Language Typology and Universals - Lec 3

All the theree fields studied till now separately, i.e lingo, C.S. and Maths all contribute to one field - A.I.

- Linguistics courses are imp.
- Typology stands to look at the differences between the languages

Examples demonstration to show the correlation between languages

- English - Ram killed Ravan (SVO)
- French - Ram a tue` Ravan (SVO)
- German - Ram hat Ravan getotet (SVO)
- Hindi - Ram ne Ravan ko maara (SOV)

Indo-Aryan and Dravidian languages follow a SOV order

- Suppose you are building a model for a SVO language, it can be extended to other similar SVO languages as well
- Similarly is the case for SOV
- This realisation could not be made until very late
- Once a model for a language was made, the same schema could be implemented for other languages
- IL-ILMT - a group of institutes working on developing language models with HIndi as the intermediate language
- Many old models were Rule-Based and Statistal in nature.
- The concept was to convert an English (or any other foreign language)

For Verb, we have TAM (Tense, Aspect, Modality)

For Noun, we have GNP (Gender, Number, Person)

Not to forget, Cases. The problem for a NLP student is when you encounter a null marker.

For Hindi, we don't have a fixed notation of gender for few words (eg. daal) and can vary across dialects.
In telugu, objects get assigned a female gender by default.
In Marathi, the verb NECESSARILY has a gender assigned while in hindi might be carried by the auxillary verb as well
Bengali DOES NOT have gender at all
Same for Malayalam
Nepali is a normal language (very nice)
Tamil has singl. and pl. (different genders have different terms)

Table

Singl.		Plu
Masc.	Fem.	

```
Human     Non-living | Human     Non-Human
du-marker  Cow       |
(Telugu)   Girl      |
           table     |
```

In telugu, WHO has separate terms for both singl. and pl. forms.

A chat on Dakkhini

Dakkhini - Has a structure similar to Hindi but on a closer inspection, it also bears much resemblance with the Dravidian languages. (hona = want)
Telugu 'ani' is a complimentiser. eg. Sita told me THAT John will come tomorrow
Bengali - je, Marathi/Nepali - ki, Tamil - enru, Malayalam - enn, Telugu - ani
In telugu, the sentence would go something like 'john will come tomorrow, sita said'
Dakkhini has a similar structure. Bolke is one such complimentiser. Many words are translated directly from telugu and other Dravidian languages to Dakkhini.

We say a lot of Indian Languages on the superficial fact that most of the IL have SOV order.

Idiolect

- A very personal style/way of speaking (even writing for that matter)
- Way of speaking restricted to an age group
- Gender
- Colors - It is said that girls are able to distinguish colors much more ... vividly?

Sociolect

- Regional specific and region-restricted
- Eg. Indians who went foreign while returning have an accent

In Historical Linguistics, we are interested in how a language changes over time.
Language change does not come about in a day. It is a continous process.

There are two studies for languages:

1. Synchronic Study - How was a language spoken at any X period in time?
2. Diachronic Study - How has a language changed from point X to point Y in time?

Changes might happen at the sound level or morphology or Structural level. DO NOT forget that there might be changes at the spellings level as well (thou art in heaven)

IMP: Philology is a sub-part of Historical Linguistics

Eg. in Telugu Bus is written as Bas but for the plural instead of buses it is written as bassulu the 'ss' is geminated before adding the 'ulu' part.
Similarly in Hindi buses are written as basein

As far as syntax changes are much lower.

Differences can be looked into:

- Within different languages from different language families
- Different languages from the same language family

- Within the same languages over a spread of time

Unviersals

This is about the similarities

To ask ma'am - How do native speakers avoid confusion that might arise because of a smaller phonetic articulations. eg. Zanjeer if spoken as janjeer can be confused with Janjir fort.