CSME 206A Natural language & Speech Processing **Spring Semester**

Lecture 5: Computational Phonology

Lecturer: K.R. Chowdhary *: Professor of CS*

5.1 Computational Phonology

Symbols are used to represent sounds that make words. How the words are spelled in different languages indicate the relation between different cultures. The sound based writing, i.e, spoken words are composed of smaller units of speech, underlines the modern theory of *phonology*.

The *Computational Phonology* comprises: algorithms for speech recognition, speech synthesis, and the linguistics.

The aim of speech recognition is to take sound wave as input and produce as output a string of words. Where as text-to-speech (TTS) synthesis is to take input the words and produce output the acoustic sound waveform. The applications of these are: speaking dictionaries, speech based interface with computers and voice based input for disabled.

We will study how the words are pronounced in terms of individual units of speech called "phones". The challenge is that different phones are pronounced in different ways in different environments. The computational phonology is study of computational means for modeling phonological rules. We will be making use of FSTs (Finite State Transducers), studied earlier, for computational phonology.

5.1.1 Speech Sound and its Phonetic Transcription

The study of pronunciation is called *phonetics*. We will represent pronunciation of words as strings of symbols, which represent phones or segments. These symbols are phonetic symbols. For example, phone p corresponds to letter p and phone l corresponds to letter l. The IPA (International Phonetic Alphabets) is used as standard for phonetics. Some examples of IPA are given in the table 5.1.

These phones are generated by sound organs of human mouth when we speak (see the notes on first day).

5.1.2 Relation between Phonological Rules and Phonemes

All the identical phones, for example, [t] are not created equal. Accordingly, the phones are pronounced differently in different contexts. For example, [t] in tunafish and [t] in starefish, are pronounced differently. In the first, it is *aspirated* and in second it is not. The aspirated is period of voicelessness after the [t] closure.

There are other contextual variants of [t]. For example, when [t] occurs between two vowels, particularly when the first is stressed, it is pronounced as tap. In tap, the tongue is curled and pressed against the alveolar ridge.

Table 5.1: Examples of IPA.

IPA Symbol	Word	IPA Transcript
Consonants		
1. [p]	Parsely	[parsli]
2. [k]	Catnip	[kaetnie]
3. [d]	dill	[dil]
4. [θ]	thistle	[θisl]
5. [v]	clove	[kloʊv]
6. [z]	hazlnet	[heizln∧t]
Vowels		
1. [i]	lily	[lili]
2. [ei]	daisy	[deizi]
3. [ʊ]	woodruff	[wʊdr∧f]
4. [u]	tulip	[tulip]
5. [aʊ]	sunflower	[s∧nflaʊr]
6. [oʊ]	lotus	[loʊtes]

Another variant of [t] occurs before the dental consonant [θ]. Here [t] becomes [t̪], where the tongue touches the back of the teeth.

So the challenge is to represent the relation between [t] and its different variants, in different contexts. We do it by the abstract class, called, *phoneme*, which is realized as different *allophones* in different contexts. We write the phonemes inside the slashes. So /t/ is phoneme whose allophones includes [t^h], [ɾ], and [t̪].

The relationship between phonemes and its allophones is specified as phonological rules, e.g., the Chomsky and Hall rules indicated:

$$/t/ \rightarrow [\underline{t}]/ - -\theta \tag{5.1}$$

The symbols /t/, [t̪] are allophone and phonemes, respectively. The surface allophone appears to the right of arrow, and phonetic environment is indicated by symbols surrounded by the - -.

The following is version of flapping rule.

$$/\left\{ \begin{array}{c} t \\ d \end{array} \right\}/ \rightarrow [ɾ]/V^{'} - - - V \tag{5.2}$$

The symbols $V^{'}$ and V stand for stressed and unstressed vowels.

5.2 Phonological Rules and Transducers

The phonological rules can be implemented by transducers as we did in the case of morphological analysis. We will use two level morphology. As a first example, we use the *flapping* rule:

$$/t/ \rightarrow [ɾ]/V^{'} - -V \tag{5.3}$$

The corresponding finite state transducer is shown in figure 5.1.

In phonetics, a *flap* or *tap* is a type of consonantal sound, which is produced with a single contraction of the muscles so that one articulator (such as the tongue) is thrown against another.

The main difference between a flap and a stop is that in a flap, there is no buildup of air pressure behind the place of articulation, and consequently no release burst. Otherwise a flap is similar to a brief stop.

Flaps also contrast with trills, where the air-stream causes the articulator to vibrate. Trills may be realized as a single contact, like a flap, but are variable, whereas a flap is limited to a single contact.

In the figure 5.1, dx (in ARPAbet) indicates a flap, "other" means any feasible pair not used any where, @ indicates any symbol not used on any edge. In IPA dx is represented as [ɾ]. The FST accepts any string in which flaps in the correct places (i.e. after a stressed vowel, and before an unstressed vowel), and rejects all other strings. The examples of flapping sound words are: butter, mutter, and ruttr, where u is stressed and e is unstressed vowel. The tt is flapping sound. However, "letter" is not a flapping sound, though tt is surrounded by vowels.

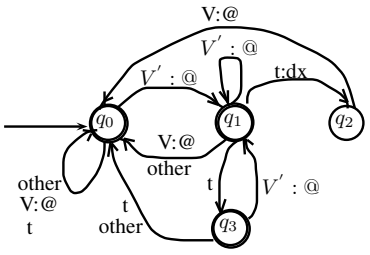

Figure 5.1: English Flapping Transducer.

A Flap is a consonant, the flap sound occurs, e.g., when number of birds flock take-off for fly, and the sound produced by feathers together for all these is flapping sound.

References

[1] D. Jurafsky and J. Martin, "Speech and Language Processing," *Pearson India*, 2002, Chapter 4.