

1 Opening thoughts

1.1 Goals

1. To learn about the basic principles and techniques of (language) documentation that can be applied to any language
2. How to plan a documentation project
3. To practice applying these techniques in a language you have no previous experience with

1.2 First, we need to think about...

- ... the difference between *documentation* and *description*¹
- ... the collaborative nature of documentation and/or description: you are an (emerging) expert in what you do; your collaborator or consultant is an expert in their language and culture
- ... the connection between ‘what’, ‘why’ and ‘who’
- ... the relationships you are creating

2 Guiding principles to (language) documentation

1. Accuracy

- You are creating a permanent record of your language and culture
- Future language learners, community linguists, educators, researchers will be relying on your work
- The more accurate you are the less time future language workers will spend correcting mistakes or clarifying data

2. Detail

- As much information as *practically* possible – you can’t do everything!
- Best *possible* recordings (text, audio, video) given the environment and equipment
- Attention to accuracy and detail is a show of respect, and instills confidence in your work

3. Consistency

- Once you have a system (transcription, recording, etc.) or way of doing things, be consistent in using it

¹Furbee, L. 2010. “Language Documentation: Theory and Practice.” In Lenore A. Grenoble and N. Louanna Furbee, Editors. *Language Documentation: Practice and Values*. Pp. 3–24. Amsterdam and Philadelphia: John Benjamins

- Consistency = efficiency: if you have a consistent, organized system, it's much easier and quicker to find things

4. Planning and goals

- Know your goals and take the time to plan how you'll reach them – *before* you sit down to do any work
- Be realistic: it's better to have smaller, more manageable projects than grand, complicated ones – this maximizes the chances of success, which in turn gives you confidence for your next project!

3 Organizing your thoughts, ideas, and plans

1. Questions to ask yourself

- What kind of project am I undertaking?
- What are the outcomes?
- What are the steps I need to take to implement it?
- How is the community involved?
- Who is this for?
- What kind of media? (i.e. text, audio, video, photo)
- What is my skill/experience/preparation level?
- Do I have or need permission/IRB and how will I obtain it?
- Timeline, funding and budget?
- Are there any ethical issues (privacy, culturally-specific sensitivities)?

2. Project design

- *Pre and post session tasks*: planning the location and setting where you will be doing your fieldwork, checking/analyzing your data after the session and planning for the next one based on this.
- *Making plans with language consultants (hours, frequency)*: planning how long your sessions will be (usually 2 hours max, plus breaks), making them feel as comfortable as possible.
- *Equipment*: what kind of equipment do you have? What kind do you need in order to achieve your outcomes?
- *Archiving and sharing*: where will your recordings and transcriptions be archived? Who will have access? Making plans to give your language consultants copies of your work together. What kind of media do they prefer?
- *Elicitation plan*: What kind of language do you plan to elicit, and how can you best prepare in advance of your session? (See below)
- *Educational/community resources*: coordinating your plans with language teachers, or thinking about how your language data can be used in an educational setting.

2. Resources

- *Hardware*: transcription notebooks, computer, smartphone, personal recording device, microphones, cameras, USB drives, etc.
- *Software*: digitization and editing software, linguistic and metalinguistic databases, etc.
- *Other tools*: storyboards for elicitation, objects or photos etc. for eliciting the names of things
- *Money*: to pay your language consultants and purchase your hardware and software

4 Types of documentation

4.1 Language

1. Elicitation

- a. Direct elicitation: *How do you say X?*
 - * For objects and sentences that ‘do not change’
 - * Usually the easiest and quickest ways to get language data, and the best way to get familiar with a language
 - * The beginnings of a dictionary
- b. Testing judgments: *Can I say Y (in this context)?*
 - * Can be used if you already know about the language
 - * Necessary for things that can’t easily be translated or directly elicited:
The book is on the table vs. *A book is on the table*

2. Observation

- a. Prompting
 - * The “Show me” method
 - Ideal for documenting all of the steps in how things are made or done: *Can you show me how to make soup?*
 - Does not require a lot of prep, but it is a lot of work in analyzing!
 - Storyboards: reduces the effect of English in an elicitation setting
 - * The “Tell me” method
 - Eliciting personal or traditional stories: *Can you tell me a story from your childhood?*
 - Also storyboards
 - Does not require a lot of prep, but it is a lot of work in analyzing!
 - Take your linguist’s hat off and be a listener!
- b. Passive observing
 - * Recording conversations
 - Supply a topic to two or more speakers and see what happens

- Little prep needed: a great way to observe spontaneous, natural language use and different kinds of language you can't elicit
- Possibly opportunities to interject with questions about the language
- Actually challenging because you can end up with way more data than you can actually practically use
- * “Fly on the wall”
 - Almost no prep needed: used to observe cultural practices where asking questions is maybe not possible or appropriate
 - Language art (i.e. song and poetry)
 - Take your linguist's hat off and be a listener!

4.2 Cultural language

- Can be a much more ‘open-ended’ methodology, often involving the documentation and use words and phrases
- Much of the actual description of cultural language is done in the ‘research’ language (i.e. English)
- Includes not only a description of the language, but also the *use* of language, often conditioned by the social context
- Both *elicitation* and *observation* are used

5 Types of media

- With language elicitation and/or observation there are (at least) four kinds of media:
 1. *Text*: this is the bare minimum. Language documentation is about creating a linguistic record, which almost always means writing things down.
 2. *Audio*: Also the minimum. Language is sound!
 3. *Video*: think about why you want to record video: how does it contribute to what you are documenting?
 4. *Photo*: not an obvious choice, but given today's technology (i.e. everyone has a camera in their phone) it is a very easy way to add another dimension to the language record
 5. *Other*: botanical samples, art

6 Best practices in documentation

- The value of teamwork: minimally, you and your consultant are a team – you work together towards reaching a common goal:
 - The meaning of a word that balances a suitable linguistic (i.e. technical) description with the descriptions your speaker offers

- The interpretation of a phrase or sentence and the deeper meanings it conveys
- Team elicitation: working in pairs can be very efficient – one can concentrate on elicitation and interaction while the other transcribes

6.1 Ideals

- **The minimum:**
 - A notebook and a pencil
- **The ideal:**
 - A notebook and a pencil
 - A way to make audio recordings
- **The deluxe ideal:**
 - A notebook and a pencil
 - A way to make audio recordings
 - A way to make video recordings or take photographs
- **Working with what you have:** *it's not necessary to have all of the latest recording gadgets or the most expensive equipment or the most complicated software!*
 - Notebooks and pencils never break down
 - If your equipment is reliable and makes decent recordings, then this is what you need

6.2 Field notes for elicitation

- Preferred: a book with binding – no looseleaf or cards etc.
- Pencil, not pen
- Formatting:
 - Page numbering
 - Example numbering
 - Language data: either a three or four line format, plus the context and speaker comments and your notes
 1. Orthography
 2. Morphological break down (parts of words)
 3. Gloss
 4. Translation
 - Speaker
 - Date, location, time
 - Timestamp
 - Optional: only use one side of the page

6.3 Elicitation plan

- If using the elicitation method, it's best to have all of your questions entered into your field notebook in advance of the session – this is actually a part of planning your elicitation!
- If using the observation method, you should make a record of the topic(s) (if you know them in advance) and any other details of the setting

6.4 Audio recording

- **The minimum:**
 - A device that reliably records audio: a laptop, smartphone, personal audio recorder, iPad, tape recorder, etc.
- **The ideal:**
 - A device that makes high quality digital recordings (a laptop, most smartphones, iPad/Phone)
 - An quality microphone (in order of ideal: headset, lapel, desktop mic on a stand, internal device mic)
 - A quiet, controlled environment
- **The deluxe ideal:**
 - A device dedicated to making high quality digital recordings (laptop, digital recorder, DAT)
 - A headset condenser mic
- Understanding how your recorder works
 - Do you know how to start and stop the recording?
 - Do you know where the files are on your computer/smartphone/iPad/Phone?
 - Do you know how to rename them?
 - Do you know how to copy or move them?
- Goals
 - Are you recording stories for a collection or for the language consultant's personal use?
 - Are you recording word lists for a talking dictionary or for a pronunciation guide?

7 A Direct Elicitation session

- Basic setup:
 - Recording device ideally on a stand (to prevent any knocking or vibrations in the table)

- Check for sufficient memory (minimum two hours)
- Test levels: the golden rule for digital audio recording: *never go in the red!*
- The mic (if internal) should be pointed towards the speaker at a distance of around 3 feet (but this can vary)
- Make sure you and your language consultant is comfortable and has water
- Direct Elicitation using two methods: basic word lists – ideal if you know very little about the language

7.1 Direct Elicitation Method 1: *Through recording*

- The basic principles of *through recording*:
 - The entire session is recorded, breaks, starts, stops, and all resulting in one file for the session
 - Advantages: no details are missed – potentially interesting and spontaneous discussions, elaborations and speaker reflections are captured (also real-time corrections)
 - Possible disadvantages: recording memory intensive, and it can be hard to locate specific things later
- 1. Start recorder: introduce yourself, your language consultant, the date and time, and location
- 2. Record this same information in your field notes
- 3. Basic questions: *How do you say X?* or *What do you call (a) Y?*

7.2 Direct Elicitation Method 2: *Rehearsed recording*

- The basic principles of *rehearsed recording*:
 - The session is broken up into parts which are first practiced and then recorded
 - Advantages: organization, which makes for efficient memory use, retrieval, cataloging, and editing
 - Possible disadvantages: effects of rehearsed speech (it can sound clear, but unnatural)
- 1. Start recorder: introduce yourself, your language consultant, the date and time, and location
- 2. Record this same information in your field notes
- 3. Stop recorder and work through the following steps:
 - a. Take three or four words or sentences
 - b. Elicit these from the speaker and transcribe them
 - c. Give the speaker a chance to practice them
 - d. Restart the recorder and ask the speaker each of the words, pronouncing each one **three times** (the magic number in direct elicitation)
 - e. Stop the recorder and repeat

8 An Observation Prompting session

- Although you could use either *through* or *rehearsed recording*, the latter works best
- Storyboards: an effective way of prompting naturalistic language with minimal interference
- 1. Start recorder: introduce yourself, your language consultant, the date and time, and location
- 2. Record this same information in your field notes
- 3. Stop recorder and work through the following steps:
 - a. Chose a storyboard (either printed out or displayed as slides on a computer/iPad)
 - b. Run through the story in English first (planning the story)
 - c. Give the speaker a chance to rehearse the story in the language (maybe even a couple of times)
 - d. Restart the recorder and ask the speaker to tell the story
 - e. Stop the recorder and repeat with another storyboard
- You may chose to do the language transcription later while listening to the recordings rather than on the spot

9 Post-session

- Transfer audio files to your computer
- Always review your recordings for quality or and malfunctions **as soon as possible** (i.e. the same day)
- Check your transcriptions and make the appropriate corrections and clarifications
- Fill in the timestamps for easy retrieval
- Identify questions for your next elicitation session
- Analysis:
 - There are a number of very useful annotation tools that you can use to analyze and enrich your descriptions
 - *Notability* is a note-taking tool that can align audio with text in realtime

10 Sharing and archiving

1. Community

- How are they directly or indirectly supporting or guiding your work?
- How will your work complement other efforts in the community (i.e. cultural, educational)
- How will you present and/or share your work?
- Are there any potentially sensitive issues?

2. Educational resources

- Possibly part of the planning of your project: producing video clips to use in a classroom
- A dictionary or grammar text that is accessible to language learners
- Publications that can be used in a classroom
- Considering what language teachers need

3. Archiving

- What is the archival media?
- Where is it being stored permanently and who has access to it?
- Issues of duplicating and citing work

4. Accountability

- Often, if you have a grant you must make arrangements for your notes and recordings to be archived or deposited with the organization that financially supports your work
- Good record keeping
- A successful project can lead to more money for future projects!

11 Three activities (for Day 2)

1. The Swadesh list (1972): a classic compilation of basic concepts for the purposes of historical-comparative linguistics

- Chose one of the two Direct Elicitation methods above and elicit as many words as you can from the Swadesh list
- *Tip:* Instead of using the (slightly boring and pedantic) *How do you say X?* or *What do you call (a) Y?* questions, try taking the word and putting it into a sentence
- See if you can discover anything about the phonemes or morphology in Gitksan

2. Eliciting a storyboard (a little more adventurous at this stage)

- There is a very nice selection of storyboards for language elicitation at totemfieldstoryboards.org (or you can draw your own or adapt pictures)
- Four storyboards are included in this package

3. What is the difference between a *yukw* and *liliget*?

- What are relevant words and phrases that are involved?
- Describe the arrangement and activities of each
- What is *gluk*
- What is the importance of song in these activities?

12 Working with previously recorded/legacy materials

- Sometimes you don't have time to ask detailed questions about a narrative or story you collected previously
- Often there are 'legacy' recordings that communities or speakers have: these are often on cassette tapes or reels
- These recordings contain a wealth of (historical) material, but recording quality and length can be an issue
- **Develop a fieldwork project using these materials!**
 - Digitize the recording: it's much easier to work with; it also makes the recordings usable with modern annotation software
 - Work with a speaker to re-record the language data
 - Collaborate on transcribing the language
 - Use the data for discovering other topics of interest (linguistic description or cultural activity)
 - Utilize the methodology from above, especially 7.2

Appendix

Common digital recording format specifications (found on almost every digital recording device and software application now):

- .wav PCM format MONO
 - This is the industry standard: it is open source, can be read by any device
 - Does not compress or alter the audio in any way
 - It also takes up the most memory (because it is not compressed)!
 - Only use other formats like .mp3 if you have space issues
- 44.1k (CD quality)
 - This is called the *sampling frequency*
 - The default on most recording equipment and software
 - 48k is higher and becoming more common, but 44.1k is more than sufficient for language recordings (and it takes less memory)
- 16 bit
 - The resolution of the sampling frequency
 - The default, and sufficient for language recordings