Task - 03[Team Report on a Unique Book Chapter]
Group - 27

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

Chapter-28
Phonetics

Members:
20101440 Name Abhijit Saha
20101554 Name Faiza Bushra
20301077 Name Ahmed Anwar
20301267 Name Rubaba Rashid

Chapter-28
Phonetics

The use of symbols representing sounds that make up words. The concept of smaller speech units underlying the spoken word is important to both speech recognition and text-to-speech algorithms. Computational perspective on phonetics includes the study of speech sounds, their production in the human vocal tract, and how they can be digitized and processed.

28.1 Speech Sounds and Phonetic Transcription

The pronunciation of words is being represented as a string of phones, which are speech sounds, each represented with symbols adapted from the Roman alphabet. The Standard phonetic representation for transcribing the languages of the world is the International Phonetic Alphabet (IPA), a developing standard initially developed in 1888. In this chapter, phones will be represented using the ARPAbet, a simple phonetic alphabet that easily uses ASCII symbols to express an American-English subset of the IPA. There are numerous IPA and ARPAbet symbols that are equivalent to popular Roman letters.

28.2 Articulatory Phonetics

This section basically describes the articulatory phonetics, which is the study of how these sounds are produced through mouth, throat, and nasal organs alter the airflow from the lungs.

28.3 Prosody

Prosody refers to the melodic and rhythmic aspects of speech, including pitch, loudness, and duration. These features help to convey important information, such as emphasis, emotion, and attitude, and can make speech more engaging and memorable for listeners. Prosody plays a crucial role in communication, helping to signal the structure and meaning of spoken language. It is also important for conveying social and cultural information, such as accents, dialects, and regional variations. Understanding prosody is essential for effective communication, both in everyday conversation and in professional settings such as public speaking and broadcasting.

28.4 Acoustic Phonetics and Signals

We begin with a brief introduction to the acoustic waveform, its digitalization, and its frequency analysis.

28.5 Phonetic Resources

Phonetically annotated corpora like TIMIT, Switchboard, and Buckeye provides time-aligned transcriptions whereas the fine-grained UNISYN dictionary that is available for research purposes can provide syllabification, stress information and pronunciations for dozens of English dialects. The LDC pronunciation dictionary provides pronunciation lexicons for Dutch, English, German, Japanese, Korean, Mandarin, Spanish and Egyptian Arabic

The Vocal Organs

As air moves through the trachea, it passes through the larynx, also known as the voice box or Adam's apple. Among the voiced and unvoiced are the alphabet and all English vowel sounds. Consonants are produced by limiting or obstructing the airflow in some way.

Consonants: Place of Articulation

Consonants are produced by restricting airflow, which can be classified into six classes based on prominent points of articulation

Consonants: Manner of Articulation

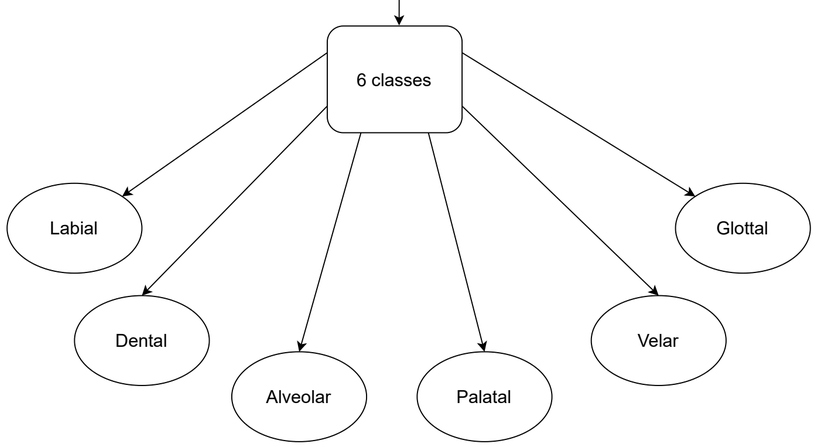
Airflow is partially restricted and Dental fricatives allow air to move between the teeth as well as around the tongue. In the palato-alveolar fricatives, the tongue is positioned behind the alveolar ridge, driving air through a groove the tongue has made.

Vowels

Vowels, like consonants, can be distinguished by the placement of the articulators throughout production. During a vowel, the tongue's highest point is near the front of the mouth. The tongue is placed higher than the other organs.
.

Syllables

A syllable is a combination of a vowel-like sound and some of the consonants that are mostly associated with it. The vowel which is the core of a syllable is called nucleus. The initial consonants are referred to as the onset. The coda is the optional consonant that comes after the nucleus.
.



Prosodic Prominence: Accent, Stress and Schwa

When American English is spoken naturally, some words and syllables are audibly louder than others in a sentence. These syllables or words are more perceptually visible to the listener because of their prominence. Accent syllables are those that have a pitch accent in them.In addition to pitch accent, other factors can also contribute to the perceptual prominence of a syllable or word in spoken English. For example, the duration of a syllable or word can also make it more prominent, as longer syllables or words tend to draw more attention to themselves.

Prosodic Structure

Prosodic structure refers to the way in which spoken language is organized into larger units, such as phrases, clauses, and sentences, based on patterns of pitch, loudness, and duration. These prosodic features help to create a sense of rhythm and melody in speech and can convey important information, such as emphasis and attitude. Prosodic structure is a key element of spoken communication, playing a crucial role in conveying meaning and maintaining listener engagement.

Tune

In prosody, tune refers to the pitch contour or melody of a sentence or utterance. It involves the rising and falling of pitch across words and phrases and can convey important information such as the speaker's attitude and emotion, as well as the syntactic structure of the sentence.By using distinct tunes, two utterances with the same emphasis and phrasing patterns can yet differ prosodically. The typical English contours for expressing surprise and conflict are more examples.

Waves

Acoustic analysis is built on the sine and cosine functions that represent a sine wave. The frequency, also known as heart rate, is the number of cycles per second that a wave repeats itself. On the graph's Y axis, a sine wave's amplitude A is represented by its highest value.

Speech Sound Waves

We can see sound waves by looking at a time-series graph of the air pressure variation. The incorrect representation of real-valued numbers as integers is referred to as quantization. The difference between two integers serves as a minimum granularity, and all values that are closer together than this quantum size are represented identically.

Frequency and Amplitude:Pitch and Loudness

Pure sine wave sound waves are further classified according to their frequency, amplitude, and additional characteristics. As the vocal folds are vibrating, we expect regular peaks in the nature of each main peak that correlate to an opening.

Interpretation of Phones from a Waveform

Visually examining a waveform can reveal a lot of information. Keep in mind that vowels often have length and volume in addition to having voices. Periodic increases in each main peak's amplitude, which coincide with vocal fold openings, enable the voicing process.
.

Spectra and the Frequency Domain

The spectrum of a signal displays each frequency component and its amplitude. The spectrum is an alternate representation of the original waveform that we use to explore the individual frequencies of a sound wave at a specific time.
.

The Source-Filter Model

In order to describe the acoustics of a sound, the source filter model replicates how the vocal tract shapes the pulses generated by the glottis. As a result, when a wave with the same fundamental frequency travels through several vocal tract locations, individual harmonics will be amplified..