**Concept Map**

Speech and Language Processing (3rd ed. draft)
Dan Jurafsky and James H. Martin

**Chapter 25**
**Phonetics**

**CSE 431**

**Task 3**
**Group 07 (Section 2)**

22241167 Fahmida Ahmed
19101491 Md. Muhtadee Faiaz Khan Soumik
19101166 Syed Ziaul Bin Bashar

**Chapter 25: Phonetics**

Speech recognition and text-to-speech algorithms are both based on the implicit notion that spoken words are made up of smaller speech units. This chapter provides a comphonetics putational viewpoint on phonetics, which is the study of speech sounds used in many languages around the world and how they are formed in the human vocal tract, realized acoustically, and processed digitally.

**25.1 Speech Sounds and Phonetics Transcription**

The International Phonetic Alphabet (IPA), an evolving standard first devised in 1888, is the preferred phonetic representation for transcribing the world's languages. The ARPAbet, a straightforward phonetic alphabet that readily uses ASCII symbols to represent an American-English subset of the IPA, will be used to represent phones in this chapter. Numerous IPA and ARPAbet symbols correspond to well-known Roman letters.

**25.2 Articulatory Phonetics**

This section provides the study of how these sounds are made as the mouth, throat, and nasal organs alter the airflow from the lungs is known as articulatory phonetics.

**25.3 Prosody**

Prosody is the study of the intonational and rhythmic aspects of language, and in particular the use of F0, energy, and duration to convey pragmatic, affective or conversation-interactional meanings. Prosody is used to mark the saliency of a particular word or phrase.

**25.4 Acoustic Phonetics and Signals**

The acoustic waveform, its digitalization, and its frequency analysis are quickly introduced at the outset.

**25.5 Phonetic Resources**

The open-source CMU Pronouncing Dictionary has pronunciations for word forms, while the fine-grained UNISYN dictionary freely available for research purposes, gives syllabifications, stress, and also pronunciations for dozens of dialects of English.

**25.2.1 The Vocal Organs**

The air travels through the larynx, also referred to as the voice box or Adam's apple, as it moves through the trachea. The alphabet and all English vowels are among the voiced and unvoiced sounds. Consonants can be voiced or unvoiced and are produced by restricting or obstructing the airflow in some way.

**25.2.2 Consonants: Place of Articulation**

Consonants are made by restricting airflow which can break down into 6 classes, with prominent places of articulation like labial, dental, alveolar, palatal, velar, and glottal.

**25.2.3 Consonants: Manners of Articulation**

Airflow is restricted but not entirely stopped in fricatives. Air can move between the teeth and around the tongue thanks to the dental fricatives. The tongue is positioned behind the alveolar ridge in the palato-alveolar fricatives and, driving air through a groove the tongue has made.

**25.2.4 Vowels**

Similar to consonants, vowels can be distinguished by the location of the articulators during production. For instance, the tongue's highest point is located near the front of the mouth during a vowel. The tongue is higher than other organs. A diphthong is a vowel in which the tongue position changes significantly as the vowel is being produced.

**25.2.5 Syllables**

A syllable is a vowel-like sound together with some of the surrounding consonants that are most closely associated with it. The vowel at the core of a syllable is called the nucleus. Initial consonants are called the onset. The coda is the optional consonant of consonants following the nucleus.

**25.3.1 Prosodic Prominence: Accent, Stress and Schwa**

Certain words and syllables are audibly more prominent than others in an American English sentence when spoken in a natural way. By prominence, these syllables or words are perceptually more noticeable to the listener. Syllables with a pitch accent are referred to as accented syllables.

**25.3.2 Prosodic Structure**

Prosodic structure is often described in terms of prosodic phrasing, meaning that an utterance has a prosodic phrase structure in a similar way to it having a syntactic phrase structure. Automatically predicting prosodic boundaries can be important for tasks like TTS.

**25.3.3 Tune**

Two utterances with the same prominence and phrasing patterns can still differ prosodically by having different tunes. Other examples include the characteristic English contours for expressing contradiction and expressing surprise.

**25.4.1 Waves**

The sine and cosine functions that depict a sine wave are the foundation of acoustic analysis. The frequency is the number of cycle per seconds, that a wave repeats itself which also called Hertz. A sine wave's amplitude A is its greatest value on the Y axis in graph.

**25.4.2 Speech Sound Waves**

Using a time-series graph of the air pressure change, we can visualize sound waves. The difference between two integers serves as a minimum granularity and any values that are closer together than this quantum size are represented identically, this representation of real-valued numbers as integers is known as quantization.

**25.4.3 Frequency and Amplitude; Pitch and Loudness**

Sound waves characterized in terms of frequency, amplitude, and the additional factors for pure sine waves. We anticipate regular peaks in the nature of each major peak, corresponding to an opening of the vocal folds, while the vocal folds are vibrating.

**25.4.4 Interpretation of Phones from a Waveform**

An examination of a waveform visually can provide a lot of information. Remember that vowels have voices, and that they frequently have length and volume. The voicing process is accomplished by periodic peaks in the amplitude of each main peak that correspond to an opening of the vocal folds.

**25.4.5 Spectra and the Frequency Domain**

Each of a signal's frequency components and their amplitudes are shown in the signal's spectrum. We use the spectrum as a tool to investigate the constituent frequencies of a sound wave at a certain time point. The spectrum is an alternate representation of the original waveform.

**25.4.6 The Source-Filter Model**

The source filter model simulates how the vocal tract shapes the pulses produced by the glottis to explain the acoustics of a sound. As a result, distinct harmonics will be increased when a wave with the same fundamental frequency passes through different vocal tract places.