

Observing and Being part of Evolution of a Language

Some Random Thoughts

- Rohit Kumar -
(rohit@iiit.net)

The way the computer users interact today i.e. by emails and chats and all the variants of these that are available, they are in the process of evolving a whole new natural language. A language that has (like any other natural language) its own nuances and details and probably also a grammar and a lexicon. Or probably the language has already evolved and is currently undergoing the dialectization phase that happens when a new language comes into contact with other languages and is used by natives of other languages.

If we think about it we are surprised to observe that we are having the honour to see and be part of the evolution of a whole new language and how is it that a language influences the society and a lot more stuff that is studied by linguists and language scientists.

So here are some observations and thoughts that strike me when I realize that we are being part of this evolution. One of the things that is typical to this new cyber language is that nobody is a native user of this language (atleast not till the matrix or its like is created☺). Nobody is taught this language as a part of their basic learning like your native language and other natural languages that you learn formally or informally. All users of this new language are natives of some other language also. Their native language effects and modifies this new language in the way it is used by a particular individual or a group. This fact is very basic to this language. As it has evolved over a common electronic medium but has been used by a large number of people with different language background this language has a level of

flexibility that is AWESOME compared to any other natural language.

This flexibility is observed when we see words from other languages getting adopted into this language along with special modifiers that are special to the new language. Not that this phenomena is very new to this language. For e.g.: the English word Refinery when used as a plural in a hindi text becomes "refinariyan". And we call it "Hindification" which is another example of the same. But the kind of modifiers that are being used in the new language are typical to it. One most observable one is the abbreviating modifier. asap, brb, cya, lol, fyi, and what not. Also it may be noted that the new language is evolving its own lexicon that has been influenced by so many social, technical and historical reasons. Many new words are getting coined rapidly for this new language and as we look for them, new dictionary are being built. People call it jargon and consider it as an addendum to the existing languages. But is it so? This is surely worth some brain-storming sessions and I am not going to discuss it out here.

This new language will evolve to be a very flexible one and the dynamics of minimization of effort and word distributions (suggested by Zipf's Law) will get to applicable on this too. The effort would probably get to be measured by the work done in moving from one key to other and how to humans try to minimize the work while still trying to maintain communication. Alternative studies may come up as to how the users are skipping characters or how are typos coming into play.

Also we note that basically this language is written and not spoken (atleast that is what it seems now. and that is probably unlike a lot other of languages). This is the second most typical thing to this language. Try imagining saying "brb" to someone on phone and "lol" in your gang and "asl" to first time you see that "pyt" sitting next to you in that lecture. Its odd. So we observe that this language is not for spoken use (atleast for now). This leads to another observation that speech is the major domain where other languages undergo variations due to influence of natives of other languages. But despite lacking the speech dimension, this language is having so much of variation which is another statement of its inherent flexibility. But spoken usage of the nuances of this language will evolve sooner than later.

One of dimension I would note here about the flexibility of this language is the script and the spellings. Universally due to wide availability the characters of the English keyboard is the official script of the language along with a basic set of emoticons. (though to some extent other scripts are also used). Now the observation here is the way the languages other than English and to some English also is scripted. There is a great deal of variation in the spellings and several possible spellings of words are readily accepted. All this could be accounted for due to the awesome pattern matching capabilities of human brain. And also to the fact that finding a nearest best match pattern is of much lesser effort than typing the whole word correctly everytime. And hence by the general the users minimize total effort. As machines lack this extensive pattern matching capability, the processing of this language is more and more difficult or if I put it the other way challenging. There are other reasons for interest in being able to process this language, one being the ability to increase usability of dialog

systems which are envisioned to be the next generation user interfaces by many. Another typicality of this language is that the study and use of this language will not be restricted to users of a particular region like most of the languages.

Various studies will start evolving an almost new sub field of language research would come up. This field will be of interest not only to the linguists but also to the N. L. P. people and computer scientists working with language. There will be lot of interest both academically and commercially to generate processor for this new language primarily due to the large amount of corpuses that are being generated in this languages. Processing the languages and generating relevant knowledge will be surely of interest. Lot of things can be thought of and weeks of brain-storming can go into it. That is not what I intend to do here.

It is arguable though. Can we call it a language at all? Or is it that it is dialect of existing languages. But then it has to be many dialects of so many languages. And the dialects within themselves will have a lot of commonalities. So probably it is correct to call it a new language which has so much of flexibility due to circumstances of its evolution that it gets influenced and influences the other languages it get in contact with through users.

Finally I note that compared to any other language which took probably hundreds and thousands of years to develop this new language has surely been a pace-ster partially due to its derivation from existing languages.

A lot of possibility to explore knowledge wealth will be opened and we will witness and will be part of it right in our own short lives. We are privileged in that way.