

ENG423A Current Issues in Linguistics

Language Documentation and Archiving Part 1 - Endangerment

Chaithra Puttaswamy

Endangered Languages






- UN has declared this decade as the Decade of Indigenous Languages.
- There are approximately 7,000 languages in the world, and 95 percent of the world's population speak 300 of them. Half the world speaks the largest 16.
- Linguists typically celebrate the tension that plays between two realities of human linguistic knowledge, universality and diversity. But linguistic diversity is not something whose future can be taken for granted. Many local languages and cultures find themselves in great peril in this era (Hale 1998: 35)

- Thomason (2015) argues that the speakers age is the most reliable indicator of language vitality. The more children there are who are learning the language as their first language, the less likely the language is to disappear any time soon.
- Speaker numbers are obtained from census reports and they do not provide any evidence of the age of the speakers.

Factors contributing to language endangerment

- **Endangerment by numbers:** Hebrew, Avestan
- **Conquest:** if the conquerers language replaces the language(s) of the conquered. -North America, Australia. Political pressure is not the sole cause of endangerment. Melting pot phenomenon.
- **Economic pressure:** “you want a job? Speak my language!”
- **Language politics:** China, Bangladesh, Sri Lanka
- **Genocide, fatal disease introduced by foreigners:** Rare with respect to linguistic communities

- **By domains of use:** school/education; media; socio-cultural space; religion; trade and commerce, village, home
- **Due to language contact-** a precondition for language endangerment.
Bilingualism is the first step towards endangerment especially when the social, political and economic relations between the two communities in contact are markedly asymmetrical.
- **Wealth and resources for language maintenance**
- **Language Attitudes:** What people think about their language – its values, its usefulness, its importance to their culture -can play a decisive role in a language's fate. (Thompson 20 : 26)
- **Loss of Linguistic diversity via standardization:** Hindi

| ¹ Degree of endangerment | | Intergenerational Language Transmission |
|---|------------------------------|---|
| | <i>safe</i> | language is spoken by all generations; intergenerational transmission is uninterrupted >> <i>not included in the Atlas</i> |
|  | <i>vulnerable</i> | most children speak the language, but it may be restricted to certain domains (e.g., home) |
|  | <i>definitely endangered</i> | children no longer learn the language as mother tongue in the home |
|  | <i>severely endangered</i> | 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 |
|  | <i>critically endangered</i> | the youngest speakers are grandparents and older, and they speak the language partially and infrequently |
|  | <i>extinct</i> | there are no speakers left >> <i>included in the Atlas if presumably extinct since the 1950s</i> |

- Vulnerable (81) - Adi, Ao, Irula, Gondi, Ho, *Bodo*, Bishnupriya Manipuri, Kokborok, Ladhaki, Mundari, Wancho.
- Definitely endangered (62) Badaga, Khalasi, Koch, Kodagu, Kolami, Lepcha, Malto, Nahali
- Severely Endangered (7) – A'tong, Remo, Mech, Geta?, Mech
- Critically endangered (47) – Andamanese, Jarawa, Koraga, Kota, Toda, Kuruba, Parji, Pengo, Aimol, Ruga, Nihali
- Extinct: Ahom, Andros, Rangkas, Sengmai, Tolcha

Endangerment and losses

- **What a community loses:** cultural identity, loss of artistic expression, loss of cultural knowledge.
- **What science loses:** Language loss as a threat to our understanding of human history, human cognition, and the natural world.

What can we do about language endangerment?

- 1) Do nothing
- 2) Documentation: criticised as museum-oriented approach; commodifying languages
- 3) Promote revitalisation: create socio-cultural contexts for language use
- 4) Relate language endangerment to wider socio-economic policies, politics, environmental pressures and so on.

Language Documentation

- “The concept of language documentation as a field of linguistic research and activity in its own right proceeds based on the assumption that the linguistic practices and traditions in a given speech community are worthy of documentation in the same way as material aspects of its culture (arts and crafts) are generally deemed worthy of documentation. The aim of a language documentation, then, is to provide a comprehensive record of the linguistic practices characteristic of a given speech community. Linguistic practices and traditions are manifest in two ways: (1) the observable linguistic behavior, manifest in everyday interaction between members of the speech community, and (2) the native speakers' metalinguistic knowledge, manifest in their ability to provide interpretations and systematizations for linguistic units and events.” – Himmelmann 1998: 166

- The goal of language documentation is not only to collect the corpus of written texts but also to inculcate the native speakers and their linguistic knowledge during corpus collection.
- In other words, spoken form of a language acquires dominance over the written form.
- Many linguists have tried to create a fine line between the processes and methodologies involved in linguistic documentation and description.
- Linguistic documentation is an urgent and a specific task.
- The primary goal is to collect data through field work.
- Also, documentary linguistics strives to create an archive which preserves the collected linguistic, cultural and traditional knowledge of the language and its speakers.

Documentation vs Description

- Descriptive linguistics on the other hand makes an attempt to the describe the structure of a language with the help of the data collected. Lehmann stresses that linguistic description aims at:
 1. essential completeness
 2. intelligibility
 3. adequacy - Linguistic description must first and foremost satisfy explanatory adequacy.

- Himmelmann (1998) and Lehmann (1999) point out that linguistic documentation and description are theoretically two separate fields.
- Woodbury on the other hand, agrees with the above mentioned proposition but considers that both are equally important aspects of linguistic fieldwork.
- Dixon (2007) lays emphasis on the description of the language.
- Austin (2010) gives equal weight to linguistic description and documentation and also considers both of them complement to each other.

- Preservation of an endangered language and dissemination of the linguistic knowledge to the younger generation also forms a part of the goal of linguistic description and documentation. A dialectological study should also be conducted to determine the mutual intelligibility between the varieties of the language. Although some linguists contest strictly that sociolinguistic enquiries do not constitute the part of linguistic fieldwork.

Collaboration with the native consultants

- There have been many instances where the 'native consultants' showed an enthusiasm for the field of linguistics and were actively involved in the documentation of their own language.
- The earliest recorded participation has been observed in Nahuatl community. The speakers of the community were encouraged by Fray Bernardino de Sahagun (1499-1590) to pen down the oral traditions of their own language.
- The scholars working during the Boas-Sapir-Bloomfield tradition collaborated frequently with their native consultants to conduct linguistic fieldwork and description of their language.
- Kenneth L. Hale (1934-2001) was a pioneer who collaborated frequently with the native speakers of Indian languages. Many such speakers went on to complete their doctoral work under the supervision of Hale.

Choosing a language

Some of the criteria of choosing a language are mentioned below:

- a. If a language is an isolate or belongs to a lesser known family, then it is an ethical duty of a fieldworker to prefer this language on the one which belongs to a well known language family.
- b. A linguist can choose a language if there is a particular aspect which interests him/her.
- c. It is also important to secure permission from the relevant authority once a fieldworker has decided his/her target language for documentation and description. This has to be taken from either from the speech community or the local governing body.

d. The remoteness of the area limits our ease of access and in turn the work also becomes hectic. Fieldworkers nowadays have to worry about the funding which will cover their conveyance and accommodation. Therefore, documentation site should be chosen carefully.

e. If a fieldworker has already worked on a particular speech community then, it is advisable to ask that fieldworker for suggestions.

f. It should also be kept in mind that one's cultural practises, religion, language, or racial background are not demeaned by any community. The fieldworker will have to develop a bit of a thick skin which will help him/her sail through the rough times.

g. It is better to have little background about the language. Ethnologue is helpful in answering some questions about the languages which have been little worked upon.

Field Preparation: Philological, Practical and Psychological

There are three main reasons behind philological preparations:

- a. A fieldworker can learn and improve himself/herself from the mistakes committed by his/her predecessors. But this is only possible in the cases where there is an existing data or texts on a particular language. This reason falls flat on the linguistic fieldwork of languages where there are no written texts.
- b. Philological preparation is important because a researcher might be interested in some aspects of the already captured written texts.
- c. Also at the same time, a researcher can re-elicite the old documents to replenish what is missing in them.

- Before heading to the linguistic field a researcher must be fully aware of the social and political situation of the area. Also the historical knowledge of the area might also come handy during the fieldwork. A deeper understanding of the speakers' socio-economic condition might also give us a hint about the factors which have led dominant languages seep into their speech. In an Indian context, caste system is an important factor which helps us to understand the social hierarchy of a speech community.

Contact Language

- A fieldwork is usually conducted with the help of a contact language. A contact language is a language which acts as a bridge for speakers of different speech communities. Therefore, it is advisable to the researcher to be at least fluent in the contact language because the knowledge of the same will enable him/her to establish verbal contact with the speakers of the target language community.
- Some of the important equipment that a researcher will require for a linguistic fieldwork are:
 - a. minimum two audio recorders
 - b. a laptop
 - c. microphone
 - d. notebook
 - e. digital camera
 - f. rechargeable batteries

Fieldwork Ethics: The Rights and Responsibilities of the Fieldworker

- In the opinion of some of the linguists like Ladefoged (1992) and Dwyer (2006), not everyone maintains a stance that language documentation of endangered languages is an important task for a linguist. This is so because many speech communities which speak less known languages are quite satisfied to learn the dominant language/s of their region for a fruitful future.
- But the number of such believers is quite minute. As mentioned by linguists like Woodbury, Hale, Himmelmann language documentation and description is of imminent need. Also, language revitalization has also enabled the speakers with increased self-esteem. Not only this, the data collected on less known languages helps in studies of language typology.

- As Keren Rice (2006) suggests that the fieldworker should repay the community with which he/she has worked. Linguists must provide their expertise to create teaching materials, discuss language planning policies with local authorities and developing teaching programs. Also it is suggested that the native speakers should be empowered to their own language. In other words, they should be trained in the field of linguistics so that they can analyse their own language. There has been an evidence (Crowley, Penfield et. al) that training a native speaker has been instrumental in the linguistic fieldwork

- The recorded narratives are attributed to the speaker/s and the reproduction of those recordings should always be approved by the speaker. This proposition becomes arduous if the fieldworker lives miles away from the community on which he/she has worked upon. Also, the fieldworker and the translator hold fifty percent each ownership to the translated/ transcribed data. This has resulted from the increase in the importance of cultural and intellectual rights.
- It is also important for the fieldworkers to decide as to how will they share the data and the subsequent analyses with fellow researchers. As we have seen read that the ownership of the data lies with the speaker but the ownership of the products created lies with the researcher. They can be: field recordings, time-aligned annotation of texts, morphological analysis, word lists, cultural notes, pictures, songs and scans of printed materials.

- Also, field materials should be archived properly so that they are accessible to the future generations. The Archive of Indigenous languages of Latin America (AILLA) has maintained different levels of data access. Some of the data is openly accessible while some of it restricted. Some data is only accessible by the speakers of specific speech community.
- One must never argue or impose one's political or religious beliefs over the consultant and his/her community. Dress is another factor which should be adhered to according to the social and traditional norms of the speech community. One must also not interfere into the local proceedings of the community and must not act as if he/she can impart his/her opinion on every matter. It is expected from the researchers to read about the experiences of other fieldworkers to gather more information about one's conduct in the field. It is the duty of a researcher to work with responsibility and respect which will ultimately be reflected in his/her work.

Outcomes of documentation

- One of the main outcome of a linguistic fieldwork is the development of various materials which may help in the revival of the language.

These materials can be:

- a. Teaching materials
- b. Grammar of the language
- c. Dictionary of the language

Native Speakers and Fieldworkers

- One of the key factor for a reliable data is to find a reliable native speaker. But the task to do so may vary from situations to situations. Abbi suggests that one way to find the native speakers is through the officials posted at that location.
- If there has been a fieldwork in the same speech community recently then it might be easier for the fieldworker as he/she can ask the previous researcher/s about their consultants.
- Along with this it is also necessary not to depend on a single consultant as they might prove to be reliable at all. Therefore the fieldworker must contact at least three to four people to make them his/her consultant. In that case, even if one proves to be not helping with the task, then the others have already been prepped for the job. Also, having multiple consultants also enables us to verify the information given from a single source. In choosing multiple consultants, it is always better to select consultants of varying age and gender. This approach is said to have enriched the data of the researchers.

- During a field session, a fieldworker must both record the data digitally and also transcribe the answers in the elicitation response. Field notes should be recorded in the notebooks which help in the future analysis of the data. Each notebook should be numbered and the relevant topic title should be written on the cover of the notebook. The notebooks may also be categorized on the basis of the subtype, for example, intonation, affixes, word order etcetera. The transcribed lines should be numbered for future references. Also after transcribing a sentence or a clause, one must keep a space below the transcription. This space may be later filled with the translation of the above written sentence or a clause. If the fieldworker is handling more than one grammar topics in one session, then he/she should write the topic title on the top most position of the page.

Issues with data collection and analysis

| | Collection | Analysis |
|------------------------------|---|---|
| Result | Corpus of utterances Notes on observation and comments by speaker or compiler on a particular form or construction | Descriptive statements illustrated by one or two examples |
| Procedure | Participant observation, elicitation, recordings; Transcription and translation of primary data | Phonetic, phonological, morphosyntactic and semantic analyses (instrumental measurements, distributional tests etc.) |
| Methodological issues | Sampling, reliability, naturalness | Definition of terms and levels, justification (adequacy) of analysis. |

Recording Metadata

- There should be a list which contains information about each and every information of a particular file. This is called metadata which should be updated after each and every session. Following are the entries which should be present in a metadata:
 - i. Date, time and place of meeting
 - ii. Name of the person who did the recording
 - Topic of recording
 - iv. Persons being recorded
 - v. List of equipment used
 - vi. Indication whether permission was obtained to carry out the session
 - vii. Cross reference to transcriptions
 - viii. Sampling rate and format

Archiving

- Austin (2006:100) explains digital archiving as “ the preparation of the recorded captured data, metadata and proceeded analysis so that the information it contains is maximally informative and explicitly expressed. “
- A digital archive is the need of the hour because with this the fieldworker can also make the data available to the native speakers and other linguists at a similar platform. Also, it also helps other linguists to check the truth in the claims made by the fieldworker. There can be archives which deal with single language and there are also archives where depositors can deposit data from various languages.

Tools for archiving

- Some of the speech transcription programs used by linguists are PRAAT or TRANSCRIBER.
- ELAN and EXMARaLDA are helpful for time aligning speech with annotations.
- FLEx and TOOLBOX are important in creating lexical databases and annotating interlinear texts.

References- Endangered Languages:

- Austin, Peter and Julia Sallabank (eds). 2011. The Cambridge Handbook of Endangered Languages. CUP
- Campbell, Lyle and Anna Belew. 2018. Cataloguing the world's endangered languages. Routledge.
- Grenoble, Lenore A and Lindsay J Whaley. 1998. Endangered languages: Language loss and community response. CUP
- Hale Ken. 1992. Language endangerment and the human value of linguistic diversity. LANGUAGE, VOLUME 68, NUMBER 1.
- Moseley, Christopher (ed.). 2010. Atlas of the World's Languages in Danger, 3rd edn. Paris, UNESCO Publishing. Online version: <http://www.unesco.org/culture/en/endangeredlanguages/atlas>
- Thomason, Sarah G. 2015. Endangered Languages: An Introduction. CUP
- Sallabank, Julia. 2013. Attitudes to endangered languages: Identities and Policies. CUP.

References – Language Documentation and archiving

- ABBI, ANVITA. 2001. *A Manual of Linguistic Fieldwork and Indian language structures*. Munich: Lincom Europa.
- AUSTIN, PETER. 2010. *Language Documentation and Description*. Vol 7. (Special Lecture Series in Language Documentation and Description). London: School of Oriental and African Studies.
- CHELLIAH, SHOBHANA L., DE REUSE, WILLEM J. 2011. *Handbook of Descriptive Linguistic Fieldwork*. Dordrecht: Springer.
- DIXON, R. M. W. 2007. Field Linguistics: a minor manual. *Sprachtypologie und Universalienforschung* (Focus on Linguistic Fieldwork, ed. By Alexandra Y. Aikhenvald) 60(1): 12-31
- DWYER, ARIENNE. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In *Essentials of Language Documentation*, ed. By Jost Gippert, Nikolaus P. Himmelmann, and Ulrike Mosel, 31-66. (Trends in Linguistics. Studies and Monographs 178). Berlin and New York: Mouton de Gruyter.
- LADEFOGED, PETER. 1992. Another view of endangered languages. *Language* 68 (4) : 809-11
- LEHMANN, CHRISTIAN. 1999. Documentation of endangered languages. A priority task for linguists. *Assidue, Arbeitspapiere des Seminars für Sprachwissenschaft der Universität Erfurt* Nr.1. Erfurt: Seminar für Sprachwissenschaft, Philosophische Fakultät, Universität Erfurt.
- NIKOLAUS P. HIMMELMANN. 1998. Documentary and Descriptive linguistics. In *Linguistics* 36. 161-195
- RICE, KEREN. 2006. Let the languages tell its story? The role of linguistic theory in writing grammars. In *Catching language: The Standing Challenge of Grammar Writing*, edited by Felix K. Ameka, Alan Dench, and Nicholas Evans, 235-268. (Trends in Linguistics. Studies and Monographs 167.) Berlin: Mouton De Gruyter.
- WOODBURY, ANTHONY C. 2003. Defining documentary linguistics. In *Language Documentation and Description*, Vol 4, ed. By Peter K Austin, 120-135. London: School of Oriental and African Studies.