NLP towards Revival and Development of Tamil

S.Sivakama Sundari

Head, Department of Computer Science, Alpha Arts and Science College, Porur, Chennai, India
siva.aasc @gmail.com

Abstract— Tamil is the most traditional and fundamental of all the languages of the world. The letters started from Stone Age era and now has reached up to internet era. The three fundamental aspects of strength are unbeaten tradition, continuity in existence and development to cope up with the current scenario. Many great achievers own Tamil as their mother tongue. Tamil adds glory to the society by virtue of its culture and societal discipline. The sons of this language are motivated to take necessary steps to revive and develop their mother tongue in the international arena.

Keywords— Natural Language Processing; Virtual University; Incubation centre; Inscriptions;

1. Introduction

Natural language Processing is a discipline of computer technology supported by artificial neural network algorithms. Mankind is bent upon earning money for survival. But every human being is inclined to do certain activities for their passion. One such activity is the promotion and development of their mother tongue; now the language in discussion is Tamil.

Tamil Virtual University has created an impetus in the development of the language. Nowadays the software engineers spend their week -ends to work in the agricultural field to develop traditional agriculture. Similarly there is a clear indication of the vigour of language development in the present day population. There may be two reasons. One is that they realize the greatness of Tamil. Other reason is that the latest innovations in computer technology in hardware and software have made the impossible to become possible. The cheapest memory, internet connectivity, cloud computing, open source software and high speed processors has made almost all the people to revive their mother tongue. Starting from the age of inscriptions in the stone and dry leaves the life and existence of Tamil language is amazing. Every linguist and every scientist must contribute their innovation and effort towards this noble cause making the language untouched and unblemished by the passage of time.

2. Survival of the language

The time of actual creation of Tamil language is not known. A lot of controversial ideas exist. But it is the most fundamental and oldest of all the languages in the world, which is accepted by people all over the world.

2.1 Electronic media

The electronic media is used to convert the inscriptions on the rocks and dry leaves (chuvadi) and books written many thousands of years ago are revived and stored permanently to excel the barrier of time and other forms of danger.

2.2 Internet

Internet is another technology which aids in the preservation, search and revival of the time tested books and literary works in Tamil. Cloud computing is the technique to make use of the memory from a remote distance.

2.3 CD/DVD

CDs and DVDs are also used to transfer the scholarly information from place to place in order to cater to the needs of the people across the globe.

3. Social Entrepreneurship

Computer technology has gone a long in assisting the development of the language.

3.1 Machine Translation

Machine translation is the most popular venture taken up by almost all the software companies and search engines to facilitate the power of better communication.

3.2 Online Dictionaries

Online dictionaries are available as freeware, shareware as well as open source. The number of words in the corpus is taken as the measure of popularity and viability of a language. The status of Tamil must be improved by addition of new words to the vocabulary.

3.3 Online error checker

The spelling errors checking and grammar error checking and online virtual keyboard are the facilities due

to the advent and use of electronic media. Tamil typing has become an easy sport to all with such facility.

3.4 Online conversion from graphics to text

The articles in the form of graphics can be converted into text using Google app. This provides the ability to edit the text and use the same for the knowledge transfer or societal improvement or information gathering.

. 4. Growth of linguistic skills

The linguistic skills are promoted by the usage of electronic equipments and software. Traditional procedures were time consuming and difficult to learn. The easiness has attracted more people into linguistic activities. Natural language is a promising field of research for the computer technologists. NLP is a tech savy and societal field of study, because it touches the lives of millions. That's precisely the aim of any research. It results in the improvement of the life style of the people.
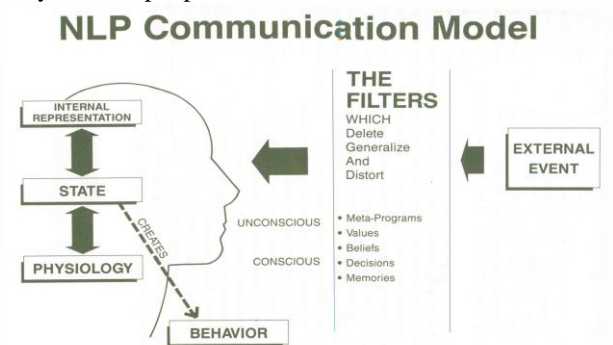

Fig.1: NLP Communication Model

5. Global village - open to all languages

The people of Japan, China and many countries of Europe like Germany and France are equipped with the online translation facility, which translates any text of other languages of the world into their mother tongue by means of equipment. This equipment senses the letters and translates and displays in the required language within a fraction of a second. Now this facility has rendered the entire world into a global village. Such advanced equipment is yet to be developed to support Tamil.

6. Personal skills honed

Anyone who studies his primary education in his mother tongue is capable of academic excellence. This fact has been realized by people all over the world and so they want to teach their children through their mother tongue, though they are geographically dispositional. This is made possible

with the advent of technology in the field of NLP. Learning their mother tongue not only gives pleasure to the child but also improves his/her skills.

7. Support of hardware

High speed processors and very cheap memory, cloud computing techniques are acting as a boon to the linguistic innovations and development.

8. Support of software

Much meaningful software is developed in order to assist the people to improve their communicative skills. Many software products provide the interface in different languages including the search engines. The next step is to develop the programme itself in Tamil or any other language of interest. Tamil Virtual University provides the incubation centre to the research scholars to develop innovative projects in to promote Tamil language.

9. Communication barrier smashed

People of the world live in different parts of the globe due to personal and official commitments. When you receive an short mail service (sms) in Tamil it makes you feel happy, when you are away from your motherland.

10. Use of electronics and wireless technology

All the wireless and electronic equipments are aiming at the human population as their target either for their sales or marketing or information explosion. These ventures can be successful only when the take the NLP as a tool with particular reference to the mother tongue bond.

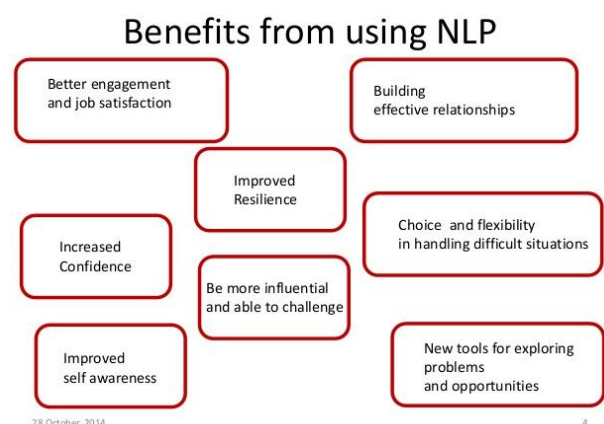### 11. Mobile ad-hoc network in aid of NLP

Fig. 2: Benefits from using NLP

Present day smart phones and mobile phones are supporting the cross language communication policies. In this time constrained life scenario, injecting the mother tongue through technology will create the realistic effects.

12. Conclusion

NLP assisted by the various services and projects offered by the Tamil Virtual University will definitely bring Tamil as the language of highest vocabulary and popularity amongst all the Indian languages.

References

[1] Alisa Kongthon, Chatchawal Sangkeettrakarn, Sarawoot Kongyoung and Choochart Haruechaiyasak. Implementing an online help desk system based on conversational agent Published by ACM 2009 Article, Bibliometrics Data Bibliometrics. Proceeding, MEDES '09 Proceedings of the International Conference on Management of Emergent Digital EcoSystems, ACM New York, NY, USA.

[2] Hutchins, J. (2005). "The history of machine translation in a nutshell" (PDF).[self-published source]

[3] Kishorjit, N., Vidya Raj RK., Nirmal Y., and Sivaji B. (2012) "Manipuri Morpheme Identification", Proceedings of the 3rd Workshop on South and Southeast Asian Natural Language Processing, pages 95–108, COLING 2012, Mumbai, December 2012.

[4] Yucong Duan, Christophe Cruz (2011), Formalizing Semantic of Natural Language through Conceptualization from Existence. International Journal of Innovation, Management and Technology(2011) 2 (1), pp. 37-42.

[5] Mittal et al., "Versatile question answering systems: seeing in synthesis", IJIIDS, 5(2), 119-142, 2011.

[6] Ciaramella, Alberto, Giancarlo Pirani, and Claudio Rullent. "Conclusions and Future Developments." Advanced Algorithms and Architectures for Speech Understanding. Springer Berlin Heidelberg, 1990. 265-274.

[7] Roberto Pieraccini, Esther Levin, Chin-Hui Lee: Stochastic representation of conceptual structure in the ATIS task,, Proc. Fourth Joint DARPA Speech and Natural Lang. Workshop, Pacific Grove, CA, Feb. 1991.

[8] Christopher D. Manning, Hinrich Schütze: Foundations of Statistical Natural Language Processing, MIT Press (1999), ISBN 978-0-262-13360-9, p. xxxi

[9] Magnini, B., Cappelli, A., Tamburini, F., Bosco, C., Mazzei, A., Lombardo, V., Bertagna, F., Toral, A., Bartalesi Lenzi, V., Sprugnoli, R. & Speranza, M. (2008, May). Evaluation of Natural Language Tools for Italian: EVALITA 2007. In Proceedings of LREC 2008.

[10] Attardi, G., Basile, V., Bosco, C., Caselli, T., Dell'Orletta, F., Montemagni, S., Patti, V., Simi, M. & Sprugnoli, R. (2015). State of the Art Language Technologies for Italian: The EVALITA 2014 Perspective. Intelligenza Artificiale, 9(1), 43-61.

[11] Cambria, Erik; White, Bebo (2014). "Jumping NLP Curves: A Review of Natural Language Processing Research" (PDF). IEEE Computational Intelligence Magazine 9 (2): 48–7.doi:10.1109/mci.2014.2307227.