Languages: Why should we care for a dying one

shukla.vitthal@stud.hs-fresenius.de

2023-02-13, HS-Fresenius: Data Science for Business

Contents

Rendered at 08 February, 2023

Wordcount: 7398

1 Abstract

Languages are crucial to mankind. They have evolved from signs and sounds to written and spoken over a period of time. They form the base of our everyday activities and existence. It does not matter where or which part of the globe you are from, language is the only way to interact and act. In my report, I am going to try and understand what is a language and why should we care for a language which is on the verge of dying. The purpose of this report is to identify the most important and influential factors causing death of a language. Hopefully, the examination and monitoring of this factor, which serves as an ostensible marker, is not just a measure but a legitimate motivation to preserve it.

2 Introduction

In this study, I want to deeply understand about what is a language and the aspects related to it. I would like to describe a bit about how did I decided on choosing this topic as my report's topic. I am a person who is highly inclined towards new languages and cultures. I was always in India where we speak more than 200 languages and I was very curious about it. I even tried to learn 2-3 different languages during my time in graduation in a different city than my own. Coming to Germany and having started to learn German, has further driven my passion towards this subject matter. Thus, I will be trying my best here to make the readers get a better insight on languages and make them more interested in such topics.

I will be looking at a data set which has multiple variables about languages and using this data set, I will try to determine a few facts.

Language is a structured system of communication. The structure of a language is its grammar and the free components are its vocabulary. Languages are the primary means by which humans communicate, and may be conveyed through a variety of methods, including spoken, sign, and written language. Many languages, including the most widely-spoken ones, have writing systems that enable sounds or signs to be recorded for later reactivation. Human language is highly variable between cultures and across time.

Language is a means of communication. It is a means of conveying our thoughts ,ideas, feelings, and emotions to other people. Jack C. Richards and Richard Schmidt define the language :"the system of human communication which consists of the structured arrangement of sounds (or their written representation) into larger units, e.g. morphemes, words, sentences, utterances. In common usage it can also refer to non-human systems of communication such as the "language" of bees, the "language" of dolphins.

To define language is not an easy task. Different linguists tried to define language variously. However, if we analyse the definitions closely, we will find that each of these definitions is incomplete in some respect or the other. These definitions will raise a large number of questions. Some of the most commonly approved definitions of language given by the experts in the field of linguistics are given below:

Edward Sapir says: "Language is a purely human and non-instinctive method of communicating ideas emotions and desires by means of voluntarily produced symbols" This definition is rather incomplete because 'ideas, emotions and desires' are not the only things communicated by language. The term language covers a wide range of implication such as body language, sign language and animal language.

According to Hall, language is "the institution whereby humans communicate and interact with each other by means of habitually used oral-auditory arbitrary symbols." Hall's defi-

nition is narrow because it regards language purely as a human institution. We know that animals do communicate. Animals have their own language .In the words of Noam Chomsky, language is "a set of (finite or infinite) sentences, each finite in length and constructed out of a finite set of elements."

Chomsky focuses on the structural features of language. He showed how language can be investigated by analyzing it into its constituent elements. Each of these linguists focuses on certain aspects of language and ignores some others. However what they have said of language is true, though not comprehensive. As an object of linguistic study, "language" has two primary meanings: an abstract concept, and a specific linguistic system, e.g. "French".

The Swiss linguist Ferdinand de Saussure, who defined the modern discipline of linguistics, first explicitly formulated the distinction using the French word langage for language as a concept, langue(competence by Chomsky) as a specific instance of a language system, and parole(Chomsky's performance) for the concrete usage of speech in a particular language(Trask, 1999:92).

Human languages have the properties of productivity and displacement, and rely on social convention and learning. Estimates of the number of human languages in the world vary between 5,000 and 7,000. The scientific study of language is called linguistics. Language is thought to have gradually diverged from earlier primate communication systems when early humans acquired the ability to form a theory of mind and shared intention. This development is sometimes thought to have coincided with an increase in brain volume, and many linguists see the structures of language as having evolved to serve specific communicative and social functions. Humans acquire language through social interaction in early childhood, and children generally speak fluently by approximately three years old. Language and culture are codependent. Therefore, in addition to its strictly communicative uses, language has social uses such as signifying group identity, social stratification, as well as use for social grooming and entertainment.

Languages evolve and diversify over time, and the history of their evolution can be reconstructed by comparing modern languages to determine which traits their ancestral languages must have had in order for the later developmental stages to occur. A group of languages that descend from a common ancestor is known as a language family; in contrast, a language that has been demonstrated to not have any living or non-living relationship with another language is called a language isolate. There are also many unclassified languages whose relationships have not been established, and spurious languages may have not existed at all.

Academic consensus holds that between 50% and 90% of languages spoken at the beginning of the 21st century will probably have become extinct by the year 2100.

3 Data

Data was originally organized and published by The Guardian, and can be accessed via

[https://unesco.org.uk/]

[https://www.theguardian.com/news/datablog/2011/apr/15/language-extinct-endangered#data]

[https://www.kaggle.com/code/devisangeetha/preserve-languages/report]

[https://www.kaggle.com/datasets/the-guardian/extinct-languages]

Data sets are important because they can hold information that can be used by a program running on the system. Data sets are also used to store information needed by applications or the operating system itself, such as source programs, macro libraries, or system variables or parameters. My data set reflects my personal interest in languages and more specifically dying ones.

The full detailed data set includes names of languages, number of speakers, the names of countries where the language is still spoken, and the degree of endangerment. The UNESCO endangerment classification is as follows:

Vulnerable: most children speak the language, but it may be restricted to certain domains (e.g., home)

Definitely endangered: children no longer learn the language as a 'mother tongue' in the home

Severely endangered: language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves

Critically endangered: the youngest speakers are grandparents and older, and they speak the language partially and infrequently

Extinct: there are no speakers left

4   Source of Data in detail

UNESCO (United Nations Educational, Scientific and Cultural Organisation) is the United Nations body responsible for coordinating international cooperation in education, science, culture, and communication. Its goal is to help prevent war by building the defenses of peace in the minds of men and women.

Founded in 1946 by the UK government and established under Article VII of UNESCO's Constitution, they are a vital part of UNESCO's unique global network of 190+ national cooperating bodies known as National Commissions. They work to support the UK government's Permanent Delegation to UNESCO and to strengthen the ties between the UK's educational, academic, scientific, cultural, creative, and artistic communities and UNESCO's global program-mes and policies.

A recent Guardian blog post asks: "How many endangered languages are there in the World and what are the chances they will die out completely?" The United Nations Education, Scientific and Cultural Organisation (UNESCO) regularly publishes a list of endangered languages, using a classification system that describes its danger (or completion) of extinction.

The Guardian is a British daily newspaper. It was founded in 1821 as The Manchester Guardian, and changed its name in 1959. Along with its sister papers The Observer and The Guardian Weekly, The Guardian is part of the Guardian Media Group, owned by the Scott Trust.

Guardian Media Group is a global news organisation that delivers fearless, investigative journalism - giving a voice to the powerless and holding power to account. Our independent ownership structure means we are entirely free from political and commercial influence. Only our values determine the stories we choose to cover – relentlessly and courageously.

The Guardian is globally renowned for its coverage of politics, the environment, science, social justice, sport and culture. Their journalists deliver agenda-setting investigations, breaking live news, compelling opinion writing and the liveliest features, as well as award-winning pod-casts, video documentaries and info graphics and visuals.

Their teams in the UK, US and Australia produce theguardian.com, Guardian Australia, Guardian US, Guardian Weekly, and the Guardian and Observer newspapers in the UK.

5 Purpose of study

It is important that we realize before it's too late and we lose a language forever. Loosing a language is like loosing a part of our history and origins. Something that was there and which helped some people to communicate and exchange their ideas but now we are loosing it. Further in the paper we are going to look in depth about our data set and what is the impact of loosing a language.

6   Framework

To understand how many languages exist or are extinct, we need to understand the data set first and for this report as we are going to use R Markdown, we need to understand what is R and how does it work. The list explain the procedure in further detail:

Logical Organization of the R Markdown Script

1. Install and load libraries of R
2. Import data
3. Process data
4. Produce outputs (tables, plots, etc.)
5. Save outputs, if applicable (.csv, .png, etc.)

6.1   Install and load libraries of R

In the very first step, I need to install the packages that will be needed for my project. So, I used the function install.packages() to install packages like tidyverse, ggplot, and then I used the library() function to load them. Tidyverse provides a consistent, easyto- learn, and integrated approach to data analysis and manipulation. The packages in the tidyverse include ggplot2 for data visualization, dplyr for data manipulation, tidyr for data cleaning, and readr for data import, among others. The goal of the tidyverse is to provide a set of tools that work together seamlessly and make data analysis more efficient and enjoyable.

```
library(ggplot2)
library(dplyr)
library(tidyr)
library(RColorBrewer)
library(stringr)
```

```
library(gridExtra)
library(maps)
library(mapdata)
library(rebus)
library(ggmap)
```

6.2   Data Importing

I use the R programming language to analyze my data. I used data on languages which was available from public sources. Before we can manipulate and analyze data with R, we need to import data. R supports a variety of file formats, including.docx,.xls,.txt, and comma separated files like.csv.I imported my data with the function read.csv() and called it lang_data

```
lang_data<-read.csv("F:/Data Science Project/data.csv",stringsAsFactors = FALSE)
```

```
str(lang_data)
```

```
## 'data.frame':    2722 obs. of  15 variables:
##  $ ID                     : int  1022 1023 383 335 382 405 2211 1586 1020 1943 ..
##  $ Name.in.English        : chr  "South Italian" "Sicilian" "Low Saxon" "Belarusi
##  $ Name.in.French         : chr  "italien du sud" "sicilien" "bas-saxon" "biéloru
##  $ Name.in.Spanish        : chr  "napolitano-calabrés" "siciliano" "bajo sajón" "
##  $ Countries              : chr  "Italy" "Italy" "Germany, Denmark, Netherlands,
##  $ Country.codes.alpha.3  : chr  "ITA" "ITA" "DEU, DNK, NLD, POL, RUS" "BRB, LVA,
##  $ ISO639.3.codes         : chr  "nap" "scn" "act, drt, frs, gos, nds, sdz, stl,
##  $ Degree.of.endangerment : chr  "Vulnerable" "Vulnerable" "Vulnerable" "Vulnerab
##  $ Alternate.names        : chr  "Neapolitan; Neapolitan-Calabrese;          ;
```

```
##  $ Name.in.the.language      : chr  "" "" "Neddersassisch" "" ...
##  $ Number.of.speakers        : int  7500000 5000000 4800000 4000000 3500000 3500000
##  $ Sources                   : chr  "" "" "" "Hienadź Cychun: Weißrussisch. -
Lexikon der Sprachen des europäischen Ostens. Herausgegeben von Miloš Okuka unt"| __trun
##  $ Latitude                  : num  41 37.4 53.4 54 45.7 ...
##  $ Longitude                 : num  15.25 14.5 10.36 27.58 9.33 ...
##  $ Description.of.the.location: chr  "Campania, Lucania (Basilicata), Abruzzi (Abruzz
```

6.3 Data Cleaning

The below code helped me to clean my data by sorting out the variables and cleaning out
empty items.

```
cnames<-c("id","name_english","name_french","name_spanish","country","country_code","iso
colnames(lang_data)<-cnames
#str(lang_data)
dim(lang_data)
```

```
## [1] 2722    15
```

7 Language Family

The world's languages can be grouped into language families consisting of languages that
can be shown to have common ancestry. Linguists recognize many hundreds of language
families, although some of them can possibly be grouped into larger units as more evidence
becomes available and in-depth studies are carried out. At present, there are also dozens of
language isolates: languages that cannot be shown to be related to any other languages in

the world. Among them are Basque, spoken in Europe, Zuni of New Mexico, Purépecha of Mexico, Ainu of Japan, Burushaski of Pakistan, and many others.

8 What is Linguistics?

Linguistics is defined as the scientific study of language .It is the systematic study of the elements of language and the principles governing their combination and organization. Linguistics provides for a rigorous experimentation with the elements or aspects of language that are actually in use by the speech community. It is based on observation and the data collected thereby from the users of the language, a scientific analysis is made by the investigator and at the end of it he comes out with a satisfactory explanation relating to his field of study. This sort of systematic study of language has rendered the traditional method language study outmoded or unfit for any theorization. (Sreekumar, 2011 :20)

8.1 Modern Linguistics

Linguistics is the systematic study of the elements of language and the principles governing their combination and organization. Philology was the older term used to refer to the study of language. Philology was rather comparative and historical. A comparative study of language focuses on the similarities and differences within a family of related languages. A historical study analyses the evolution of a family of languages or the changes that occur within a particular language, over a long course of time. Saussure introduced new concepts and procedures in analyzing language. The following are some of the major terms and concepts introduced by him.(Aitchson, 1999:5)

Language is a system for communication using sounds or gestures that are put together in meaningful ways according to set rules. Trough language ,people can express their feelings , write poetry and novel, and even think. Language is essentially human , although

possibly not limited to humans. Linguistics on the other hand is the scientific way of studying language .This science try to answer the question such as What is the knowledge of language(competence), and how this knowledge put to use (performance).

9 Death of a language

The importance of a language rises and ebbs in direct proportion to the importance enjoyed by the people who speak that language. Of late, many different languages are being spoken on a mass basis by a large number of people. Therefore, it is becoming difficult for smaller languages to survive. However, the possibility of the revival of smaller languages is not entirely gloomy. There are a significant number of languages that have been successfully revitalized, and there are ongoing initiatives to reclaim other endangered and no-longer spoken languages.

Language death is considered to be something sad and quite unfortunate. For many speakers of widely spoken languages, such a case may be difficult to grasp. However, this case is real, and it happens around the world. For example, more than eleven percent of the world's languages have less than one hundred fifty speakers each. Additionally, there are a significant number of languages that are likely to be no-longer spoken within the next century (Nettle & Romaine 2000, p. 40, cited in Hoffmann 2009, p. 6). In such cases, there is someone who will be the last speaker and feel a great sadness (Krauss 1992, p. 6, cited in Hoffmann 2009, p. 6).

However, the loss of languages not only affect those people who speak them, but also linguists and anthropological communities. As an illustration, linguists may learn lots regarding human language by examining the forms that could be found in endangered and dead languages. However, with every loss of a language, the data in linguists' pool for study and their ability to make discoveries about the world gradually shrink. By reclaiming endangered languages and no-longer spoken languages, linguists are likely to gain a great source of information.

Moreover, researchers from other disciplines can benefit greatly from endangered languages' speakers, who often have useful and helpful knowledge about environmental science and medicine which could be unknown to modern science (Hoffmann 2009, p. 7).

9.1   What is language death?

There are many ways in which languages die. First of all, if people do not speak their language in their community and pass it to the next generation, their language may decline and eventually die. The second important point is that if a group of speakers of a certain language do not have access to favorable conditions where they can earn a livelihood then that language may slide into irrelevance (Nettle & Romaine 2000, p. 5). The third factor which must be considered is that if a group of speakers does not believe that language forms an important component of their uniqueness just like what happened in different populations in Wales and Ireland then the desire to ensure the survival of their language gradually recedes (Singh & Jones 2005, p. 84). The fourth point is that the language which is not passed from parent to children will ultimately become a dilapidated version of itself.

In addition to this, in case a language has just one remaining speaker then it will be deemed to have reached the end of its journey or be effectively dead whereas a language for which there are no present speakers is deemed to be extinct or dead(Crystal 2000, p. 20-21). Also, all the languages that are no longer in use today and there are no written documents or records of them can also be classified as extinct (Crystal 2000, p.2).

In certain circumstances or scenarios individuals speak their native tongue in only some specific settings like for religious purposes and do not use it to interact with others. In such scenarios it is predicted that such languages will continue to remain viable for a reasonable time period and may regain certain degree of importance as vernacular tongues just like the instances of Ancient Greek, Church Slavonic, Coptic and Latin (Wurm, 1991, p.1). The above mentioned languages have been utilized primarily for academic purposes or religious

practises; an example of this is Israel where increased use of Hebrew language is being seen on a day to day basis (Nettle & Romaine 2000, p. 188).

In certain other scenarios no speakers are available for certain languages and as a consequence of this they can't be used for limited purposes also, only a few expressions or words are used. The group under which such languages are placed is that of not being totally dead or not being totally alive since written documents are available and there is a likelihood of restoration of their previous standing, in other words they can be classified as sleeping languages which are likely to rise from slumber. An example of this is Miami language, an Algonquin language which is used widely in Oklahoma and had been facing a decline in the last 30 years but is undergoing a revival at present (Leonard 2008, p. 2).

9.2 How many languages are currently in use?

Nowadays, there are about 3,000 to 10,000 languages that are widely spoken; however, there are a number of sources that give a figure between five thousand and seven thousand (Gordon, 2005). Grenoble & Whaley (1998) give a figure between 5,000 and 6,000; Ruhlen (1991, p. 391) believes that the figure is 5,000; and the Global Language Register suggests about 10,000 that includes some dialects as separate languages. In addition, Dixon (1997, cited in Černý 2010, p. 52) estimates that there are approximately 5,000 to 6,000 living languages in the world. However, the differences between dialect and language are largely dependent on sociopolitical factors (Crystal, 2000, p.8). Clearly, there is no significant agreement about the number of languages currently in use, but it may range between 6,000 and 10,000 languages.

9.3 How many languages are endangered?

About 50% of the languages in the world have less than 10,000 speakers and there are 548 languages which can boast of less than 100 people who can speak it (Gordon, 2005). In addition, some of these languages are at risk of disappearing, and over the past few centuries

the rate of language death has significantly increased (Wurm 1991, p.1). According to some dire estimates, there are more than four thousand languages that may become disappear by the end of the twenty-first century (Krauss 1992, p.7), and people are likely to use languages such as Spanish, Mandarin Chinese, Arabic, and English (Grenoble & Whaley, 1998). The Foundation for Endangered Languages conducted a study which focusses on the situation and postulated that greater than 50% of the languages present globally at present are facing the threat of demise and that the majority of languages will disappear within the span of a few generations (Foundation for Endangered Languages 2009).

9.4   How and why do languages die?

There are a number of reasons or events that contribute to the deaths of languages around the world. First of all, languages are likely to die quickly when all or most of their speakers pass away because of wars, genocides, or reprisals (Wurm 1991, p. 2). As an example, in 1932 in El Salvador, the speakers of Pipil (Nawat) stopped speaking their mother language since they feared further reprisals (Nettle & Romaine 2000, p. 6). Another reason for language death is the spread of infectious diseases among communities of people who do not have adequate immunity to resist such illnesses. Infectious diseases have a noticeable impact on indigenous peoples and their languages. An example is the impact of European entry in the Americas, about 90% of the native population passed away due to illnesses that they contracted from Europe and or their livestock (Crystal 2000, p.72).

Whenever a situation occurs where a geographical area or nation is conquered or overtaken by colonial masters then in the majority of cases it is the language spoken by the new entrants which becomes more important and the natives are under pressure to adopt the language of the new entrants. Also, the native speakers might feel that they have to incorporate the languages and traditions of the newcomers. In addition, this kind of scenario has a number of financial advantages associated with use of mastering the language of the newcomers like job

opportunities, facilities and availability of new products. This scenario may also be witnessed when individuals migrate to a different nation where a completely different language is the language of conversation (Wurm 1991, p 5; Crystal 2000, p.77).

Another reason for language death is language shift or the process of assimilation that usually happens due to conquest or colonization. During this period, some bilingualism often occurs wherein local people retain their own language(s) while learning the language of the incoming people. In majority of such situations, the time period associated with this change is different but when the younger generation of native population is more at ease with the language of new entrants instead of their native tongue then there is a greater possibility of them considering the native tongue as defunct for serving their purposes. In addition, people may start to look down upon their mother tongue and only utilize it in very few circumstances. Also, the number of people able to speak a language decreases when languages are not transmitted from parents to their children. Hence, the final result will be that the language will slowly disappear as it will be used in very few situations and the people who are still able to speak it may have to face segregation and may be forced to limit their social interaction. The native tongues may be altered and changed to a simpler form due to the influence of the dominant language in terms of grammatical formations or incorporation of a large number of idioms or words from it and if this change in the native language is not halted then there is a greater possibility of disappearance and demise of the native form (Crystal 2000, p. 79; Wurm 1991, p.13).

Many scenarios have been played out globally where languages have been knowingly crushed. For instance, in Kenya schools, any child caught speaking Gikuyu was fined, caned, or forced to carry a slate saying, "I am a donkey," or "I am stupid." For another example, in Welsh schools, any student caught speaking Welsh in or around school had to wear a sign inscribed "Welsh Not." The punished child had to wear the slate until the end of the day. Similar systems were used in many places; for example, Breton in Brittany: if anyone speaks their native languages at school in these communities, their mouths will be washed out with soap

and water.

As a result, some native speakers become ashamed of their mother tongue; they become embarrassed and reluctant to use it. In addition, they feel their languages are inferior and lacking in importance, and because of that they do not pass their language to the next generation because they do not want them suffering in a similar way (Singh & Jones 2005, p. 83; Crystal 2000, p. 84).

9.5   Why should we worry about languages dying?

Because languages express identity: Language plays a leading role in defining people's identity, and the loss of language may impact people in negative ways. According to Sir James Henare, who expressed the influence of the loss of the Maori language, "The language is the life force of our Maori culture and mana ['power']. If the language dies, as some predict, what do we have left to us? Then, I ask our own people who are we?" (Nettle & Romaine 2000, p. 23). However, language can influence people in positive ways as well. As an illustration, people can have an emotional experience and sense of pride and identity when they reacquire their ancestors' mother tongue. For example, Aden Ridgeway, who is an Aboriginal politician said, "Language is power; let us have ours!" (Ridgeway 2009).

Another example is a participant who attended a workshop in May 2012 (Port Lincoln, Eyre Peninsula, South Australia) about Barngarla Aboriginal language reclamation who expressed that learning the language was "liberating" and that she felt a sense of identity. Moreover, another one said that their ancestors are happy (Zuckermann 2013). Identity is a way of being that makes the individuals of any community recognizably the same. People can discern each other's identities by looking at customs like dress, ritual, beliefs, and the whole panoply of personal behaviors, but language is the most ubiquitous among all these behaviors. It is available even when people cannot see each other, as in talking in the dark or shouting at a distance. Language is the primary symbol, or register, or index of identity (Crystal 2000,

p. 39). There are many metaphors that describe language as a way of discerning identity. According to Dixon (1997), "A language is the emblem of its speakers," and Barthes (1977) said, "Language is a skin" (Crystal 2000, p. 40).

There are many proverbs that capture the importance of identity, and without language, that identity would be lost. As an example, a Welsh proverb says, "Cenedl heb iaith, cenedl heb galon." It means, "A nation without a language is a nation without a heart" (Crystal 2000, p. 36). Since languages provide valuable information to the mankind's search for facts: Language can reflect a unique interpretation of human existence, and this may give people another motive to care about languages' deaths. Linguists and academics can learn a great deal from them; for example, endangered and no-longer-spoken languages may provide various levels of intellectual profundity, more so than widely spoken languages.

According to a famous proverb from Slovakia, "With each newly learned language, you acquire a new soul." In addition, there is a proverb from France that can describe how language can play an important role in knowledge: "A man who knows two languages is worth two men." According to Emerson (1860), who takes up this theme, "As many languages as he has, as many friends, as many arts and trades, so many times is he a man." These messages and proverbs state clearly that there is a lot of data and knowledge that can be enjoyed and learned by experiencing other languages (Crystal 2000, p. 44).

The loss of languages has a significant influence on human knowledge. In other words, languages' deaths can affect academics and linguists who study the human capacity of other languages. As an illustration, the more academics and linguists know about the varied methods of using other languages, the more knowledge they will have about human creativity. According to Nettle & Romaine (2000, p. 11), "Linguistic diversity gives us unique perspectives into the mind because it reveals the many creative ways in which humans organize and categorize their experience." Clearly, the fewer languages or methods of documentation that linguists know, the less diversity they can work with, and the less diversity academics can

draw on to learn about human creativity in language. Moreover, no-longer-spoken languages or endangered languages contain a huge amount of data that can be beneficial for the field of linguistic study. For instance, at the beginning of linguistics as a field of study, the majority of scholars focussed on Indo-European languages as they were associated with many different languages, had strong interrelationship and demonstrated many similarities (Nettle & Romaine 2000, p. 28). But, languages that are on the brink of extinction or those which have become defunct have huge amount of information stored within them which can act as a source of important and beneficial information to both linguistic researchers and scholars. According to Nettle and Romaine (2000, p. 56), "Many Hawaiians have now forgotten more of that local knowledge accumulated and handed down orally over the past 2,000 years than western scientists will ever learn."

As these data were communicated orally, they have been forgotten along with the language. If a language is allowed to become extinct, then all this important and useful information will disappear with no chances of recovery. In a nut shell, it can be said that linguistic researchers and people who speak such endangered tongues might have to face the threat of loss of valuable information and facts in case such languages become extinct. In addition, both linguists and speakers of such languages can work together in order to be able to help each other to reclaim these languages; however, the situation is not that simple (Hoffmann 2009, p. 22).

Because of aesthetic reasons: Languages play a vital role in defining people's cultures; in other words, language ties closely with culture, and when communities lose their language, they lose a number of aspects of its culture. As an illustration, there are a considerable number of practices that depend on language. For instance, losing language is likely to lead to the loss of many verbal art forms such as poetry, and traditional songs. Even if the language was written, it may lose some aspects such as its tales unless they are translated into other languages (Hoffmann 2009, p. 21). Language loss has substantial implications for both the individual and the collective identities of any community. Ken Hale, who worked

with a considerable number of endangered languages, said, "When you lose a language, you lose a culture, intellectual wealth, a work of art. It's like dropping a bomb on a museum, the Louvre" (Ken Hale 2001, cited in Zuckermann 2013). According to Boroditsky and Gaby (2010, cited in Zuckermann 2013), the Kuuk Thaayorre, which is a language spoken in Pormpuraaw in Cape York, its speakers always use cardinal directions instead of 'left' or 'right.' The speaker of Kuuk Thaayorre always can understand where he or she is situated, and this use of directions influences one's awareness of time. This can indicate that each language has various levels of intellectual profundity

9.6 What can save languages from dying?

Despite the fact that there are a number of forces that can cause death to indigenous languages, many forces which can support endangered languages. As an example, languages are less likely to disappear if their speakers feel that it is necessary and valuable to save their languages. Furthermore, language can be preserved by implementing methods such as using the language in media and technology and gaining official and community support (Hoffmann 2009, p.18). As per the view of Krauss (1992) according to whom television acts as a nerve gas with the ability to impact cultural beliefs, both television and radio can play a vital role in maintaining the importance of a language. If a program is prepared in a native tongue then there will be more chances for speakers to hear their language being spoken. Also, use of native tongue in mass media in place of the commonly used dominant language proves that the native tongues can function successfully in the modern era as well. Various charitable groups have lent their support to the cause of utilization of endangered tongues in media. An example is the Guatemala Radio Project which provides backing to local radio stations, helps in broadcasting in native tongue like news, programmes related to education and health and classical music, all reinforcing pride in Mayan heritage." Therefore, it is clear that media can play a potential role in enhancing and promoting indigenous language and

culture (Cultural survival, 2009).

10   Endangerment Degree

Language endangerment occurs when a language is at risk of falling out of use as its speakers die out or shift to speaking another language. Language loss occurs when the language has no more native speakers, and becomes a dead language. If eventually no one speaks the language at all, it becomes an extinct language. While languages have always gone extinct throughout human history, they have been disappearing at an accelerated rate in the 20th and 21st centuries due to the processes of globalization and neo-colonialism, where the economically powerful languages dominate other languages.

The more commonly spoken languages dominate the less commonly spoken languages, so the less commonly spoken languages eventually disappear from populations. Of the between 6,000 and 7,000 languages spoken as of 2010, between 50 and 90% of those are expected to have become extinct by the year 2100. The top 20 languages, those spoken by more than 50 million speakers each, are spoken by 50% of the world's population, whereas many of the other languages are spoken by small communities, most of them with less than 10,000 speakers.

Notwithstanding claims that the world would be better off if most adopted a single common lingua franca, such as English or Esperanto, there is a consensus that the loss of languages harms the cultural diversity of the world. It is a common belief, going back to the biblical narrative of the tower of Babel in the Old Testament, that linguistic diversity causes political conflict,[34] but this is contradicted by the fact that many of the world's major episodes of violence have taken place in situations with low linguistic diversity, such as the Yugoslav and American Civil War, or the genocide of Rwanda, whereas many of the most stable political units have been highly multilingual.

Together, the eight countries in red contain more than 50% of the world's languages. The areas in blue are the most linguistically diverse in the world, and the locations of most of the world's endangered languages.

11 Plots

11.1 Find out how many speakers exists under each degree, obviously extinct won't have any speakers at all.

```
lang_data %>% group_by(endangerment_degree)%>% arrange(desc(Number_of_speakers))%>%   ggp
```
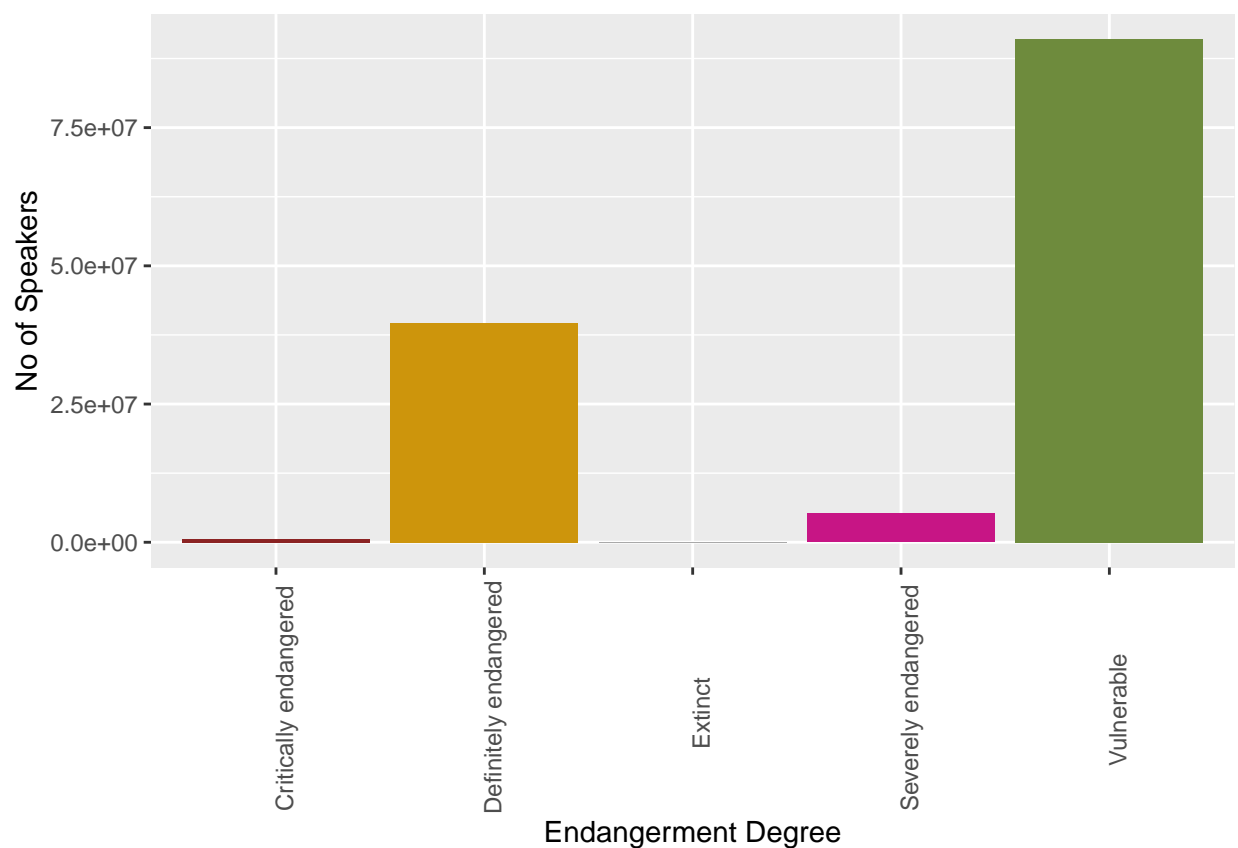
## 11.2 Find out the number of languages under endangerment degree

```
degree_data<-lang_data %>% group_by(endangerment_degree) %>% summarise(lang_count=n_dist
```

```
ggplot(degree_data,aes(x=endangerment_degree,y=lang_count,fill=endangerment_degree))+geo
```
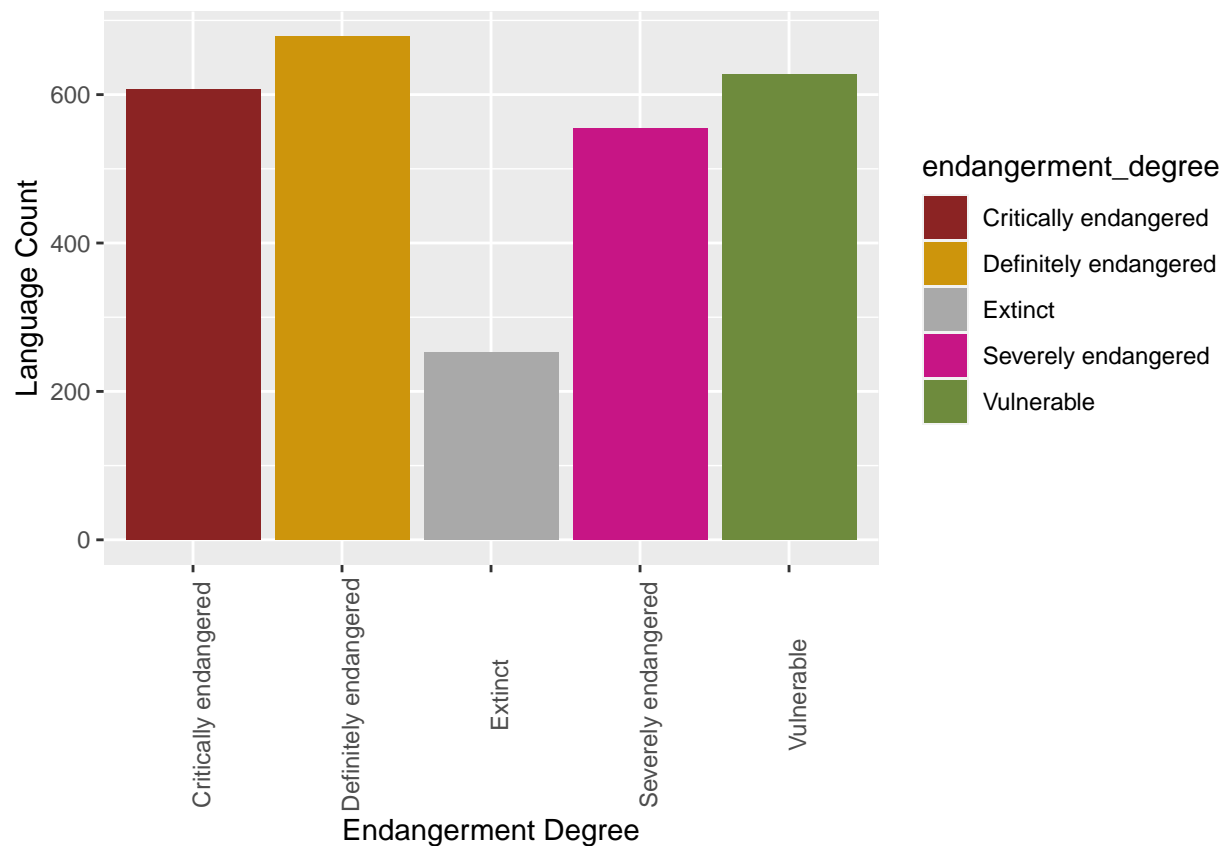

Around **250** languages are under **Extinct** category, for which measures should be taken to preseve it.

11.3 Which country has got more extinct languages

```
lang_data %>% filter(endangerment_degree=="Extinct")%>%group_by(country)%>% summarize(n=
ggplot(aes(x=factor(country,levels=country),y=n))+geom_bar(stat="identity",width=0.4,col
labs(x="Country",y="No of Languages",title="Countrywise Extinct Languages")
```
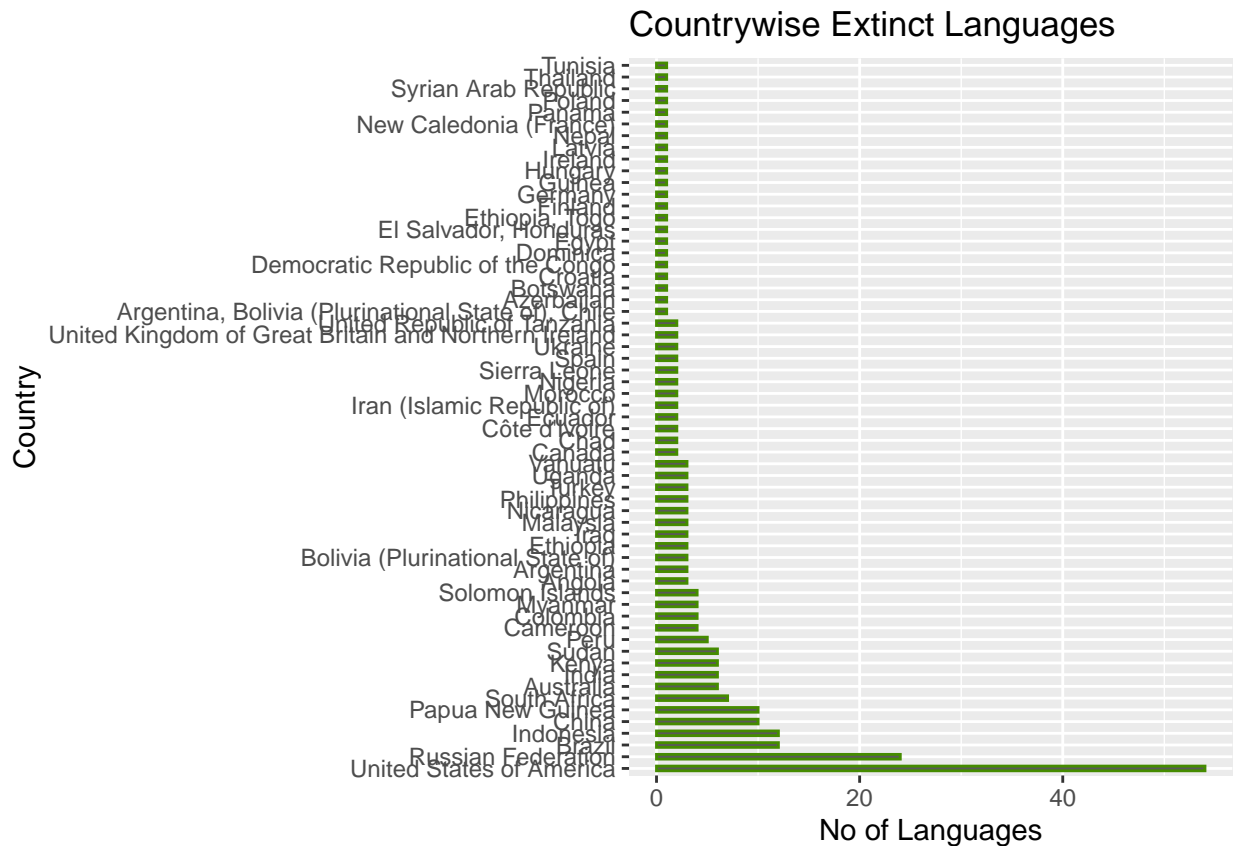

11.4 Mapping languages grouped by Endangerment Degree

```
cols<-c("brown4","darkgoldenrod3","blue","mediumvioletred","darkolivegreen4")
cols_m1<-cols[factor(lang_data$endangerment_degree)]
map("world", fill=TRUE, col="lightgrey", bg="lightblue", ylim=c(-60, 90), mar=c(0,0,0,0)

points(lang_data$longitude,lang_data$latitude, col=cols_m1, pch=16)
legend(x=-50,y=-30,legend=paste(rep(c("Definitely Endangered","Critically Endangered","E
```

```
col=rep(cols,times=1),pch=16,bty="0",cex=0.5,pt.cex=0.7)
```


Legend:
- Definitely Endangered
- Critically Endangered
- Extinct
- Severely Endangered
- Vulnerable
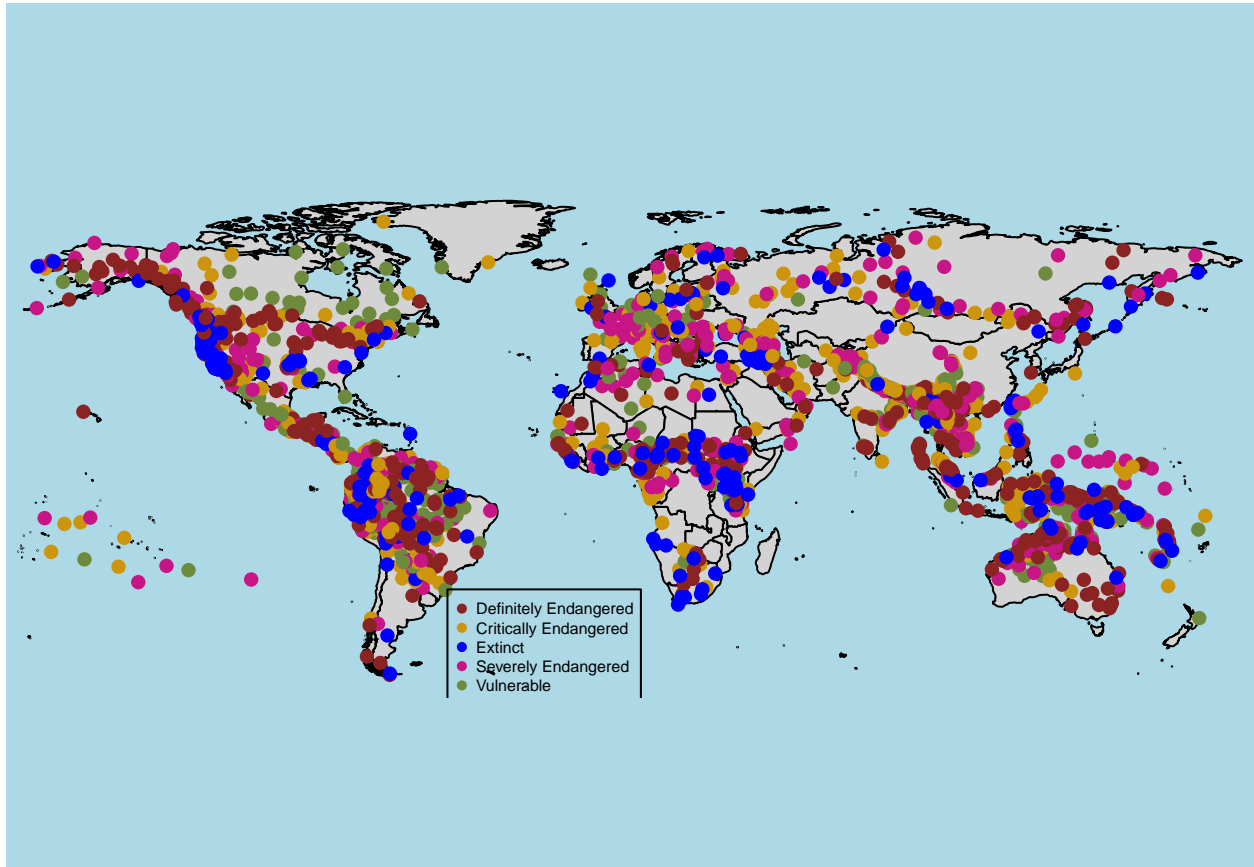
12 Conclusion

To sum up, loss of language is classified as a terrible loss. However, there are many languages becoming increasingly widely spoken; thus, it is becoming more difficult to revive some languages, such as Aboriginal languages. Reviving, revitalizing, and maintaining these kinds of languages is not easy; however, it is still possible. There are a number of factors that can revive endangered and no-longer-spoken languages, including the community of language, the time frame in which the revival begins, the level of official and community support, the dedication and motivation of groups and individuals, and official governmental recognition for the language.

According to Grinevald (1998, p. 143), "Bridging the gap between academic linguistics and community wants and efforts is surely one of the major challenges of the linguistic profession as it faces the situation of endangered languages at the turn of the new century." A significant number of challenges that can be found, but by working together, endangered language speakers and linguists can benefit simultaneously.

This paper has aimed to explore themes of language death in various respects. They are the meaning of language death, the number of languages that are currently in use, the number of languages that are endangered, and how and why languages are dying. Why should we care about languages dying? What can be done to save languages from dying? Indeed, significant research has highlighted the problems surrounding the loss of languages, and it has provided some constructive solutions to reclaim and revive dead and endangered languages. In fact, in linguistics, a great deal of time has elapsed.

Therefore, linguists should begin to put their acts together before such languages wind down. As Dorian (1998, p. 21) said, "Having waited too long before undertaking to rally support for threatened languages, we may find ourselves eulogizing extinct languages whose living uniqueness we had hoped instead to celebrate."

The linguists on the front line, who work in this field, should intervene to reclaim dead and endangered languages because time is running out. It might be late for some languages, but linguists may hold in their hands the future of other languages. Therefore, linguistics needs much support in order to study and revive languages that are definitely in danger of disappearing or becoming severely endangered, critically extinct, or critically endangered. In addition, the raising of public awareness will play an important role, and hopefully, this research may play its part in that.

13   Obstacles I overcame

This project has some obstacles that I can overcome, but others that I cannot, as I will explain in the next chapter.

Data quality: The data I have collected may be incomplete, inaccurate, or inconsistent, which can make it difficult to conduct reliable analysis. To address issues with data quality, I conducted a thorough data cleaning and pre-processing step. This may include checking for missing values, out lairs, and inconsistencies and addressing them accordingly.

Lack of domain knowledge: Because language data is such a rich quality data and I am a very young and inexperienced student, I took some time to understand the data and read it in a proper way.To overcome this obstacle, I did some research in the language sector and familiarized myself with key concepts and terminology.

New to coding and R: This was my first time using R and its tools for a report. In fact, this was my first time using any kind of machine language in my life. This was totally new and complex for me and I tried to learn it in a very short span of time and thus I might have done mistakes here and there but still it was a great learning journey.

Additional: I was struggling to knit my report to PDf as well as create a table of contents and do a word count. I looked into multiple websites, youtube, FAQs etc and by trial and eroor, I was able to figure it out and get the output. This maybe a small issue but for me it was a big achievement considering my newness to R. Also, while knitting to pdf, I was getting an error saying some sigma special character is causing error. I figured out to use latex_engine: xelatex and my PDF worked just fine.

14 Remaining Challenges, Problems, and Weaknesses

Limited data size: With only a limited set of data variables, and my limited knowledge, I was only able to produce simple output. Nevertheless, this was a great learning experience for me and has opened up new possibilities for me to look into.

Time constraints: I had limited time to complete the project, so I needed to prioritize and focus on the most important tasks and impact ful aspects of my project.This may have caused me to miss some information, but I decided to set realistic goals and deadlines for each task to stay on track.

15 How I Would proceed with My Work If I Had More Time and Resources

If I had more time, I would have tried to explore more features of R and tried to use more types of packages and complex analysis such as regression analysis or data augmentation. I would have produced more plots on my own and look into other works available on my topic and derive even more rich data and its analysis. Also, one aspect would be to learn how to use JabRef and BibTex for citations. I would try to learn and improve my work by using such amazing tools which makes the work alot more easier and organised.

16 Results

I would like to conclude this report by saying a big thank you to my professor Mr. Huber who not only motivated us to believe in ourselves and push our limits, but also showed us that anything can be achieved if we try hard enough. For a non-technical person like me, if I can create a report using a machine language, so can anyone.

Apart from this, I hope this work which I have produced is up to the expectation of the university and the professor. Thank you for your time.

17 References

Aitchson .J (1999) Linguistics,Teach yourself, Library of Congress : London.

Crystal, David (2008) A Dictionary of Linguistics and Phonetics. UK: Blackwell Publisher.

Finch. G (1997), How to study Linguistics, PALGRAVE MACMILLAN: New York.

Mamkjaer .K (1995), The Linguistics Encyclopedia , Routledge: London

R.L.Trask. (1999) ,KEY CONCEPTS IN LANGUAGE AND LINGUISTICS, Routledge: New York.

Sri. Sreekumar. P.(2011) LANGUAGE & LINGUISTICS, S.V. College of Advanced Studies Cheruvannur, Calicut.

Victoria Fromkin, Robert Rodman ,Nina Hyams.(2014) An Introduction to language ,10th ed Wadswort: New York.

Widdoson H.G (1996)Linguistics. New York: Oxford university press.

Yule, George (2010) The Study of Language. Cambridge:Cambridge University Press

Černý, M 2010. Language death versus language survival: A global perspective. Beyond Globalisation: Exploring the Limits of Globalisation in the Regional Context, pp.51-57.

Crystal, D 2000. Language death. 1st ed. Cambridge, UK: Cambridge University Press.

Cultural survival, 2009. Community Radio Project | Cultural Survival.

Dorian, N 1998. Western language ideologies and small-language prospects. In Grenoble and Whaley (eds.), 3– 21.

Fishman, J 1991. Reversing language shift. 1st ed. Clevedon: Multilingual Matters.

Foundation for Endangered Languages, 2009. Manifesto - Foundation For Endangered Languages -.

Gordon, R 2005. Statistical Summaries. Ethnologue

Grenoble, L & Whaley, L 1998. Endangered languages. 1st ed. Cambridge [England]: Cambridge University Press.

Grinevald, C 1998. Language endangerment in South America: a programmatic approach. Endangered languages: Current issues and future prospects, pp.124–159.

Hoffmann, M 2009. Endangered Languages, Linguistics, and Culture: Researching and Reviving the Unami Language of the Lenape.

Jones, M. and Singh, I 2005. Exploring language change. 1st ed. London: Routledge.

Kolbert, E 2005. Last Words: A language dies. The New Yorker June 6, 2005.

Krauss, M 1992. The world's languages in crisis. Language, 18(1), pp.4–10.

Leonard, W 2008. When is an" Extinct Language" not extinct. Sustaining linguistic diversity: Endangered and minority languages and language varieties, pp.23–33.

Nettle, D & Romaine, S 2000. Vanishing voices. 1st ed. New York: Oxford University Press.

Nettle, D & Romaine, S 2000. Vanishing Voices: The Extinction of the World's Languages. 1st ed. Oxford [England]: Oxford University Press.

Ridgeway, A 2009. Language is power; let us have ours. The Sydney Morning Herald [online].

Ruhlen, M 1991. A guide to the world's languages. 1st ed. Stanford, Calif.: Stanford University Press.

UNESCO, 2009. UNESCO Atlas of the World's Languages in danger.

Wurm, S 1991. Language death and disappearance: causes and circumstances. Diogenes, 39(153), pp.1–18.

Zuckermann, G 2013. Historical and moral arguments for language reclamation. History and Philosophy of the Language Sciences.

Almurashi Wael, 2017. International Journal of English Language and Linguistics Research Vol.5, No 5