. Upper Limit => 1170 Hz, i.e., soprano

References

- Ladefoged, P., and I. Maddieson (1996): *The Sounds of World's Languages* **Blackwell** Oxford, UK
- Laver, J. (1994): *Principles of Phonetics* **Cambridge University Press** Cambridge, UK
- Wikipedia (2021): Articulatory Phonetics

Auditory Phonetics

Overview

1. Focus of Auditory Phonetics Subfield: *Auditory Phonetics* is the branch of phonetics concerned with the hearing of speech sounds and with speech perception. It thus entails the study of relationships between speech stimuli and a listener's response to such stimuli mediated by the mechanisms of peripheral and the central auditory systems, including certain acoustic areas of the brain (Wikipedia (2021)).
2. One of Three Phonetics Branches: It is said to comprise one of three main branches of phonetics along with acoustic and articulatory phonetics (O'Connor (1973), Mack (2004)).

Physical Scales and Auditory Sensations

1. Auditory Sensation vs. Acoustic Properties: There is no direct connection between the auditory sensations and the physical properties of sound that give rise to them. While the physical/acoustic properties are objectively measurable, auditory sensations are subjective and can only be studied by asking listeners to report on their perceptions (Denes and Pinson (1993)).
2. Physical Property/Auditory Perception Mapping: The table below shows some correspondence between physical properties and auditory sensations.

Physical Property	**Auditory Perception**

Amplitude or Intensity	Loudness
Fundamental Frequency	Pitch
Spectral Structure	Sound Quality
Duration	Length

Segmental and Suprasegmental

1. Segmental/Prosodic Aspects of Speech: Auditory phonetics is concerned with both segmental – chiefly vowels and consonants – and prosodic – such as stress, tone, rhythm, and intonation – aspects of speech. While it is possible to study the auditory perception of these phenomena without context, in continuous speech all these variables are processed in parallel with significant variability and complex interactions between them (Wood (1974), Elman and McClelland (1982)).
2. Example - Intrinsic Frequencies of Vowel Formants: For example, it has been observed that vowels, which are usually described as different from each other in the frequencies of their formants, also have intrinsic values of fundamental frequency – and presumably therefore of pitch – that are different according to the height of the vowel.
3. Fundamental Vowel Frequencies - Open/Cloud: Thus, open vowels have a lower fundamental frequency than closed vowels in a given context (Turner and Verhoeven (2011)), and vowel recognition is likely to interact with the perception of prosody.

In Speech Research

1. Auditory Phonetics vs. Speech Perception: If there is a distinction to be made between auditory phonetics and speech perception, it is that the former is more closely associated with the traditional non-instrumental approaches to phonology and other aspects of linguistics, while the latter is closer to experimental laboratory-based study.
2. Instrument Usage in Auditory Phonetics: Consequently, the term *auditory phonetics* is used to refer to the study of speech without the use of instrumental analysis: the researcher may use if technology such as recording equipment, or even a simple paper and pencil – as used by Labov (1966) in the study of English in New York department stores, but will not use laboratory techniques such as spectrography or speech synthesis, or methods such as EEG and fMRI that allows phoneticians to directly study brain's response to sound.
3. Research using Auditory Analysis: Most research in sociolinguistics and dialectology has ben based on auditory analysis of data, and almost all pronunciation dictionaries are based on impressionistic, auditory analysis of how words are pronounced.
4. Definitions given by Pike/Pilch: It is possible to claim an advantage for auditory analysis over instrumental: "Auditory analysis is essential to phonetic study since the ear can register all those features of sound waves, and only those features, which are above the threshold of audibility … whereas analysis by instruments must always be checked against auditory reaction". Pilch (1978) attempted to define auditory phonetics in such a way as to avoid any reference to acoustic parameters.
5. Auditory Training in Speech Phones: In the auditory analysis of phonetic data such as recordings of speech, it is clearly an advantage to have been trained in analytical listening. Practical phonetic training has since the 19th century been seen as an essential foundation for phonetic analysis and for the teaching of pronunciation; it is still a significant part of modern phonetics.
6. Examples: Trainings in Cardinal Vowels: The best-known system of auditory training has been in the system of cardinal vowels; there is disagreement about the auditory and the articulatory factors in the underlying system, but the importance of auditory training for those who use it are indisputable (Ladefoged (1967)).
7. Training in Prosodic Speech Factors: Training in the auditory analysis of prosodic factors such as pitch and rhythm is also important.

8. Instrumental Approach in Prosody Research: Not all research on prosody has been based on auditory techniques; some pioneering work on prosodic features using laboratory instruments was carried out in 20th century, e.g., Elizabeth Uldall's work on synthesized intonation contours, Fry's work on stress perception (Fry (1954)), or Jones' early work on analyzing pitch contours by means of manually operating the pickup arm of a gramophone to listen repeatedly to individual syllables, checking where necessary against a tuning fork (Jones (1909)).
9. Computer Approaches in Prosody Research: However, a good majority of the work on prosody has been based on the auditory analysis until the recent arrival of approaches explicitly based on computer analysis of the acoustic signal, such as ToBI, INTSINT, or the IPO system (t'Hart, Collier, and Cohen (1996)).

References

- Denes, P., and E. Pinson (1993): *The Speech Chain 2nd Edition* **W. H. Freeman** New York, NY
- Elman, J., and J. McClelland (1982): Exploiting Lawful Variability in the Speech Wave, in: *Invariance and Variability in Speech Processes (Editors: Perkell, J. S., and D. Klatt)* **Erlbaum** Mahwah, NJ
- Fry, D. (1954): Duration and Intensity as Physical correlates of Linguistic Stress *Journal of the Acoustical Society of America* **27 (4)** 765-768
- Jones, D. (1909): *Intonation Curves* **Teubner** Leipzig, Germany
- Labov, W. (1966): *The Social Stratification of English in New York City* **Center for Applied Linguistics** Washington, DC
- Ladefoged, P. (1967): *Three Areas of Experimental Phonetics* **Oxford University Press** Oxford, UK
- Mack, M. (2004): Auditory Phonetics, in: *The Linguistics Encyclopedia (Editor: Malmkjaer, K.)* **Routledge** Oxfordshire, UK

- O'Connor, J. D. (1973): *Phonetics 1st Edition* **Penguin** New York, NY
- Pike, K. (1943): *Phonetics* **University of Michigan** Ann Arbor, MI
- Pilch, H. (1978): Auditory Phonetics *Word* **29 (2)** 148-160
- t'Hart, J., R. Collier, and A. Cohen (1990): *A Perceptual Study of Intonation* **Cambridge University Press** Cambridge, UK
- Turner, P., and J. Verhoeven (2013): Intrinsic Vowel Pitch: A Gradient Feature of Vowel Systems? *Proceedings of the International Congress of Phonetic Sciences* 2038-2041
- Wikipedia (2021): Auditory Phonetics
- Wood, C. C. (1974): Parallel Processing of Auditory and Phonetic Information in Speech Discrimination *Perception and Psychophysics* **15 (3)** 501-508

Phonology - Introduction

Overview

1. Alternate Meanings of the Term Phonology: *Phonology* is a branch of linguistics that studies how languages or dialects systematically organize their sounds – or signs, in sign languages. The term also refers to the sound system of any particular language variety.
2. Focus of the Phonology Field: At one time, the study of phonology only related to the study of systems of phonemes in spoken languages. Now it may relate to:
    a. Any linguistic analysis either at a level beneath the word – including syllables, onset and rime, articulatory gestures, articulatory features, mora, etc. OR
    b. All levels of language where sounds or signs are structured to convey linguistic meaning (Brentari, Fenlon, and Cormier (2018)).
3. Phonological Equivalents in Sign Languages: Sign languages have a phonological system equivalent to the system of sounds in spoken languages. The building blocks of signs are specifications for movement, location, and hand shape (Stokoe (1978)).

Terminology

1. Phonology as a Language System Component: The word *phonology* – as in *the phonology of English* – can also refer to the phonological system – the sound system – of a given language. This is one of the fundamental systems that a language is considered to comprise, like its syntax, its morphology, and its vocabulary.
2. Distinction between Phonology and Phonetics: Phonology is often distinguished from *phonetics*. While phonetics concerns the physical production, acoustic transmission, and perception of the sounds of speech (Lass (1998), Carr (2003)), phonology describes the way sounds function within a given language or across languages to encode meaning.

3. Distinction between Theoretical/Descriptive Linguistics: For many linguists, phonetics belongs to descriptive linguistics, and phonology to theoretical linguistics, although establishing a phonological system of a language is necessarily an application of theoretical principles to the application of phonetic evidence.
4. Conflation between Phonology and Phonetics: This distinction was not always made, particularly before the development of the modern concept of the phoneme in the mid-20th century.
5. Crossover of Phonology with Phonetics: Some sub-fields of modern phonology have a cross-over with phonetics in descriptive disciplines such as psycholinguistics and speech perception, resulting in specific areas such as articulatory phonology or laboratory phonology.

Derivation and Definitions

1. Origin of the Term Phonology: The word *phonology* comes from the ancient Greek φωυη, *phone, voice, sound*, and the suffix -*logy*, which is from the Greek λογος, *logos, word, speech*, *subject of discussion*.
2. Trubetzkoy's Definition of the Term: Trubetzkoy (1939) defines phonology as *the study of sound pertaining to the system of language*, as opposed to phonetics, which is *the study of sound pertaining to the act of speech* – the distinction between *language* and *speech* being basically Saussure's distinction between *langue* and *parole*.
3. Lass Definition of the Term: Lass (1998) writes the phonology broadly refers to the sub-discipline of linguistics concerned with the sounds of language, while in more narrow terms, *phonology proper is concerned with the function, behavior, and organization of sounds as linguistic items*.
4. Definition of Clark, Yallop, and Fletcher: According to Clark, Yallop, and Fletcher (2007), it means the systematic use of sound to encode meaning in any spoken human language, or the field of linguistics studying this use.

Analysis of Phonemes

1. Decomposed Units of Distinctive Sounds: An important part of traditional, pre-generative schools of philosophy is studying which sounds can be grouped into distinctive units within a language; these units are known as phonemes.
2. Example: Phoneme Units in English: For example, in English the *p* sound in *pot* is aspirated, while that in *spot* is not aspirated. However, English speakers treat both sounds as variations/allophones – of the same phonological category, that is of the phoneme *p*. Traditionally, it would be argued that if an aspirated *p* were interchanged with an unaspirated *p* in *spot*, native English speakers will still hear the same words; that is, the two sounds are perceived as *the same* p.
3. Phoneme Units in other Languages: In some other languages, however, these two sounds are perceived as different, as they are consequently assigned to different phonemes. For example, in Thai, Hindi, and Quechua, there are minimal pairs of words for which the aspiration is the only contrasting feature – two words can have different meanings but with the only difference in pronunciation being that one has an aspirated sound where the other has an unaspirated one.
4. Sound Inventory of Native Speakers: Part of the phonological study of language therefore involves looking at data – phonetic transcriptions of the speech of native speakers – and trying to decide what the underlying phonemes are and what the sound inventory of the language is.
5. Criteria for Identifying Minimal Pairs: The presence or absence of minimal pairs, as mentioned above, is a frequently used criteria for deciding whether two sounds should be assigned to the same phoneme. However, other considerations often need to be taken into account as well.
6. Historical Evolution of Language Phonemes: The particular contrasts which are phonemic in a language can change over time. At one time, [f] and [v], two sounds that have the same place and the manner of articulation and differ in voicing only, were allophones of the same phoneme in English, but later come to belong to separate phonemes. This is one of the main factors of historical change of languages as described in historical linguistics.

7. Interchanging the Allophones of Phonemes: The findings and insights of speech perception and articulation research complicate the traditional and somewhat intuitive idea of interchangeable allophones being perceived as the same phoneme.
8. Gibberish resulting from Allophone Switch: First, interchanged allophones of the same phoneme can result in unrecognizable words.
9. Highly Co-articulated Low-level Speech: Second, actual speech, even at a word level, is highly co-articulated, so it is problematic to be able to splice words into simple segments without affecting speech perception.
10. Assigning Sounds to Individual Phonemes: Different linguists therefore take different approaches to the problem of assigning sounds to phonemes.
11. Constraints around Allophone Sounds: For example, they differ in the extent to which they require the allophones to be phonetically similar.
12. Equivalence with the Brain Functions: There are also differing ideas as to whether this grouping of sounds is purely a tool for linguistic analysis, or reflects an actual process in the way human brain processes a language.
13. Idea behind Morphophonemes and Morphophonology: Since the early 1960s. theoretical linguistics have moved away from the traditional concept of a phoneme, preferring to consider the basic units at a more abstract level, as a component of morphemes; these units are called *morphophonemes*, and analysis using this approach is called morphophonology.

Other Topics in Phonology

1. Aspects of Phonological Studies - #1: In addition to the minimal units that can serve the purpose of differentiating meaning – the phonemes, phonology studies how sounds alternate, i.e., replace one another in different forms of the same morpheme – allomorphs, as well as, for example, syllable, stress, feature geometry, and intonation.
2. Aspects of Phonological Studies - #2: Phonology also includes such topics as phonotactics – the phonological constraints on what sounds can appear in what positions in a given language – and phonological alternation – how the pronunciation of a sound changes through the application of phonological rules, sometimes in a given order which can be feeding or

bleeding (Goldsmith (1995)), as well as prosody, the study of supra-segmentals, and topics such as stress and intonation.

3. Phonology applied to Sign Languages: The principles of phonology can be applied independently of modality because they are designed to serve as general analytical tools, not language specific ones. The same principles have been applied to analysis of the sign languages, even though the sub-lexical units are not instantiated as speech sounds (Wikipedia (2020)).

References

- Brentari, J. Fenlon, and K. Cormier (2018): *Sign Language Phonology* **Oxford Research Encyclopedia of Linguistics**
- Carr, P. (2003): *English Phonetics and Phonology: An Introduction* **Blackwell Publishing**
- Clark, J., C. Yallop, and J. Fletcher (2007): *An Introduction to Phonetics and Phonology 3rd edition* **Blackwell Publishing**
- Goldsmith, J. A. (1995): *The Handbook of Phonological Theory* **Blackwell Publishers**
- Lass, R. (1998): *Phonology: An Introduction to Basic Concepts* **Oxford University Press**
- Stokoe, W. C. (1978): *Sign Language Structure – An Outline of the Visual Communication Systems for the American Deaf* **University of Buffalo**
- Trubetzkoy, N. (1939): *Grundzuge der Phonologie* **Travaux der Cercle Linguistique de Prague 7**
- Wikipedia (2020): Phonology

Phoneme

Overview

1. <u>Definition of a Phoneme</u>: A *phoneme* is a unit of sound that distinguishes one word from another in a particular language (Wikipedia (2020)).

2. <u>Meaning - Contrast through Minimal Pair</u>: For example, in most dialects of English, with the notable exception of West Midlands and the Northwest of England (Wells (1982)), the sound patterns for *sin* and *sing* are two separate words that are distinguished by the substitution of one phoneme */n/* for another phoneme */nj/*. Two like this that differ in meaning through the contrast of a single phoneme form a *minimal pair*.

3. <u>Phonetic Variants of a single Phoneme</u>: If, in another language, any two sequences differ only by the pronunciation of their final sounds *[n]* and $[n_j]$ are perceived as being the same in the meaning, then these two sounds are interpreted as phonetic variants of a single phoneme in that language.

4. <u>Notation for Representing a Phoneme</u>: Phonemes that are established by the use of minimal pairs, such as *tap* vs *tab* or *pat* vs *bat*, are written between slashes: */p/*, */b/*. To show pronunciation, linguists use square brackets: $[p^h]$ - indicating an aspirated *p* in *pat*.

5. <u>Analyzing a Language through its Phonemenes</u>: There are two different views as to exactly what phonemes are and how a given language should be analyzed in *phonemic* - or *phonematic* - terms.

6. <u>Conceptual Idea behind a Phoneme</u>: However, a phoneme is generally regarded as an abstraction of a set - or equivalence class - of speech sounds - *phones* - that are perceived as

equivalent to each other in a given language. For example, the English *k* sounds in the words *kill* and *skill* are not identical - as described below - but they are distributional variants of a single phoneme */k/*.

7. Allophonic Variants of a Phoneme: Speech sounds that differ but do not create a meaningful change in the word are known as *allophones* of the same phoneme.

8. Factors Contributing to Allophonic Realization: Allophonic variation may be conditioned, in which case a certain phoneme is realized as a certain allophone in particular phonological environments, or it may otherwise be free, and may vary by speaker of dialect.

9. Language Phonemes versus Speech Sounds: Therefore, phonemes are often considered to constitute an abstract underlying representation for segments of words, while speech sounds makeup the corresponding phonetic realization, or the surface form.

Notation

1. Phonemes vs. Speech Sound Representation: Phonemes are conventionally placed between slashes in transcription, whereas speech sounds - phones - are placed in square brackets. Thus, */p omega integral/* represents a sequence of three phonemes, */p/*, */omega/*, */integral/* - the word *push* in Standard English, and $[p^h$ *omega integral]* represents the phonetic sequence of sounds $[p^h]$ - aspirated *p* - *[v]*, *[integral]* - the usual pronunciation of *push*.

2. Orthography Representation using Grapheme Units: This should not be confused with the similar convention of the use of angle brackets to enclose the units of orthography, graphemes. For example, <*f*> represents the written letter - grapheme - *f*.

3. IPA Based Phoneme Symbol Set: The symbols for particular phonemes are often taken from the International Phonemic Alphabet (IPA), the same set of symbols most commonly used for phones. For computer-typing purposes, systems such as X-SAMPA exist to represent IPA

symbols using only ASCII characters.

4. Custom Transcription of Language Phonemes: However, descriptions of particular languages may use different conventional symbols to represent the phonemes of those languages. For languages whose writing systems employ the phonemic principle, ordinary letters my be used to denote phonemes, although this approach is often hampered by the complicated relationship between orthography and pronunciation.

Assignment of Speech Sounds to Phonemes

1. Uniqueness of Meaning/Speech Unit: A phoneme is a sound or a group of different sounds perceived to have the same function by the speakers of the language or dialect in question.

2. Example of Phoneme in English: An example is the English phoneme */k/*, which occurs in words such as *c*at, *k*it, s*c*at, and s*k*it. Although most native speakers do not notice this, in most English dialects, the *c/k* dialects are not identical; in *kit* $[k^h it]$ the sound is aspirated, but in *skill [skil]*, it is unaspirated.

3. Similarity of the Speech Sounds: The words, therefore, contain different *speech sounds*, or *phones*, transcribed $[k^h]$ for the aspirated form and *[k]* for the unaspirated one.

4. Variations of the Phonemic Unit: These sounds are nonetheless considered to belong to the same phoneme, because if the speaker used one instead of the other, the meaning of the word would not change; using the aspirated form $[k^h]$ in *skill* might sound odd, but the word would still be recognized.

5. Phone Change Induced by Meaning Difference: By contrast, some other sounds would cause a change in meaning if substituted; for example, substitution of the sound *[t]* would produce a different word still, and that sound must therefore be considered to represent a different phoneme - the phoneme */t/*.

6. Scheme for Identification of Phonemes: The following simplified procedure is used for determining whether two sounds represent the same or different phonemes.

7. Determining Phonemic Status of Sounds:

8. English Allophones *[k]* and *[k^h]*: The above shows that in English *[k]* and *[k^h]* are allophones of a single phoneme */k/*.

9. *[k]* and *[k^h]* as Non-phonemic: In some languages, however, *[k]* and *[k^h]* are perceived as different sounds, and substituting one for another can change the meaning of the word. In those languages, therefore, the sounds represent different phonemes.

10. Icelandic Language *[k]* and *[k^h]*: For example, in Icelandic, *[k^h]* is the first sound of *ka'tur*, meaning cheerful, but *[k]* is the first sound of *ga'tur*, meaning *riddler*. Icelandic, therefore, has two separate phonemes */k^h/* and */k/*.

Minimal Pair

1. Existence of the Minimal Pair: A pair of words like *ka'tur* and *ga'tur* above that differ only in one phone is called the minimal pair for the two alternative phones in question - in this case *[k]* and *[k^h]*.

2. Minimal Pair Check for Phonemes: The existence of minimal pairs is a common test to decide whether two phones represent different phonemes or are allophones of the same phoneme.

3. Minimal Pair Presence - English Example: To take another example, the minimal pair *t*ip and *d*ip illustrate that in English, *[t]* and *[d]* belong to separate phonemes */t/* and */d/*; since both words have different meanings. English speakers must be conscious of the distinction between the two sounds.

4. Minimal Pair Absence - Korean Example: In other languages, however, including Korean, both sounds *[t]* and *[d]* occur, but no such minimal pair exists.

5. *[t]* and *[d]* as Allophones: The lack of minimal pairs distinguishing *[t]* and *[d]* in Korean provides evidence that they are allophones of a single phoneme */t/*. The word */tada/* is pronounced *[tada]*, for example.

6. Perception Variation across different Languages: That is, when they hear this word, Korean speakers perceive the same sound in both the beginning and the end of the word, but English speakers perceive different sounds in these two locations.

7. Minimal Pairs in ASL Expressions: Sign languages, such as American Sign Languages ASL also have minimal Pairs, different only in exactly one of the sign parameters: handshape, movement, location, palm orientation, and non-manual signal or marker.

8. Parameters Guiding ASL Minimal Pair: A minimal pair may exist in the sign language if the basic sign remains the same, but one of the parameters changes.

9. Phonetic Marker Dissimilarity for Phonemes: However, the absence of minimal pairs for a given pair of phones does not always mean that they belong to the sane phoneme: they may be so dissimilar phonetically that it is unlikely for speakers to perceive them as the same sound.

10. Phonetic Marker Example in English: For example, English has no minimal pair for the sounds *[h]* - as in *h*at - and *[n,]* - as in ba*ng*, and the fact that they can be shown to be in complementary distribution could be argued for their being allophones of the same phoneme. However, they are so dissimilar phonetically that they are considered separate phonemes (Wells (1982)).

11. Case of "Near Minimal Pairs": Phonologists have sometimes had to recourse to "near minimal pairs" to show that speakers of the language perceive the two sounds as significantly different even if no minimal pair exists in the lexicon.

12. Near Minimal Pair English Example: It is virtually impossible to find a minimal pair to distinguish */integral/* from */z/*, yet it seems uncontroversial to claim that the two consonants

are distinct phonemes. The two words *pleasure* and *pressure* can serve as a minimal pair (Wells (1982)).

Suprasegmental Phonemes

1. Suprasegmental Phonemes Impact Word Meanings: Besides segmental phonemes such as vowels and consonants, there re also suprasegmental features of pronunciation - such as tone and stress, syllable boundaries, and other forms of juncture, nasalization, and vowel harmony - which, in many languages, can change the meaning of the words and so are phonemic.
2. Phonemic Stress Impacting Word Meanings: *Phonemic stress* is encountered in languages such as English. For example, the word *invite* stressed on the second syllable is a verb, but when stressed on the first syllable - without changing any of the individual sounds - it becomes a noun.
3. Phonemic Specification of the Word: The position of the stress in the word affects the meaning, so a full phonemic specification - providing enough detail to enable the word to be pronounced unambiguously - would include indication of the position of the stress: */inv'ait/* for the verb, */'invait/* for the noun.
4. Languages where Stress is Non-phonemic: In other languages, such as French, word stress cannot have this function - its position is generally predictable - and is therefore not phonemic, and is not usually indicated in dictionaries.
5. Phonemic Tones Impacting Word Meanings: *Phonemic tones* are found in languages such as Mandarin Chinese, in which a given syllable can have 5 different tonal pronunciations.
6. Phonetic Variants of the Word *Ma*:
7. Meanings Induced by Tonal Variations: Here, the character pronounced *m~~a~~* - high level pitch - means *mother*; *ma'*, rising pitch, means *hemp*; *mau*, falling then rising, means *horse*; *ma`*,

falling, means *scold*; and *ma*, neutral tone, is an interrogative particle. The tone phonemes in such languages are called *tonemes*.

8. Phonemic Intonation Impacting Word Meanings: Languages such as English do not have phonemic tones, although they use intonation for functions such as emphasis and attitude.

Distribution of Allophones

1. Complementary Distribution of Allophones: When a phoneme has more than one allophone, the only actually heard at the given occurrence of that phoneme may be dependent on the phonetic environment, i.e., surrounding sounds; allophones which normally cannot appear in the same environment are said to be in the complementary distribution.
2. Free Variation in the Allophones: In other cases, the choice of the allophone may be dependent on the individual speaker or other unpredictable factors - such allophones are said to be in free variation, but allophones are still selected in a specific phonetic context, not the other way around.

Background and Related Ideas

1. Meaning of the Greek Word: The term *phoneme* - from the ancient Greek *pho-ne-ma* - "sound made, utterance, thing spoken, speech, language" (Liddell and Scott (1940)) - was reportedly used first by Dufriche-Desgenettes in 1873, but it referred only to a speech sound.
2. Fonema - Basic Unit of Psychophonetics: The term *phoneme* as an abstraction was developed by the Polish linguist Jan Niecislaw Baudouin de Courtenay and his student Nikolaj

Kruszewski during 1875-1895 (Jones (1957)). The term used by these two was *fonema*, the basic unit of what they called *psychophonetics*.

3. <u>Modern Usage of the Word Phoneme</u>: Jones (1919) became the first linguist in the Western world to use the word *phoneme* in the current sense.

4. <u>Elaboration of the Phoneme Concept</u>: The concept of the phoneme was then elaborated in the works of Nikolai Trubetzkoy and others of the Prague during the years 1926-1935, and in those of the structuralists like Ferdinand de Saussure, Edward Sapir, and Leonard Bloomfield.

5. <u>Psycholinguistic Role for Phonemes</u>: Some structuralists - though not Sapir - rejected the idea of a cognitive or psycholinguistic function for the phoneme (Twaddell (1935), Harris (1951)).

6. <u>Deprecation/Enhancement of the Phoneme Concept</u>: Later, it was used and redefined in generative linguistics, most famously by Chomsky and Halle (1968), and remains central to many accounts of the development of modern phonology. As a theoretical concept or model, however, it has been supplemented and even replaced by others (Clark and Yallop (1995)).

7. <u>Decomposition of Phonemes into Features</u>: Some linguists - such as Jakobson and Halle (1968) - proposed may be further decomposed into features, such features being the minimal constituents of language.

8. <u>Evolution of Sub-phonemic Features</u>: Features overlap each other in time, as do suprasegmental phonemes in oral languages and many phonemes in sign languages.

9. <u>Schemes for Extracting the Features</u>: Features can be characterized in different ways: Jakobson, Fant, and Halle (1952) described them in acoustic terms, Chomsky and Halle used a predominantly articulatory basis, though retaining some acoustic features, while Ladefoged's system (Ladefoged (2006)) is purely an articulatory system apart from the use of the acoustic term 'sibilant'.

10. <u>Duration Chronemes and Tone Phonemes</u>: In the description of some languages, the term *chroneme* has been used to indicate the contrastive length or *duration* of phonemes. In

languages in which tonemes are phonemic, the tone phonemes may be called tonemes.

11. Widespread Acceptance of the Above Terms: Though not all scholars working on such languages use these terms, they are by no means obsolete.

12. Other Fundamental Units in Linguistics: By analogy with the phoneme, linguists have proposed other sorts of underlying objects, giving them names with the suffix *-eme*, such as *morpheme* and *grapheme*. These are sometimes called emic units.

13. Generalization of Emics and Etics: The latter term was first used by Pike (1967), who generalized the concepts of emic and etic descriptions - from *phonemic* and *phonetic* respectively - to applications outside linguistics.

Restrictions on Occurrence

1. Phonotactic Combinations Constraints - Restricted Phonemes: Languages do not allow generally words or syllables to be built of any arbitrary sequence of phonemes; there are phonotactic restrictions on which sequences are possible and in which environments certain phonemes can occur. Phonemes that are significantly limited by such restrictions may be called *restricted phonemes*.

2. Phonemic Restrictions in English #1: In English, examples of such restrictions include: /n_j/, as in *si**ng***, occurs only at the end of a syllable, never at the beginning - in many other languages, such as Maori, Swahili, Tagalog, and Thai, /n_j/ can appear word-initially.

3. Phonemic Restrictions in English #2: /h/ occurs only before vowels and at the beginning of a syllable, never at the end - a few languages, such as Arabic or Romanian, allow /h/ syllable-finally.

4. Phonemic Restrictions in English #3: In non-rhotic dialects, /inverse_r/ can only occur before a vowel, never before a consonant.

5. <u>Phonemic Restrictions in English #4</u>: /w/ or /j/ occur only before a vowel, never at the end of a syllable - except in interpretations where a word like *boy* is analyzed as /b mirrored_c j/.

6. <u>Analysis using Neutralization and Archiphonemes</u>: Some phonotactic restrictions can alternatively be analyzed as cases of neutralization. In the below section on Neutralization and archiphonemes, a particular example of the occurrence of the three English nasals before stops is shown.

Biuniqueness

1. <u>Meaning of the Biuniqueness Requirement</u>: Biuniqueness is a requirement of the classical structuralist phonemics. It means that a given phone, wherever it occurs, must be unambiguously assigned to one and only one phoneme. In other words, the mapping between phones and phonemes is required to be many-ti-one rather than many-to-many.

2. <u>Controversial Nature of the Postulate</u>: The notion of biuniqueness was controversial among some pre-generative linguists and was prominently challenged by Halle and Chomsky in the late 1950s and early 1960s.

3. <u>Alveolar Flaps as a Counter-point</u>: An example of the problems arising frm the biuniqueness requirement is provided by the phenomenon of flapping in North American English. This may cause either /t/ or /d/ - in the appropriate environments - to be realized with the phone [snipped r] - an alveolar flap.

4. <u>Non-contrastive Phonemes - Contextual Realization</u>: For example, the same flap sound may be heard in the words *hi**tt**ing* and *bi**dd**ing*, although it is intended to realize the phoneme /t/ in the first word and /d/ in the second. This appears to contradict biuniqueness. The next section has a detailed discussion of such cases.

Neutralization an Archiphonemes

1. Neutralization of the Phonemic Contrast: Phonemes that are contrastive in certain environments may not be contrastive in all environments. In environments where they do not contrast, the contrast is said to be *neutralized*. In these positions, it may become less which phoneme a given phone represents.

2. Non-realized Phonemes - Absolute Neutralization: *Absolute Neutralization* is a phenomenon in which a segment of the underlying realization is not realized in any of phonetic representations.

3. Non-contrastive Phonemes - Contextual Realization: The term was introduced by Kiparsky (1968), and contrasts with *contextual neutralization* where some phonemes are not contrastive in certain environments.

4. Representation using Under-specification - Archiphoneme: Some phonologists prefer not to specify a unique phoneme in such cases, since to do so would mean providing redundant or even arbitrary information - instead they use the technique of under-specification. An *archiphoneme* is an object sometimes used to represent an under-specified phoneme.

5. Example: Stressed/Unstressed Contrastive Realizations: An example of neutralization is provided by the Russian vowels /a/ and /o/. These phonemes are contrasting in stressed syllables, but in unstressed syllables the contrast is lost, since both are reduced to the same sound - usually [flipped_e] - owing to vowel reduction in Russian.

6. Factors Impacting the Phonemic Assignment: In order to assign such an instance of [flipped_e] to one of the phonemes /a/ and /o/, it is necessary to consider morphological factors, such as the vowels that occur in other forms of the words, or which inflectional pattern is followed. In some cases, this may not even provide an unambiguous answer.

7. Using Under-specification for Description: A description using the approach of under-specification would not attempt to assign [flipped_e] to a specific phoneme in some or all of

the cases, although it may be assigned to an archiphoneme, written something like //A//, which reflects two neutralized phonemes in this position.

8. <u>English Example - Contrasting Nasal Phonemes</u>: A somewhat different example is found in English, with the three nasal phonemes /m, n, n_j/. In word-final position, all these contrast, as shown by the minimal triplet *sum* /s^m/, *sun* /s^n/, *sung* /s^n_j/.

9. <u>Exclusiveness of Nasals Preceding Stops</u>: However, before a stop such as /p, t, k/ - provided there is no morpheme boundary between them - only one of the nasals is possible in any given position: /m/ before /p/, /n/ before /t/ or /d/, and /n_j/ before /k/, as in *limp*, *link*, *link*, - /limp/, /lint/, /lin_jk/.

10. <u>Non-contrastive Nature of these Phonemes</u>: The nasals are therefore not contrastive in these environments, and according to some theorists this makes it inappropriate to assign the nasal phones heard here to any of the phonemes - even though, in this case, the phonetic evidence is unambiguous.

11. <u>Archiphonemic Representation of these Nasals</u>: Instead, they may analyze these phones as belonging to a single archiphoneme, written something like //N//, and state the underlying representations of *limp*, *link*, *link* to be //liNp//, //liNt//, //liNk//.

12. <u>Alternate Notation for Representing Archiphonemes</u>: This latter type of analysis is often associated with Nikolai Trubetzkoy of the Prague School. Archiphonemes are often notated with a capital letter within double virgules or pipes, as with examples //A// and //N// given above. Other ways the second of these has been notated include |m-n-n_j|, {m, n, n_j}, and //n*//.

13. <u>English Example - Alveolar Flap Phonemes</u>: Another example from English, but this time involving complete phonetic convergence as in the Russian example, is the flapping of /t/ and /d/ in some accents of American English - described above under Biuniqueness.

14. <u>Phonemes Implied by Consistent Flapping</u>: Here the word *betting* and *bedding* may ne pronounced ['b epsilon snipped_r i n_j]. Under the generative grammar theory of linguistics, if a speaker applies such flapping consistently, morphological evidence - the pronunciation of the related forms *bet* and *bed*, for example - would reveal which phoneme the flap represents,

once it is known which morpheme is being used (Dinnsen (1985)).

15. Archiphoneme Approach to Flap Determination: However, other theorists would prefer not to make such a determination, and simply asign the flap in both cases to a single archiphoneme, written - for example - //D??.

16. English Example - Plosives Succeeding /s/: Further mergers in English are plosives after /s/, where /p, t, k/ conflate with /b, d, g/, as suggested by the alternative spellings *sketti* and *sghetti*. That is, there is no particular reason to transcribe *spin* as /'spin/ rather than as /'sbin/, other than its historical development, and it may be less ambiguously transcribed as //'sBin//.

Morphemes

1. Sub-division into Morphophonemes and Morphemes: A *morphophoneme* is a theoretical unit at a deeper level of abstraction than traditional phonemes, and it taken to be a unit from which morphemes are built up.

2. Dividing Allomorphs to Uncover Morphophonemes: A morphophoneme within a morpheme can be expressed in different ways in different allomorphs of that morpheme - according to morphophonological rules.

3. Morphophonemic Representation of English Plurals: For example, the English plural morpheme -*s* appearing in words such as *cats* and *dogs* can be considered to be a single morphophoneme, which might be transcribed - for example - //z// or |z|, and which is realized as phonemically |s| after most voiceless consonants - as in *ca**ts*** - and as |z| in other cases - as in *dog**s***.

Number of Phonemes in Different Languages

1. Phones Produced by Natural Languages: All known languages use only a small subset of the many possible sounds that the human speech organs can produce, and, because of allophones, the number of distinct phonemes will generally be smaller than the number of identifiably different sounds.

2. Phonemic Inventory Range across Languages: Different languages vary considerably in the number of phonemes that have in their systems, although apparent variation might sometime result from the different approaches taken by the linguists doing the analysis. The total phonemic inventory in languages varies from as few as 11 in Rotokas and Piraha to as many as 141 in !Xu~ (Crystal (2010)).

3. Lowest Count of Vowel Phonemes: The number of phonemically distinct vowels can be as low as 2, as in Ubuyk and Arrernte.

4. Highest Count of Vowel Phonemes: At the other extreme, the Bantu language Ngwe has 14 vowel qualities, 12 of which may occur long or short, making 26 oral vowels, plus 6 nasalized vowels, long and short, making a total of 38 vowels; while !Xo'o~ achieves 31 pure vowels, not counting the additional variation by vowel length, by varying the phonation.

5. Lowest Count of Consonant Phonemes: As regards consonant phonemes, Puinavae and the Papuan language Tauade each have just 7, and Rotokas has only 6.

6. Highest Count of Consonant Phonemes: !Xo'o~, on the other hand, has somewhere around 77, and Ubykh 81.

7. Vowel Phoneme Range in English: The English language uses a rather large set of 13-21 vowel phonemes, including diphthongs, although its 22-26 consonants are close to average.

8. Phonemes due to Tones/Stress: Some languages, such as French, have no phonemic tone or stress; while Cantonese and several other Kim-Sui languages have 9 tones, and one of the Kui languages, Wobe', has been claimed to have 14 (Bearth and Link (1980)), though this is disputed (Singler (1984)).

9. Common Vowel/Consonant Phoneme Set: The most common vowel system consists of 5 vowels /i/, /e/, /a/, /o/, /u/. The most common consonants are /p/, /t/, /k/, /m/, /n/ (Moran, McCloy, and Wright (2014)).

10. Languages that lack Common Consonants: Relatively few languages lack any of these consonants, although it does happen: for example, Arabic lacks /p/, standard Hawaiian lacks /t/, Mohawk and Tlingit lack /p/ and /m/, Hupa lacks both /p/ and a simple /k/, colloquial Samoan lacks /t/ and /n/, while Rotokas and Quileate lack /m/ and /n/.

The Non-Uniqueness of Phonemic Solutions

1. Uniqueness of the Phonemic Construct: During the development of the phoneme theory in the mid-20th century phonologists were concerned not only with the procedures and the principles involved in producing a phonemic analysis of the sounds in a given language, but also with reality or uniqueness of the phonemic solution.

2. Pike's Statement on Phonemic Uniqueness: Some writers took the position expressed by Pike (1947): "There is only one accurate phonemic analysis of a given set of data", while others believed that different analysis, equally valid, could be made for the same data.

3. Chao's Statement on Phonemic Uniqueness: Chao (1934) stated: "Given the sounds of a language, there are usually more than one possible way of reducing them to a set of phonemes, and those different systems or solutions are not simply correct or incorrect, but may be regarded as only being good or bad for various purposes".

4. Analysis Using English Vocal System: Householder (1952) referred to this debate within linguistics as "God's truth vs. hocus-pocus". Different analysis of the English vowel system may be used to illustrate this.

5. Wikipedia on English Vowel Phonemes: The article on English phonology (Wikipedia

(2021)) states that English has a particularly large number of vwel phonemes, and that there are 20 vowel phonemes in Received Pronunciation, 14-16 in General American, and 20-21 in Australian English; the previous section indicated that the English language uses a rather large set of 13-21 vowel phonemes.

6. <u>Alternate Transcription of English Phonemes</u>: Although these figures are often quoted as a scientific fact, they actually reflect only one of many possible analysis, and Wikipedia (2021) suggests an alternate analysis in which some diphthongs and long vowels may be interpreted as comprising a short vowel linked to either /j/ or /w/.

7. <u>Vowel Phonemes for RP</u>: The transcription system for British English (RP) devised by Lindsay in 2017 and used in the CUBE pronunciation dictionary also treats diphthongs as composed of a vowel plus /j/ or /w/.

8. <u>Exposition of Trager and Smith</u>: The fullest exposition of this approach is found in Trager and Smith (1951) where all long vowels and diphthongs - "complex nuclei" - are made up of a short vowel combined with either /j/, /w/, or /h/ - plus /r/ for rhotic accents, each thus comprising two phonemes. They write: "The conclusion is inescapable that complex nuclei consist each of two phonemes, one of the short vowels followed by one of the 3 glides".

9. <u>Alternate Transcriptions for the Words</u>: The transcription for the vowel normally transcribed /ai/ would be instead /aj/, /a v with horns/ would be /aw/, and /a:/ would be /ah/.

10. <u>Significantly Reduced Count of Vowels</u>: The consequence of this approach is that English could theoretically have only 7 vowel phonemes, which is symbolized /i/, /e/, /a/, /o/, /u/, /^/, and /flipped-e/, or even 6 if schwa were treated as an allophone of /^/ or of other short vowels, a figure that would put English much closer to the average number of vowel phonemes in other languages.

11. <u>Competing Basis for Phonemic Analysis</u>: In the same period there was disagreement about the correct basis for phonemic analysis.

12. <u>Analysis Using only Sound Elements</u>: The structuralist position was that the analysis should be made purely on the basis of the sound elements and their distribution, with no reference to extraneous factors such as grammar, morphology, or the intuitions of native speakers; this

position is associated with Bloomfield (1933).

13. Analysis Using Phonetic Segment Distribution: Harris (1951) claimed that it is possible to discover phonemes of a language purely by examining the distribution of phonetic segments.

14. Twaddell's Statement on Mentalistic Approaches: Referring to the mentalistic definitions of a phoneme, Twaddell (1935) states: "Such a definition is invalid because a) we have no right to guess about the linguistic workings of an inaccessible 'mind', and b) we can secure no advantage from such guesses. The linguistic processes of the 'mind' are as such quite simply unobservable, and introspection about linguistic processes is notoriously a fire in the wooden stove".

15. Value Attributed to Native Speaker's Intuition: Using English [n_j] as an example, Sapir (1925) argued that, despite the superficial appearance that this belongs to a group of nasal consonants, "no native English speaking person can be made to feel in his bones that it belongs to a single series with /m/ and /n/ ... It still feels like n,g".

16. Emergence of Mentalist over Structuralist: The theory of generative phonology which emerged in the 1960's explicitly rejected the Structuralist approach to phonology and favored the mentalistic or cognitive view of Sapir (Chomsky (1964), Chomsky and Halle (1968)).

Correspondence Between Letters and Phonemes

1. Equivalence of Phonemes to Graphemes: Phonemes are considered to be the basis for alphabetic writing systems. In such systems, the written symbols - graphemes - represent, in principle, the phonemes of the language being written.

2. Alphabet System for Classical Latin: This is most obviously the case when the alphabet was invented with a particular language in mind; for example, the Latin alphabet was designed for Classical Latin, and therefore the Latin of that period enjoyed a near one-to-one

correspondence between phonemes and graphemes in most cases, though the devisers of the alphabet chose not to represent the phonemic effect of vowel length.

3. <u>Established Orthography vs. Evolving Phonemes</u>: However, because changes in spoken language are not often accompanied by changes in established orthography - as well as other reasons, including dialect differences, the effects of morphophonology on orthography, and the use of foreign spellings for some loanwords - the correspondence between spelling and pronunciation in a given language may be highly distorted; this is the case with English, for example.

4. <u>Correspondence between Phonemes and their Symbols</u>: The correspondence between symbols and phonemes in alphabetic writing systems is not necessarily a one-to-one correspondence.

5. <u>Phonemes Constructed using Letter Combinations</u>: A phoneme might be represented by a combination of 2 or more letters - digraph, trigraph, etc., like <sh> in English or <sch> in German, both representing phoneme /integral/.

6. <u>Single Symbol Representing Multiple Phonemes</u>: Also, a single symbol may represent two phonemes, as in English <x> representing |gz| or /ks/.

7. <u>Complications arising from Pronunciation Rules</u>: There may also exist spelling/pronunciation rules - such as those for the pronunciation of <c> in Italian - that further complicate the correspondence of letters to phonemes, although they need not affect the ability to predict the pronunciation from spelling and vice versa, provided the rules are known.

In Sign Languages

1. <u>Sign Language Articulation Feature Bundles</u>: Sign language phonemes are bundles of articulation features. Stokoe was the first scholar to describe the phonemic system of ASL.

2. Identifiers from Tab/Dez/Sig: He identifies the bundles *tab* - elements of location, from Latin *tabula*, *dez* - the handshape, from *designator*, *sig* - the motion, from *signation*. Some researchers also discern *ori* - orientation, facial expression, or mouthing.

3. Sign Phonemes and Minimal Pairs: Just as with spoken languages, when features are combined, they create phonemes. As in spoken languages, sign languages have minimal pairs which differ in only one phoneme.

4. Examples - Father/Mother ASL Signs: For instance, the ASL signs for *father* - https://media.spreadthesign.com/video/mp4/13/455635.mp4 - and *mother* - https://media.spreadthesign.com/vides/mp4/13/48601.mp4 - differ minimally with respect to location while handshape and movement are identical; location is thus contrastive.

5. Limitations of Stokoe's terminology/Notation: Stokoe's terminology and notation system are no longer used by researchers to describe the phonemes of sign language; Stokoe's research, while still considered seminal, has been found to not characterize American Sign Language of other sign languages sufficiently (Clayton and Lucas (2000)).

6. Enhancement to Sign Language Phonology: For instance, non-manual features are not included in Stokoe's classification. More sophisticated models of sign language phonology have since been proposed by Sandler (1989), Brentari (1998), and van der Kooij (2002).

Chereme

1. The Basic Unit of Sign: *Cherology* and *chereme* - from Ancient Greek *chi epsilon i rho* "hand" - are synonyms of phonology and phoneme previously used in sign languages. A *chereme*, as the basic unit of signed communication, is functionally and psychologically equivalent to the phonemes of oral languages, and has been replaced by that term in the academic literature.

2. <u>Cherology - Study of Sign Cheremes</u>: *Cherology*, and the study of *cheremes* in language, is thus equivalent to phonology. The terms, are, not in use anymore. Instead, the terms *phonology* and *phoneme* - or *distinctive feature* - are used to stress the linguistic similarities between signed and spoken languages (Bross (2015)).
3. <u>Acceptance of the above Terminology</u>: The terms were coined by Stokoe (1960) at Gallaudet University to describe sign languages as true and full languages. Once a controversial idea, the position is now universally accepted in linguistics. Stokoe's terminology, however, has been largely abandoned.

References

- Bearth, T., and C. Link (1980): The Tone Puzzle of Wobe *Studies in African Linguistics* **11 (2)** 147-207
- Bloomfield, L. (1933): *Language* **Henry Holt**
- Brentari, D. (1998): *A Prosodic Model of Sign Language Phonology* **MIT Press**
- Bross, F. (2015): Chereme, in: *Dictionaries of Linguistics and Communication Science, T. A. Hall and B. Pompino-Marschall editors* **Mouton de Gruyter**
- Chao, Y. R. (1934): The non-uniqueness of Phonemic Solutions of Phonetic Systems *Academia Sinica* **4 (4)** 363-397
- Chomsky, N. (1964): *Current Issues in Linguistic Theory* **Mouton**
- Chomsky, N. and M. Halle (1968): *The Sound Pattern of English* **Harper and Row**
- Clark, J., and C. Yallop (1995): *An Introduction to Phonetics and Phonology, 2nd Edition* **Blackwell**

- Clayton, V., and C. Lucas (2000): *Linguistics of the American Sign Language - An Introduction 3rd Edition* **Gallaudet University Press** Washington D.C.
- Crystal, D. (2010): *The Cambridge Encyclopedia of Language 3rd Edition* **Cambridge** UK
- Dinnsen, D. (1985): A Re-examination of Phonological Neutralization *Journal of Linguistics* **21 (2)** 265-279
- Harris, Z. (1951): *Methods in Structural Linguistics* **Chicago University Press**
- Householder, F. W. (1952): Review of Methds in Structural Linguistics by Zellig W. Harris *International Journal of American Linguistics* **18** 260-268
- Jakobson, R, G. Fant, and M. Halle (1952): *Preliminaries to Speech Analysis* **MIT**
- Jakobson, R., and M. Halle (1968): Phonology in Relation to Phonetics, in *Manual of Phonetics, editor: B. Malmberg* **North-Holland**
- Jones, D. (1919): The Phonetic Structure of the Sechuana Language *Transactions of the Philological Society* **28 (1)** 99-106
- Jones, D. (1957): The History and the Meaning of the Term Phoneme *Le Maitre Phonetique* **35 (72)** 1-20
- Kiparsky, P. (1968): Linguistic Universals and Linguistic Change, in: *Universals in Linguistic Theory (E. Bach and R. T. Harms - editors)* **Holt, Rinehart, and Winston** New York
- Ladefoged, P. (2006): *A Course in Phonetics 5th Edition* **Thomson**
- Liddell, H. G., and R. Scott (1940): *A Greek-English Lexicon* **Clarendon Press** Oxford
- Moran, S., D. McCloy, and R. Wright (2014): *PHOIBLE Online* **Max Planck Institute for Revolutionary Anthropology** Leipzig, Germany
- Pike, K. L. (1947): *Phonemics* **University of Michigan**
- Pike, K. L. (1967): *Language in Relation to a Unified Theory of Human Behavior* **Mouton**

- Sandler, W. (1989): *Phonological Representation of the Sign: Linearity and Nonlinearity in American Sign Language* **Fortis**
- Sapir, E. (1925): Sound Patterns in Language *Language* **1 (37)** 37-51
- Singler, J. V. (1984): On the Underlying Representation of Contour Tones in Wobe *Studies in African Linguistics* **15 (1)** 59-75
- Stokoe, W. C. (1960): *Sign Language Structure: An Outline of the Visual Communication Systems of the American Deaf* **University of Buffalo**
- Trager, G. and H. Smith (1951): *An Outline of English Structure* **American Council of Learned Societies**
- Twaddell, R. F. (1935): *On Defining the Phoneme* **Linguistic Society of America**
- van der Kooij, E. (2002): *Phonological Categories in the Sign Language of the Netherlands - the Role of Phonetic Implementation and Iconicity* **Leiden University**
- Wells, J. (1982): *Accents of English* **Cambridge University Press**
- Wikipedia (2020): Phoneme (https://en.wikipedia.org/wiki/Phoneme)
- Wikipedia (2021): English Phonology (https://en.wikipedia.org/wiki/English_phonology)