

FM – Phonology

Preliminaries to data collection with language consultants

Lolke van der Veen – Pither Medjo Mvé

Outline

1. Choosing a language
2. Before going off to the field
3. Choosing your consultant(s)
4. Money makes the world go round? The benefit of transparency
5. Filling out the consultant's personal file (i.e. ethnolinguistic profile)
6. What kinds of material to collect
7. Conducting the sessions
8. Handling the notebook and the diary
9. Recording and archiving the data
10. Handling the “aftersession”

1. Choosing a language

- Languages poorly studied or not yet described
 - Enormous amount of endangered languages
 - Urgency of the task: establishing a priority (criteria?)
 - Number of researchers available
 - Funding agencies, foundations supporting this kind of work
 - Funding may sometimes determine the choice of the language
 - Submitting well-defined projects
- Languages already described, but...
 - You may want to focus on specific topics
- Same basic approach in both cases, but different strategies
- Sometimes, the demand may come from a specific language community or from governmental institutions
- **Fundamental issues:**
 - What will be the contribution to:
 - Linguistic theory?
 - Linguistic typology?

- Documentation of the world's endangered languages?
- What are the wishes and desires of the soliciting community / institution?

- Detailed preliminary inquiry/research is required:
 - Look for documentation available and make up a detailed list of references (Web, libraries, etc.)
 - Crucial stage! Accumulating knowledge!
 - If possible, prepare an overview of topics studied
 - Read as much as possible about these topics
 - Classification issues, typological features, areal features, etc.
 - Issues in phonology, morphology, syntax, etc.
 - If possible, get in touch with people who have been working on the language/these topics
 - Get information about the “shape” (vitality) of the language
 - UNESCO criteria, Krauss's criteria
 - Number of (fluent) speakers (*difficult issue!*)
 - Geographic location (access, sanitary conditions, etc.)
 - Distribution in space (grouped, scattered)
 - Geographically close languages
 - Linguistically close languages (typology and/or descent)
 - Linguistic affiliation and main features of the family
 - Social organization of the ethnolinguistic community, its economical and political situation
 - Multilingualism?
 - Major cultural features:
 - Kinship system, mating strategies, main modes of subsistence, beliefs, cosmology, etc.
 - Traditions and structures to be respected, possibly shocking or rude behavior to avoid
 - www.ethnologue.com: *Ethnologue, the Languages of the World*
 - Generally, a useful site
 - *Summer Institute of Linguistics*
 - Much information available (regular updates)
 - Online database that can be searched in different ways
 - Gaps, overestimations, underestimations, errors
 - Excellent tool for a first approach, but always stay critical!

2. Before going off to the field

- Research permits (if necessary)

- Visa (if necessary)
 - N.B. Obtaining visa may be conditioned by research permit
 - Administrative steps take (a lot of) time

- Whenever possible, written invitation from official institution(s)

- If possible, inform local authorities

- Get local administrative authorities informed (respect, personal security)
- Do not forget the local, traditional authorities
- Always do this after arrival!

- Choose best season
 - Climate and travel conditions
 - Availability of speakers (consultants in particular)

- See a medical specialist for prophylaxy, etc.

- Purchase and prepare equipment

3. Choosing your consultant(s)

- **Always observe local hierarchies**
 - Check their willingness to let you choose
 - Sometimes, impossible to choose yourself!
 - In that case, comply with the judgments of the elders of the community

- If local authorities let you choose...
 - Take your time
 - Observe
 - Do a little testing
 - Experience and intuition

- Normally, work with several assistants (different age, different sex, different villages if useful, etc.)

- Choose one or two main, regular consultants
 - Regular data collection (main corpus)

- Other, additional consultants
 - Checking the data
 - Measuring representativeness
 - Linguistic variation is a fact that should never be ignored or concealed
 - Good descriptions never ignore variation but take it into account

- Linguistic competence of main consultant(s)
 - Knowledge and use of the language
 - N.B. Several factors are to be taken into account: categories of speakers
 - Fluent speakers, semi-speakers, etc.
 - Grammatical competence vs. communicative competence: good knowledge of grammar, but bad story teller ; etc.

- How did (s)he learn his/her language
 - Check the nature of conditions
- His/her age
 - Minor criterion, in some cases
- Linguistic competence of main consultant(s) (continued)
 - Time spent within the language community
 - If relevant, time spent outside the language community
 - Profession and training
 - You may want to avoid linguists! (☺ ☺ ☺)
 - Male/female?
 - For acoustics, male voices may be preferred! But there are many more criteria that may motivate your choice...
 - Working with the opposite sex may sometimes be a source of misunderstandings
- Linguistic competence of main consultant(s) (continued)
 - Other languages spoken
 - Local vernacular languages, other languages
 - Multilingualism and its possible consequences!
 - Speakers may mix up closely related language varieties in their heads

4. Compensation

- Aim for maximum transparency
 - A written agreement may be useful (check local customs)
 - Agree on the number and the duration of the sessions
 - Agree on the kind of activities
 - Agree on nature and compensation (money, goods, assistance)
- Choose an appropriate compensation
 - Comply with local customs and local conditions (rate, etc.)
 - Money? Observe local standards
 - Goods? Observe local needs, assistant's needs
- Keeping a register
 - Have it signed at the end of each working session

5. Consultant's personal file (i.e. ethnolinguistic profile)

- Name(s) (culturally determined, sometimes several (competing) systems)
- Sex
- Date and place of birth
- Ethnic affiliation, clan, lineage, sublineage (etc.)
- Mother tongue (not always the adequate term!)
- Other languages spoken and extent of knowledge and practice (vernacular and other languages)
- Time spent in the language community
- If relevant, time spent outside the community
- Major geographical changes and time spent in other places
- Conditions of language acquisition
- Profession(s)
- Position and role within the community

- Name of **father** (biological, social)
- Date and place of birth
- Language(s) spoken
- Ethnic affiliation, clan, lineage, etc.
- Profession(s)
- Social position

- Name of **mother** (biological, social)
- Date and place of birth
- Language(s) spoken
- Ethnic affiliation, clan, lineage, etc.
- Profession(s)
- Social position

6. Kinds of material to collect

- Theoretically, two options:
 - (1) Written sources
 - Written languages
 - Collect sufficient and good-quality documents
 - Published grammars, material produced by local writers
 - (2) Oral sources
 - Data collected with the help of one or several language consultants

- Concerns 60% of the languages of the world

- In many cases, therefore, **option 2** is the only viable one
 - Not merely the so-called “primitive” or “exotic” languages
 - E.g. Oral languages (“patois”) in France
 - What is a “primitive” language? (eurocentric view)
 - What is an “exotic” language? (eurocentric view)
 - Even oral French!
 - Try to find a grammar of oral, spoken French! ...

- As for option 2:
 - Data collected by (other) fieldworkers
 - Data to be collected (in the field) by means of rigorous and systemic linguistic and/or ethnolinguistic fieldwork
 - Special training is required for data collection procedures
 - Mixture of technique and art/creativity: discovery procedure, experience and intuition
 - Discovery procedure must be made explicit

- **Words or more natural type of linguistic productions?**

- Ideally, natural productions would be the best! Long stretches of spontaneous speech... Dialogues...

- But this, of course, is not feasible right from the start...
 - Too long, too much new information, delimitation (segmentation) problems, too many contexts, etc.

- For a first sketch of the phonology and the basic morphology of the language, use **word lists** (i.e. lists of glosses)
 - Several word lists available: e.g. Swadesh, Greenberg, etc.
 - Most of these lists are supposed to contain so-called “core vocabulary” (presumably culturally neutral)
 - Choose the most suitable one(s) for the geographical and cultural environment of the language to be studied
 - Useful tools to start with, but with more or less severe limitations

- Morris Swadesh (*Swedish linguist active in the 1950s*)
 - Lists comprising 100 or 200 words (multiples of 100): core vocabulary (resistance to borrowing)

- List composed by Joseph Greenberg (*†2005*)
 - 600 words
 - List devised for work on (Black) African Languages

- Lists elaborated by Thomas & Bouquiaux (Africa)

- Other lists, for other parts of the world
 - Contact local researchers

- Look on the Internet
- Get familiar with them
- Also lists for eliciting cultural (specialized) vocabulary
 - Kinship, flora, fauna, technology, art, medicine, etc.
 - To be used later on
- A word on **core vocabulary**
- Core vocabulary is supposed to comprise:
 - *Body parts*
 - *Animals (easily perceived or domesticated)*
 - *Parts of animals used by man*
 - *Kinship (basic)*
 - *Physical entities present in the immediate surroundings*
 - *Numerals*
 - *Basic actions: movement, perception, etc.*
 - *States and basic properties*
 - *Pronouns or pronominal forms*
 - *Adverbs (space, time)*
- **Why core vocabulary?**
 - Culturally neutral?
 - Shared conceptual material?
 - More resistant to borrowing, most stable part of a language's lexicon
 - Frequency of use (which probably contributes to the preceding point)
 - Not always the same words that are most frequently used from one language to another!
- **Potential pitfalls while working with word lists**
 - Try not to collect the data from just one speaker (representativeness)
 - Listen very carefully: new sounds, new combinations of sounds, suprasegmental information
 - Different contexts, different results, possibly: with or without determiners, singular/plural, with or without presentatives
 - Does the assistant understand the gloss/entry?
 - Local French, local English: more or less different meanings!
 - E.g. “rosée” in Gabon: ‘quite heavy rain’ vs. France: ‘dew’
 - Polysemic glosses
 - E.g. “homme” (French): “human being”, “male individual”, “husband”
 - Direct equivalents may not exist: e.g. “hand”, “arm”, “leg”, “foot”, “finger” (reality being sliced up in different ways, in different cultures)
 - Not necessarily one-to-one matches
 - One gloss may produce several words
 - One word may cover several glosses
 - Homonymy

- E.g. “voler” in French: ‘to fly’ and ‘to steal’
 - When working on several related languages at the same time: cases of semantic shift, regional variations, possible impact of multilingualism
 - E.g. “kidneys” – “liver” – “heart” – “lungs”; “chest” – “chest bone”
 - So, always carefully check whether the gloss has been correctly understood
 - When using lists that do not group glosses according to some semantic principle: try to introduce some kind of grouping yourself
 - Facilitates your consultant’s job...
- A word on **cultural (specialized) vocabulary**
 - Flora and fauna; sociopolitical organization; kinship; technologies (*hunting, collecting, fishing, basketry, pottery, etc.*); beliefs, religion, cosmology; diseases and medical treatment; Arts (*music, dance, sculpture, etc.*); etc., etc.
 - Types of vocabulary which may reveal signs of contact between languages (substratum, adstratum, superstratum) and/or signs of diffusion
 - Cultural perspectives (Anthropological Linguistics)
- Once you have a good idea of the lexical phonology, you should start collecting and analyzing larger stretches of speech
 - Certain phonological phenomena (processes, constraints) only emerge at the postlexical level, i.e. beyond the word level
 - Certain types of allophonic variation
 - Tone-related phenomena
 - Intonation
 - Etc.
- And last, but not least ... a hopefully **useful advice!**
- Even if your intention is to gather data for phonological analysis, try to get an idea as precise as possible of the **basic morphology** right from the start
 - K. Pike’s position: no phonological analysis without at least some knowledge of the morphosyntax of the language
 - Though the entities of morphology and phonology are rarely congruent, they often interact
 - Such interactions may give rise to variations (alternations)
 - E.g. for Bantu: the basic structure of the noun is nominal prefix followed by noun stem
 - Example of frequent stem structure: (Px) + C₁V₁C₂V₂
 - Segments attested in each of these position: restrictions, constraints

7. Conducting the sessions

- **Be aware of your role as inquirer and reflect on possible consequences**
 - Your personality, character, image, ways of communicating, own cultural habits, etc.
 - Inquirer is a participant to whom people react
 - ‘Accommodation theory’ (*Giles 1994*): tendency to adapt one’s linguistic productions according to the person one is talking to...
 - ‘Observer’s paradox’ (*Labov 1972*): people do not behave naturally when they feel being observed...
- **Linguistic fieldwork is intrusive...**
- The language consultant(s) is/are just ordinary people, so let’s treat and respect him/her/them as such
 - They can be happy, sad, annoyed
 - They can be tired, bored
 - In short, they have their limits, they are no machines
- The consultant is someone who probably grew up in a different culture, so mind:
 - The kind of questions you ask...
 - The way you get along with the opposite sex, with elderly people...
 - The way you react to presumed errors, long digressions, being late, slow progression, etc.
- A **well-prepared protocol** is essential, but it should leave some flexibility (“adaptability”) for improvisation and (limited) digressions
 - Interesting information may emerge thanks to (short) digressions
- Working sessions
 - **Be directive, in a reasonable way**
 - Linguist should remain in charge and direct the session (which may sometimes be rather difficult!)
 - Leave some room for semi-directiveness
 - Spontaneous remarks and comments from the consultant(s) are often very instructive
 - How much time per session, per consultant?
 - One hour maximum, as a rule
 - Possibility to have two or more sessions per day
 - If longer sessions, plan pauses
 - Take into account age and obligations related to profession, family, etc.
 - What a regular session should look like
 - Partially depends on kind of elicitation
 - It should be well prepared and well planned
 - Items to be checked
 - Questions to be asked
 - Number and nature of items to be collected
 - Try to vary the activities as much as possible
 - Do not have your consultant repeat the same item more than three times

- You can make a recording!
- Keep track of all interesting remarks and phenomena that come up during the session
 - Even material not directly related to the objectives of the work (within reasonable limits, of course)

8. Handling notebook and diary

- What you most basically need: a **notebook**, a **diary**, **pens** with different colour ink
- The notebook
 - Avoid spiral notebooks
 - Left pages for comments, remarks, ideas, and additional data
 - Right pages for data in (narrow) phonetic transcription
 - Write in ink, not with pencil
 - Each modification, even the slightest, is to be carefully dated
 - Cross out, never erase!
 - Transcription: segmental level, suprasegmental level
 - Always leave enough space between the items for possible additional information (e.g. *special plural form, variants*)
 - Suprasegmentals: prefer iconic representation for melody (F_0), especially in the beginning
- The diary
 - Relevant information about the consultant(s)
 - Relevant information about the language and the cultural practices
 - Keeping track of progress made, session after session, day after day
 - Emerging questions and (working) hypotheses
 - Working out new questions
 - Data to be checked
 - Detailed plan for the next session
- Keep your notebook and diary up-to-date
 - An every-day exercise
 - Your task
 - Presence of consultant(s) is not required
 - Transfer observations, questions, ideas and hypotheses to the diary
 - Prepare the next session
 - Checklists
 - Questions to ask
 - Number and nature of items to collect

9. Recording and archiving

- It is crucial to make good quality audio recordings of all types of material collected
 - Use high-standard equipment (*do not forget the microphone!*)
 - The quality of the recordings is crucial for future utilizations (analysis, archiving)
 - Modern recording devices allow to obtain directly the “.wav” format that can be used for PRAAT (e.g.)
 - Get acquainted with the equipment and the techniques
 - Find the best possible conditions for recording
 - Use headphones to test background noise
 - Label each recording (*name of assistant, place and date, type of data, recording protocol, future utilizations*)
 - Plan recording sessions regularly
 - Keep copies of the recording in several places (PC, external hard drive, international archiving centres)
 - Audiovisual recordings may be preferred in certain cases, especially for documenting specific linguistic practices

10. Handling the “aftersession”

- Keeping your notebook and diary up-to-date, on a daily basis
 - See above section on “Handling the notebook and the diary”
- You may want to use traditional (paper) files
 - Very laborious
 - Several files for one entry!
 - Quite old-fashioned...
- A much better option: a computer database
 - Several computer programmes available allowing to sort out data very rapidly and in a straightforward manner
 - Shoebox/Toolbox (© SIL)
 - Filemaker®
 - Excel
 - Etc.
 - Searches
 - Complex searches, viz. cooccurences between segments
 - Statistical information (*frequency of specific features, combinations*)
 - Linguistic phenomena are (almost) never completely regular!
- Analyzing the data
 - Ultimate goal: analyze and describe the structure and the function of the linguistic entities, and compare the results with what is already known about languages and language

- Activity that is carried out in **several stages**, where one continuously navigates between the data and the description
- Time-consuming activity
 - Patience
 - Perseverance
 - Expertise
 - Experience
 - Intuition
- Work out **schemes** on the basis of recurrent structures in the data of the language
- Study of (different types of) data gradually leads to **hypotheses** about the language as a system (i.e. language is a theoretical construction)
 - Sound structure (syllables, words), sound inventories
 - Phonology: segments, suprasegmentals, distinctive feature, rules or constraints
 - Morphology (verb, noun, etc.)
 - Syntax
 - Lexicon and lexical semantics
 - Pragmatics

Suggestions for further reading

- Bowern, Claire. (2008). *Linguistic fieldwork: a practical guide*. Hampshire: Palgrave Macmillan.
[Particularly recommended.]
- Crowley, Terry. (2007). *Field linguistics: a beginner's guide*. Oxford: Oxford University Press.
- Dobrin, Lise M. (2008). From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84.300-24.
- Newman, Paul and Martha Ratcliff (eds) (2001). *Linguistic fieldwork*. Cambridge: Cambridge University Press.
- Payne, Thomas E. (2006). *Exploring Language Structure. A Student's Guide*. Cambridge: Cambridge University Press.
- Payne, Thomas E. (1997). *Describing morphosyntax*. Cambridge: Cambridge University Press.
- Vaux, Bert and Cooper, Justin (1999). *Introduction to Linguistic Field Methods*. München: Lincom Europa.