

ULAB Conference 2026, *University of York*
8th-10th April

**FIELD METHODS IN ACTION:
DISCOVERING THE STRUCTURE OF AN
UNKNOWN LANGUAGE**

An introduction to language elicitation

Daniel Clayton

PG Researcher



AIMS OF THIS WORKSHOP

During this three-day workshop, you will:

- understand why linguistic fieldwork is necessary;
- learn basic techniques for eliciting language structures from languages you don't speak;
- spend time with a native language consultant who speaks a language you know nothing (or very little) about and collect linguistic data from them;
- analyse and summarise this data to present to the rest of the participants.

WHAT IS FIELD LINGUISTICS?

“*Field linguistics* is primarily a way of obtaining linguistic data and studying linguistic phenomena. It involves two participants: the speaker (or speakers) of a language and the linguistic researcher [or field linguist]. The means of carrying on investigation is the most direct possible, by personal contact. The speaker of the language, *the informant* [or *consultant*], is the source of information and the evaluator of utterances put to him by the investigator.” (Samarin 1967)

WHY CARRY OUT LINGUISTIC DOCUMENTATION?

- Ethnologue estimates 3,000–3,200 of the world’s languages are currently endangered.
- This corresponds to roughly 40–46% of all languages.
- Hundreds of languages remain poorly or entirely undocumented.
- Many endangered languages have little or no descriptive material. The Endangered Languages Project explicitly notes that there are languages “researchers don’t know about” and so the number could be higher than we think.
- But even discounting these undocumented ones, there are likely to be 3,000–5,000 languages that are underdescribed to some degree, or that have received very little attention.

LINGUISTIC HOTSPOTS

According to data triangulated from Ethnologue, UNESCO, and large-scale surveys like Hammarström et al. (Glottolog) and Himmelmann (1998, 2006), the following geographical areas are most in need of attention from linguistic documentation:

- Amazonia
- New Guinea (PNG and West Papua)
- Sub-Saharan Africa

LINGUISTIC HOTSPOT: AMAZONIA (MY AREA)

- The Amazon region is home to some 350 different languages (Aikhenvald 2012: 21)
- Much genetic diversity can be found:
 - According to Aikhenvald, there are six major linguistic families, but many more smaller families and isolates.

Endangerment

- Very high: ~70–80%
- Many languages have: <1,000 speakers, often <100 speakers

LINGUISTIC HOTSPOT: AMAZONIA (MY AREA)

Documentation status

- ~30–50% of Amazonian languages are still undocumented or only minimally documented, and the majority are underdescribed.
- Many languages:
 - have only wordlists or short sketches
 - lack reference grammars
 - lack texts/corpora

Why is this?

- Late and uneven fieldwork (mostly post-1960s)
- Difficult access (geography; politics; permits and permission (especially for Brazil); dangers, e.g. drug cartels)
- Rapid shift to Spanish/Portuguese

WHAT DO FIELDWORK AND DOCUMENTATION INVOLVE?

Fieldwork involves:

- establishing contact with a language community and getting their (and sometimes the authorities' permission) to work on their language;
- “the [eliciting of and] collection of accurate data in an ethical manner” (Bower 2008).

Language documentation involves:

- analysing this data to reach conclusions about how the language works: its phoneme inventory, its phonological rules, its morpho-syntactic structures, its semantics, etc.
- producing some kind of output (e.g. word lists, a dictionary, a sketch, a reference grammar, a learner's grammar, a series of texts in the language, recordings, or a corpus) that is both beneficial to the language community itself and to the wider linguistic community.
- The ultimate goal is to provide a record of the language for the purposes of language description, which can then feed theoretical work, reverse endangerment status, or revitalise languages close to dormancy or those that have previously become dormant.

WHAT ARE YOU GOING TO DO DURING THIS WORKSHOP?

1. Split into groups and each group will work with a different language consultant. Each group will work on a different language.
2. First elicit a basic word list of concrete everyday lexical items (but don't spend too much time on this as vocabulary will amass during the sentence elicitation).
3. Decide as a group what feature you'd like to work on (e.g. verb agreement, tense/aspect, nominal agreement, constituent order, negation, questions, passives, demonstratives, nominal and verbal paradigms, (i.e. person, number, gender inflections), equative clauses, transitives vs intransitives vs ditransitives, morphological case, syllable structure, (sound inventory), etc.).
4. Carry out your elicitation. (I will explain how to do this next.)
5. Find a time to meet up at your convenience (during the conference, in the evening remotely) and collaboratively prepare a short presentation on Google Slides about what you discovered about this language. These will be presented on Friday.

HOW TO CONDUCT YOUR ELICITATION SESSIONS

- STEP 1** Working collaboratively, ask your consultant to translate sentences based around your chosen language feature from English to their native language. Document this either using IPA or your own orthographical system. You may ask the consultant to repeat each sentence as many times as necessary.
- STEP 2** Gradually change each sentence minimally to infer differences in word and sentence forms.
- STEP 3** Use these minimal pairs to make hypotheses about how you think that grammatical feature in the target language works.
- STEP 4** Test your hypothesis with more sentences, (e.g. change the verb, noun, etc.) to confirm or negate your hypothesis. Revise your hypothesis if necessary.

WORKED EXAMPLE 1: PERSON / NUMBER PARADIGMS + BASIC WORD ORDER

- | | |
|---|--------------------------|
| 1. I eat the rice. | 'e amu vau 'i te raiti |
| 2. He eats the rice. | 'e amu 'oia 'i te raiti |
| 3. She eats the rice. | 'e amu 'oia 'i te raiti |
| 4. We (not including you) eat the rice. | 'e amu mātou 'i te raiti |
| 5. We (including you) eat the banana. | 'e amu tātou 'i te mai'a |
| 6. We two (you and me) eat the banana. | 'e amu tāua 'i te mai'a |
| 7. We two (but not you) eat the banana. | 'e amu māua 'i te mai'a |
| 8. I drink the water. | 'e inu vau 'i te pape |

WORKED EXAMPLE 1: PERSON / NUMBER PARADIGMS + BASIC WORD ORDER IN TAHITIAN

- | | |
|---|--------------------------|
| 1. I eat the rice. | 'e amu vau 'i te raiti |
| 2. He eats the rice. | 'e amu 'oia 'i te raiti |
| 3. She eats the rice. | 'e amu 'oia 'i te raiti |
| 4. We (not including you) eat the rice. | 'e amu mātou 'i te raiti |
| 5. We (including you) eat the banana. | 'e amu tātou 'i te mai'a |
| 6. We two (you and me) eat the banana. | 'e amu tāua 'i te mai'a |
| 7. We two (but not you) eat the banana. | 'e amu māua 'i te mai'a |
| 8. I drink the water. | 'e inu vau 'i te pape |

What can you infer from this data?

WORKED EXAMPLE 1: PERSON / NUMBER PARADIGMS + BASIC WORD ORDER IN TAHITIAN

- | | |
|---|---------------------------------|
| 1. I eat the rice. | 'e amu vau 'i te raiti |
| 2. He eats the rice. | 'e amu 'oia 'i te raiti |
| 3. She eats the rice. | 'e amu 'oia 'i te raiti |
| 4. We (not including you) eat the rice. | 'e amu mātou 'i te raiti |
| 5. We (including you) eat the banana. | 'e amu tātou 'i te mai'a |
| 6. We two (you and me) eat the banana. | 'e amu tāua 'i te mai'a |
| 7. We two (but not you) eat the banana. | 'e amu māua 'i te mai'a |
| 8. I drink the water. | 'e inu vau 'i te pape |

What can you infer from this data? **pronoun forms**

WORKED EXAMPLE 1: PERSON / NUMBER PARADIGMS + BASIC WORD ORDER IN TAHITIAN

- | | |
|---|--|
| 1. I eat the rice. | 'e amu vau 'i te raiti |
| 2. He eats the rice. | 'e amu 'oia 'i te raiti |
| 3. She eats the rice. | 'e amu 'oia 'i te raiti |
| 4. We (not including you) eat the rice. | 'e amu mātou 'i te raiti |
| 5. We (including you) eat the banana. | 'e amu tātou 'i te mai'a |
| 6. We two (you and me) eat the banana. | 'e amu tāua 'i te mai'a |
| 7. We two (but not you) eat the banana. | 'e amu māua 'i te mai'a |
| 8. I drink the water. | 'e inu vau 'i te pape |

What can you infer from this data? **pronoun forms**

word order: VSO

WORKED EXAMPLE 1: PERSON / NUMBER PARADIGMS + BASIC WORD ORDER IN TAHITIAN

- | | |
|---|--|
| 1. I eat the rice. | 'e amu vau 'i te raiti |
| 2. He eats the rice. | 'e amu 'oia 'i te raiti |
| 3. She eats the rice. | 'e amu 'oia 'i te raiti |
| 4. We (not including you) eat the rice. | 'e amu mātou 'i te raiti |
| 5. We (including you) eat the banana. | 'e amu tātou 'i te mai'a |
| 6. We two (you and me) eat the banana. | 'e amu tāua 'i te mai'a |
| 7. We two (but not you) eat the banana. | 'e amu māua 'i te mai'a |
| 8. I drink the water. | 'e inu vau 'i te pape |

With more data, you can work out what the sentence initial 'e (aspect marker) means and the nominal-initial 'i and te (object marker and definite article respectively).

WORKED EXAMPLE II: VERBAL AGREEMENT

1. I read a book.
2. You read a book.
3. The boy reads a book.
4. The boys read a book.
5. We (including you) read a book.
6. We (excluding you) read a book.

Che amoñe'ẽ peteĩ aranduka.

Nde remoñe'ẽ peteĩ aranduka.

Mitãkaria'y omoñe'ẽ peteĩ aranduka.

Mitãkaria'ykuéra omoñe'ẽ peteĩ aranduka.

Ñande ñamoñe'ẽ peteĩ aranduka.

Ore romoñe'ẽ peteĩ aranduka.

WORKED EXAMPLE II: VERBAL AGREEMENT IN PARAGUAYAN GUARANÍ

1. I read a book.
2. You read a book.
3. The boy reads a book.
4. The boys read a book.
5. We (including you) read a book.
6. We (excluding you) read a book.

Che amoñe'ẽ peteĩ aranduka.

Nde remoñe'ẽ peteĩ aranduka.

Mitãkaria'y omoñe'ẽ peteĩ aranduka.

Mitãkaria'ykuéra omoñe'ẽ peteĩ aranduka.

Ñande ñamoñe'ẽ peteĩ aranduka.

Ore romoñe'ẽ peteĩ aranduka.

What can you infer from this data?

WORKED EXAMPLE II: VERBAL AGREEMENT IN PARAGUAYAN GUARANÍ

1. I read a book.
2. You read a book.
3. The boy reads a book.
4. The boys read a book.
5. We (including you) read a book.
6. We (excluding you) read a book.

Che **a**moñe'ẽ peteĩ aranduka.

Nde **re**moñe'ẽ peteĩ aranduka.

Mitãkaria'y **o**moñe'ẽ peteĩ aranduka.

Mitãkaria'ykuéra **o**moñe'ẽ peteĩ aranduka.

Ñande **ña**moñe'ẽ peteĩ aranduka.

Ore **ro**moñe'ẽ peteĩ aranduka.

What can you infer from this data? **verbal prefixes for person/number agreement**

WORKED EXAMPLE II: VERBAL AGREEMENT IN PARAGUAYAN GUARANÍ

1. I read a book.
2. You read a book.
3. The boy reads a book.
4. The boys read a book.
5. We (including you) read a book.
6. We (excluding you) read a book.

Che **a**moñe'ẽ peteĩ aranduka.

Nde **re**moñe'ẽ peteĩ aranduka.

Mitãkaria'y **o**moñe'ẽ peteĩ aranduka.

Mitãkaria'y**kuéra** **o**moñe'ẽ peteĩ aranduka.

Ñande **ña**moñe'ẽ peteĩ aranduka.

Ore **ro**moñe'ẽ peteĩ aranduka.

What can you infer from this data? **verbal prefixes for person/number agreement**
bonus – indication of nominal plural inflection

A FEW GROUND RULES

1. Do not ask your language consultant what the language is they speak. It's a secret (for now). This prevents cheating but also makes it more fun. They will reveal it to you at the last session. You may be able to make some educated guesses, but try not to let it dominate your work.
2. If, from the start, you realise you recognise the language AND speak it (fluently or to some degree), let your language consultant know, and we will swap you to another group. But do not let slip to your fellow participants what the language is.
3. Work collaboratively as much as possible and give everyone a chance to get involved.
4. Be polite and respectful to your language consultants. They are sharing a wonderful thing with us!

REFERENCES

- Aikhenvald, A. Y. (2012). *The languages of the Amazon*. Oxford University Press.
- Bowern, C. (2008). *Linguistic fieldwork: A practical guide*. Palgrave Macmillan.
- Eberhard, D. M., Simons, G. F., & Fennig, C. D. (Eds.). (2023). *Ethnologue: Languages of the world* (26th ed.). SIL International.
- Endangered Languages Project. (n.d.). *Endangered languages*. <https://www.endangeredlanguages.com>
- Hammarström, H., Forkel, R., Haspelmath, M., & Bank, S. (Eds.). (2023). *Glottolog 4.8*. Max Planck Institute for Evolutionary Anthropology. <https://glottolog.org>
- Himmelman, N. P. (1998). Documentary and descriptive linguistics. *Linguistics*, 36(1), 161–195.
- Himmelman, N. P. (2006). Language documentation: What is it and what is it good for? In J. Gippert, N. P. Himmelman, & U. Mosel (Eds.), *Essentials of language documentation* (pp. 1–30). Mouton de Gruyter.
- Samarin, W. J. (1967). *Field linguistics: A guide to linguistic field work*. Holt, Rinehart and Winston.

STAY IN TOUCH



My personal linguistics
website



My linguistics
Instagram