Concept Map

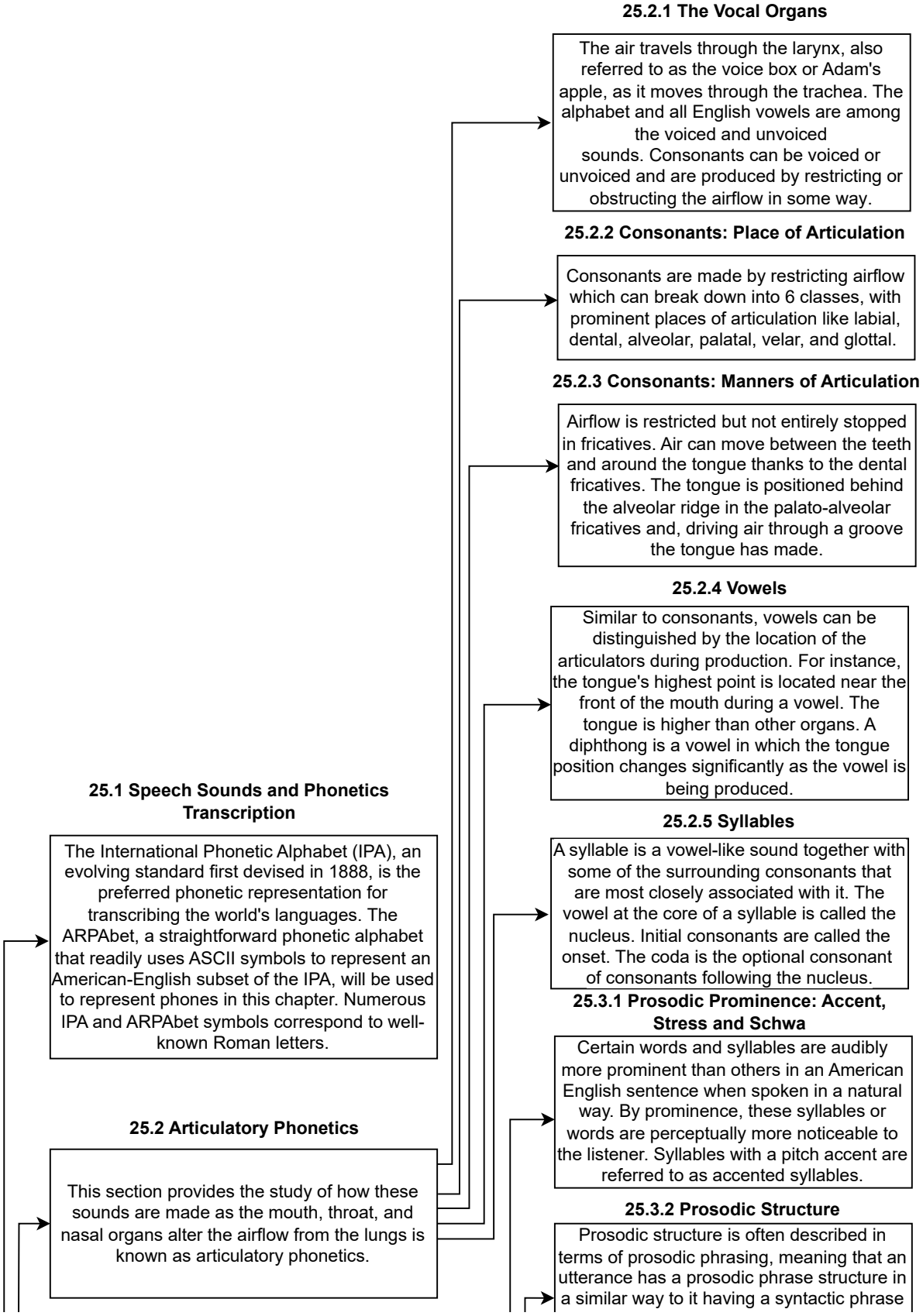
Speech and Language Processing (3rd ed. draft)
Dan Jurafsky and James H. Martin

Chapter 28
Phonetics

CSE 424

Task 3
Group 11

20101225 - Mahin Shahriar Efaz
20101238 - Aditi Saha Ria
23141036 - Md. Anonto Shuvo



structure. Automatically predicting prosodic boundaries can be important for tasks like TTS.

Chapter 28: 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 computational 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.

25.3 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.

25.3.3 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 Acoustic Phonetics and Signals

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

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.

25.5 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.

25.4.3 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.