

Predictive Analytical Study on Endangered Languages

SANKAR RAJEEV

KI (Deemed-to-be) University, Guntur, India

ABSTRACT

It is predictive that the biological function exercised by the voice box translates a noise into intelligible expressions (Tomasello 1996; Hauser, Chomsky & Fitch 2002; "Language" 1992). Representation of experience by words was explored at length by Gorgias and Plato in ancient Greece. Rousseau went on to illustrate that language stemmed from emotions, whereas Kant considered that language emerged from careful and logical thought. The structuralism view of language was first introduced by Ferdinand de Saussure (Trask 2007; Bett 2010; Devitt & Sterelny 1999; Hauser & Fitch 2003; Pinker 1994).

Predictive analytics methods give a rough judgement that of the available and in-use languages a large majority are likely to be dead by 2100. How does it happen? Is it due to cultural invasion, scientific adjustment or technological convenience?

It is not very difficult to prevent a language from disintegrating or becoming endangered. Desire of a community to conserve its language has numerous options available. Take for example Modern Hebrew which was revived after generations modified and adaptable to the present requirements.

Keywords: Intelligible expressions, language, endangered, Predictive analytics, technological convenience

1. INTRODUCTION

1. Development of language can possibly be explained as a consequence of enlargement of human brain. Scientifically it has been identified that Broca's and Wernicke's areas influence the processing of language. When human beings interact in a society the formative product is a means of communication and language emerge out of that. Social stratification, social grooming and entertainment modify the language and take it to a new level.
2. Educated guess from various sources indicates that at present we have over 6000 human languages in the world (Tomasello 1996; Hauser, Chomsky & Fitch 2002; "Language" 1992) The bigger question is whether we can consider that language is an exclusive domain of human beings or have we inherited it from animals who occupied a place on the earth before our arrival. It is predictive that the biological function exercised by the voice box translates a noise into intelligible expressions. When these expressions got wider acceptance possibly words came up and then the necessity was felt to assign specific meaning to an intelligible sound which caused the evolution of letters, group of letters became words and groups of words became phrases sentences etc., Another close adjunct to language is a dialect. In fact it may be strongly considered that dialect is the forefather of language and in all probability a popular dialect became a language or groups of dialects owing to regional influence or demographic pattern or power centres whether local regional or trans-regional made a particular dialect stronger and base for a particular language. All languages rely on the process of semiosis to relate signs to particular meanings.
3. Another possibility could be that oral, manual and tactile languages contain a phonological system that governs how symbols are used to form sequences known as words or morphemes, and a syntactic system that governs how words and morphemes are combined to form phrases and utterances. Human language has the properties of

productivity and displacement, and relies entirely on social convention and learning. Its complex structure affords a much wider range of expressions than any known system of animal communication. Language is thought to have originated when early hominins started gradually changing their primate communication systems, acquiring the ability to form a theory of other minds and a shared intentionality (Tomasello 1996; Hauser, Chomsky & Fitch 2002; “Language” 1992; Trask 2007; Bett 2010; Devitt & Sterelny 1999; Hauser & Fitch 2003; Pinker 1994).

2. LITERATURE REVIEW

1. The *Ethnologue* establishes linguistic groups based on studies of mutual intelligibility, and therefore often includes more categories than more conservative classifications. Language is a system that consists of the development, acquisition, maintenance and use of complex systems of communication, particularly the human ability to do so; and language is any specific example of such a system (Tomasello 1996; Hauser, Chomsky, & Fitch 2002; “Language” 1992).
2. Interpretation of foundation of language has been attempted by numerous enthusiasts in diverse approaches. Study carried out in a logical and scientific manner is termed as Linguistics. Representation of experience by words was explored at length by Gorgias and Plato in ancient Greece. Rousseau went on to illustrate that language stemmed from emotions, whereas Kant considered that language emerged from careful and logical thought. The structuralism view of language was first introduced by Ferdinand de Saussure, (Smith 2007) and his structuralism remains foundational for many approaches to language (Agha 2006). Certain other views are that Steven Pinker considers that language is in some way related to animal cognition (Aikhenvald 2001). Certain others like Michael Tomasello considers it having developed from animal communication (Aitchison 2001). Some people estimate that age of spoken language is as old

as 1,00,000 years (Allerton 1989) and yet there are others who considers it as 2.5 million years (Anderson 2012; Aronoff & Fudeman 2011). A unique concept of “living language” has been created which means that there must be at least one person whose first language is that language. As of 2016, *Ethnologue* cataloged 7,097 living human languages (Baker 2001).

3. STATEMENT OF THE PROBLEMA

Death of a language is slow but sure process

4. METHODOLOGY

1. In this particular research problem on the rapid degeneration and disappearance of language, we would rather facilitate ourselves with “data analysis” rather than subjecting the problem on to any conventional methods of scientific testing on the process of death of a language. Thereupon we will drive ourselves to the analytics. It is universal that scope of analytics is violently spread, but it is an acceptable fact that in this field, there is all-embracing application of mathematics, statistics, as well as use of descriptive techniques and predictive models. From the data so mined using this process researcher gets insight from data, which are used to recommend action or to accelerate decision making in complex situations. We may try and use neural networks to do predictive modelling.
2. Demographic studies, linguistic segmentation, conjoint analysis and other techniques allow usage of large amounts of survey and panel data to understand the causative complexity. Here the factors involved are the people who use the language, technology which makes it popular by its usage and personality factors hence we would be liberal to add diminutive portions of people analytics, profile analytics, portfolio analytics, and risk associated with no-usage of a language in social media-risk analytics involving predictive models. As fallout digital analytics and software analytics

definitely form essential aspects of it. Extraction of data is a subsidiary factor, by which we use techniques to identify and probe on the characteristic of the data and how the gradual degradation because of non-acceptability of a language in technological field leads to its rusting, decay and death.

3. In this research predictive analytics is largely being adopted in which an array of analytical and statistical techniques used for developing models that have been used to guess future procedures or dealings. Now we have forms of predictive models which a situated for the vent or behaviour we want to accurately predict. They are measurable and we can dovetail them to produce a death score for the language we want to predict and these being scientific assure a reasonable accuracy. Nearly all predictive models produce a score; a higher score indicates that a given event or behaviour is very likely to occur. Another feature of decay of language is predictive behaviour. Hence using predictive analytics and data mining we try to suggest on trends and envisage future activities or consequences. Here we can gradually develop a scoring pattern at convenient frequency and develop death scores for languages. While developing a death score for a language it means that we have a reasonable estimate of reaching a conclusion whether a language will survive, and if so how long. Scalability, analytic optimization and partitioning will be important parameters in these techniques.

5. RESULTS AND DISCUSSIONS

1. Evolution of language can be juxtaposed with history of civilization. It is often true that evolution of language is a dynamic process and re-enacted by analysing the development of modern language. We find a subtle change in which it goes through many plateaus for stabilising. These plateaus are necessary to give depth and self correction for the growth of the language. If we go along with emergence of language, the trip-wire would be intelligible sounds which became familiar and followed a set pattern. Predictive analytics methods give a rough judgement that of the

available and in-use languages a large majority are likely to be dead by 2100. How does it happen? Is it due to cultural invasion, scientific adjustment or technological convenience? We may find that it is a combination of all these and certain sociological factors. For the sake of convenience if the world can be covered by fewer languages it may be for good. Yes, people debate the loss of cultural identity. It may not be significant. Take for example 900 languages spoken in Papua New Guinea involves that every kilometre you travel a person is subject to certain unintelligible sounds and to recognise it requires assistance. Such a situation is not a happy one. It is often interpreted that language is a mental faculty which due to biological and physiological development we see that only human beings can interpret language. There are strong biological bases for producing and understanding utterances. As per Charles Hockett the design features of human language set it apart from other known system of communication (Brown & Ogilvie 2008). As a matter of fact animal do communicate. A study carried out reveals that communication systems used by other animals are closed systems that consist of a finite, usually very limited, number of possible ideas that can be expressed (Brown & Ogilvie 2008). In contrast, human language is open-ended. It allows humans to produce a vast range of utterances from a finite set of elements, and to create new words and sentences. The reason why it is true is caused by the fact that human communication along with sound is accompanied with gestures to amplify it. There have been evidences to substantiate that communication systems have been passed on to human beings from great apes, because as per research pre-human australopithecines had arrangements in place for exchange of ideas which was very similar to the animals. The interesting aspect is that the pre-historic man introduced a distinction between diachronic and synchronic analyses of language thus emerged the modern discipline of linguistics. Speaking is a complicated process in almost all languages it needs a fine control of vocal and neurological apparatus because production of sound requires capabilities

of controlling lips and the ability to acoustically decode sounds (Campbell 2002). The genetic base which has a great role in language production is FOXP2, which if affected by mutations may result in congenital language disorder (Devitt & Sterelny 1999). Neurolinguists have incorporated non-insidious techniques like MRI to study the language process in individuals. Acoustic fragment of speech has two essential elements namely; segmental and suprasegmental elements. This is essentially sequential flow of distinct letters even though actually there are no clear boundaries between segments or audible pauses between words (Evans & Levinson 2009). The way in which meaningful elements can be combined within a language is governed by rules. The rules for the internal structure of words are called morphology and those for sentences and phrases as syntax (Goldsmith 1995). Language does not survive in a stand-alone environment hence it is reliant on its use by others. Largely a child learns language from their elders or peers. Language as a means of communication is transmitted from generations to generations and in this process diversifies to a new language. In this way more often than not some of the languages vanish or perish. One of the tests of a dominant language is the ability to accommodate elements from other languages through a process of diffusion. This generally happens when speakers of different languages comes into contact with each other on a common platform. In this process it is estimated that we generally get used to acquiring more than one language as a matter of processes or coincidence or convenience. Whatever may be the cause, the end result is classification of languages as first language, second language and so on. In this process, there is likelihood that many small languages take a back seat and dominant language take their place making a transformation from third language to first language or in any other combination. This process of enabling oneself to transform his first language from an existing one to an ever encompassing one results in decay and endangerment of a small language. If it is realised at this stage, then the

endangerment is averted as is being done for some languages or it results in death of the language (Hockett 1960).

2. In all societies more often than not a child gets adapted to the world by response of his senses to familiar sounds, may be the mother or the close environment in which it is brought up. As considered by the equity theory, whichever gives a pleasant experience is repeated and gradually gets accommodated. This may be the reason that an easily adaptable language like Hindi is abhorrence to many because most of the teachers on the subject complicate it will too many stringent grammatical restriction that one chooses to opt for a more flexible language. Hindi is just one example, there may be many like this and slowly disuse and repugnance results in decay and death of the language.
3. Acquisition of a language starts with the adjustment of foetus with familiar sounds. It has been also identified that unborn foetuses are capable of acquiring language to certain degree (Moseley 2010; Nerlich 2010). It is natural that at the early stages of conception and foetal formation presence of chosen few is seen in the vicinity of a new born child. Naturally as in all cases a child tends to more of these familiar human speeches rather than any other. Usually it takes about a year and a half for a child to acquire 50 odd communicative holophrases which are more generally phrases or half sentences.
4. Regularity of contact between people with different language origins results in diffusion of linguistic traits between those languages and multilingualism. Political arrangements made monolingualism the order of the day-one people, one state and one language. This was essentially a feature of populations inhabiting small islands. With the evolution of means of transport, interaction between people of different geographical discontinuities gathered momentum for trade, adventure etc and gave rise to the concept of language diffusion and ethno-national state. So it is not surprising that we have just about 200+ countries in this world analogous to 6000+ languages. This results in close contact of one language with another and natural fallout is some languages

loose out in this race and get rusted, endangered and when the last person who spoke that language dies, the death of the language takes place. Such arrangements have taken place in Balkan language area, Mesoamerican language area, Ethiopian language area and South Asia (Olson 1996; Romaine 2001). In this gradual process it is likely that language convergence, borrowing and relexification takes place which results in emergence of a new language.

5. As per study carried out the rapid rate at which ethnic languages are disappearing in the 20th and 21st century compared to earlier days this can be attributed to globalization, neo-colonisation and adaptability of a language to technological development (Saussure 1983). In such cases economically powerful languages dominate weaker ones and endangerment of that language takes place. In this the native speakers of that particular language shifts out to a more powerful and economically vibrant language. Then we will be left with a language with no native speakers and it becomes a dead language. As this process continues and if no one subscribes to that language then it becomes an extinct language. As in 2010 (Tomasello 1996) there are 6000 to 7000 languages spoken all over the world. Of these 20 languages which are spoken by over 5 crore people account for 50% of the world's population. Many other languages are spoken by far less number of people and suffice to say that these languages are the potentially endangered language. Over a period of time these less commonly spoken languages disappear leading to extinction leaving only a matter of interest to linguists or researchers.
6. There are two schools of thoughts on the extinction of language, with one supporting the requirement of fewer languages and the other opposing. As per the narrative in Old Testament (Ulbaek 1998) language diversity causes political conflict. This is contradicted by the fact that major episodes of violence like in Middle East, Syria, Yugoslavia, Rwanda or for that matter the American Civil War which had very low linguistic diversity and large majority of the stable political units like India have been multilingual (Van Valin

& Robert 2001; Zentella 2002). There is also a claim that language loss results in strangulation of cultural diversity. We may seek an all encompassing concept of trilingualism, which will maintain the cultural diversity on the one hand and help in availability of a national and international alternative when needed. Whether this step should be voluntarily adapted or by the influence of state machinery is debatable.

7. We are familiar with the synonyms language death, language extinction, linguicide or glottophagy. All leads to the death of an existing language owing to lack of subscribers. When the potency of a language reduces and the level of linguistic competence degrade attraction for that language reduces and people start looking out for languages with competence to adopt. This is also results in language death. A language shift of this nature results in language death, whose manifestation may be any one of these;
 - Rusting, decay and gradual language death: slowly, over a period of time
 - Lack of usage of the majority of people contributing death
 - Death as a consequence of State sponsored activity
 - Radical language death
 - Linguicide (also known as sudden death, language genocide, physical language death, biological language death)
 - Language attrition: the loss of proficiency in a language at the individual level
8. It may also be possible that the dominant language of a bilingual or trilingual concept gets the allegiance which results in seizure of the heritage language resulting into linguicide. Languages with a small, geographically isolated population of speakers can die when their speakers are wiped out by genocide, disease, or natural disaster.
9. Linguists distinguish between language “death” and the process where a language becomes a “dead language”

through normal language change, a linguistic phenomenon analogous to extinction. Possibly when the normal development of a language morphs into a state of different identity it is considered as a different language leaving the old one with no speakers.

6. WAY AHEAD

1. It is not very difficult to prevent a language from disintegrating or becoming endangered. Desire of a community to conserve its language has numerous options available. Take for example Modern Hebrew which was revived after generations modified and adaptable to the present requirements. Similar is the case with Irish, which has a strong state sponsored support for its re-emergence despite the strong challenge it faces from English. Another well established but not so opulent case is that of Maori called *kohanga reo*, “language nests” in New Zealand. This model is extended to Alaska, Hawaii, and elsewhere. A growing number of conferences, workshops, and publications now offer support for individuals, schools, and communities trying to preserve languages.
2. Then we have another phenomenon wherein new languages grow to form and replace dead ones like Tok Pisin, the national language of Papua New Guinea. Every group of related languages is separated from every other group by at least 5000 years of development, usually more. As per rough judgement “if English were to become the sole language of every person on earth, it would take tens of thousands of years to produce anything like the diversity that is our heritage-assuming we could somehow reproduce the conditions under which this diversity grew”.

7. CONCLUSION

1. Dominance of a particular community may impel them to adopt a strategic blunder of imposing a language on another section of population. In Asia we had two such incidents in

the recent past, of which one relates to the imposition of Urdu by tormenters of Pakistan to those who resided in East, which resulted in the emergence of Bangladesh and another in the offing which is the imposition of Hindi on South India. These incidents regenerate the native language and consider the other as outsider which causes its degeneration in that particular part. In Europe alone 119 languages have become extinct, primarily owing to political dominance of certain major European powers. Similar is the case in North and South America as well as Africa. Despite colonialism, Asian countries by and large withstood the cultural imposition leaving aside the fence sitters who have formed new religious communities, which were foreign to these areas. With the might of German Army under Hitler breaking the backbone of hegemonic tendencies of European powers and the emergence of China as a force to reckon these temporary transformations have taken a back seat.

2. It is natural that unlike Chinese or Japanese who have core culture philosophy; most of the languages are likely to disappear due to technological development. It is often found that cross cultural relationship may induce the gradual death of a language. Another factor may be due to language aping followed by the migrants to European as well as American countries. Here the technological acceptance of a language superimposes the cultural or ethnic value and after two or three generations, these migrants find their native language vanishing. These vanishing languages thereafter become endangered. However certain languages may more often than not become the mother language for so many other languages and perish like Greek, Latin, Sanskrit etc., A cruel way of a language becoming extinct is by outright genocide e.g., Tasmanian languages, Kalaallisut in Greenland, Kurdish language in Turkey, Native American languages, or certain African languages. A rough estimate gives a staggering figure of 5000-6703 languages spoken in the world. Predictive analysis indicate that of these a figures are likely to reduce by half and of which we may have just about 100 spoken by majority and rest being subscribed to by just over

few thousand people. By some estimates, 80% of the world's languages may vanish within the next century. Language extinction have occurred throughout the civilization, initially influenced by wars and now by technological revolution. Language extinction is a double edged phenomenon, wherein there are positive and negative fall outs. Technically speaking rather than emotional value or the ego state at which one is it would be better economically and otherwise to have fewer and lesser languages.

REFERENCES

- Agha, A. 2006. *Language and Social Relations*. Cambridge: Cambridge University Press.
- Aikhenvald, A. 2001. Introduction. In Y. Alexandra Aikhenvald & R. M. W. Dixon (Eds.), *Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics* (pp. 1-26). Oxford: Oxford University Press.
- Aitchison, J. 2001. *Language Change: Progress or Decay?* 3rd ed. Cambridge, New York, Melbourne: Cambridge University Press.
- Allerton, D. J. 1989. Language as form and pattern: Grammar and its categories. In N. E. Collinge (Ed.), *An Encyclopedia of Language*. London: Routledge.
- Anderson, S. 2012. *Languages: A Very Short Introduction*. Oxford: Oxford University Press.
- Aronoff, M. & Fudeman, K. 2011. *What is Morphology*. John Wiley & Sons.
- Baker, M. C. 2001. Syntax. In Mark Aronoff & Janie Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 265-295). London: Blackwell.
- Bauer, Laurie (2003). *Introducing linguistic morphology* (2nd ed.). Washington, D.C.: Georgetown University Press. ISBN 0-87840-343-4.
- Bett, R. 2010. Plato and his predecessors. In Alex Barber & Robert J. Stainton (Eds.), *Concise Encyclopedia of Philosophy of Language and Linguistics* (pp. 569-70). Oxford: Elsevier.
- Brown, K. & Ogilvie, S. (eds.) 2008. *Concise Encyclopedia of Languages of the World*. London: Elsevier Science.
- Campbell, L. 2002. Areal linguistics. In Bernard Comrie, Neil J. Smelser & Paul B. Balte (Eds.), *International Encyclopedia of Social and Behavioral Sciences* (pp. 729-733). Oxford: Pergamon.

- Devitt, M & Sterelny, K. 1999. *Language and Reality: An Introduction to the Philosophy of Language*. Boston: MIT Press.
- Evans, N & Levinson, S. C. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32/5, 429-492.
- Goldsmith, J. A. 1995. Phonological theory. In John A. Goldsmith (Ed.), *The Handbook of Phonological Theory*. London: Blackwell Publishers.
- Hauser, M. D. Chomsky, N., Fitch, W. T. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298/5598, 1569-1579.
- . & Fitch, W. T. 2003. What are the uniquely human components of the language faculty?. In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution: The States of the Art*. Oxford: Oxford University Press.
- Hockett, C. F. 1960. Logical considerations in the study of animal communication. In W. E. Lanyon & W. N. Tavolga (Eds.), *Animals Sounds and Animal Communication* (pp. 392-430). Washington D.C.: AIBS.
- Language. 1992. *The American Heritage Dictionary of the English Language*, 3rd ed. Boston: Houghton Mifflin Company.
- Moseley, C. (ed.) 2010. *Atlas of the World's Languages in Danger*. 3rd ed. Paris: UNESCO Publishing.
- Nerlich, B. 2010. History of pragmatics. In L. Cummings (Ed.), *The Pragmatics Encyclopedia* (pp. 192-193). London/New York: Routledge.
- Olson, D. R. 1996. Language and literacy: What writing does to language and mind. *Annual Review of Applied Linguistics*, 16, 3-13.
- Pinker, S. 1994. *The Language Instinct: How the Mind Creates Language*. New York: Perennial.
- Romaine, S. 2001. Multilingualism. In Mark Aronoff & Janie Rees-Miller (Eds.), *The Handbook of Linguistics* (pp. 512-533). London: Blackwell.
- Saussure, F. 1983. *Course in General Linguistics*. Trans. Roy Harris. La Salle, Illinois: Open Court.
- Smith, J. 2007. Sound change and the history of English. *SELIM*, 15, 151-156.
- Tomasello, M. 1996. The cultural roots of language. In B. Velichkovsky & D. Rumbaugh (Eds.), *Communicating Meaning: The Evolution and Development of Language* (pp. 275-308). Mahwah, NJ: Erlbaum.

- Trask, R. L. 2007. *Language and Linguistics: The Key Concepts*. 2nd ed. London: Routledge
- Ulbaek, I. 1998. The origin of language and cognition. In J. R. Hurford & C. Knight (Eds.), *Approaches to the Evolution of Language* (pp. 30-43). Cambridge: Cambridge University Press.
- Van Valin & Robert, D. 2001. Functional linguistics. In Mark Aronoff & Janie Rees-Miller (Eds.) *The Handbook of Linguistics* (pp. 319-337). USA: Blackwell.
- Zentella, A. C. 2002. Spanish in New York. In Ofelia García & Joshua Fishman (Eds.), *The Multilingual Apple: Languages in New York City*. Berlin: Walter de Gruyter.

DR. SANKAR RAJEEV

ASSOCIATE PROFESSOR MBA,
KL (DEEMED-TO-BE) UNIVERSITY,
GREEN FIELDS, VADDESWAREM,
GUNTUR 522502.

E-MAIL: <SANKARRAJEEV2010@REDIFFMAIL.COM>

PH: 9496742629