Putting practice into words: Fieldwork methodology in grammatical descriptions

ICLDC 4 February 26-March 1, 2015 Honolulu

Lauren Gawne(1), Barbara F. Kelly(2), Andrea Berez(3), & Tyler Heston(3)

(1) NTU Singapore, (2) The University of Melbourne, (3) The University of Hawaii
Introduction

In science, claims must be falsifiable, verifiable, and reproducible.

Linguistic science values reproducibility too.

Today we will look at the state of reproducibility and methodology in language documentation research.
Reproducibility in Documentation

Language Documentation can make claims reproducible:

“[Language] documentation [...] will ensure that the collection and presentation of primary data receive the theoretical and practical attention they deserve.” (Himmelmann 1998:164)

“[...] it is our professional responsibility to provide the data on which our claims are based [...] It enhances the scientific basis of the linguists’ work.” (Theiberger 2009: 365-6)

“Establishing open archives for primary data is in the interest of making analyses accountable.” (Himmelmann 2006:6)
We have good models for fieldwork

On valuing reproducibility

Despite this, most descriptive publications make reproducibility difficult.

Boasian history (cf. Woodbury 2011):

➢ Raw textual data separate module from the descriptive grammar that generalizes over it
➢ No tradition of linking generalizing claims to data
➢ Old habits are hard to break!
Putting practice into words

What methodological features do linguists report on in their writing?

How do researchers link their writing back to the underlying data?
Data Sample

50 Published Grammars
50 PhD Dissertations
also 200+ journal articles (not discussed today)

All published/awarded 2003-2012
Grammars: A variety of publishers, institutions, languages
Articles: 9 Journals, range of areal and theoretical focii
Data Coding

1. Presence of methodology:

1.1 Books
Presence of explicit methodology rated on 1-5 scale (1= negligible, 5= comprehensive)

1.2 Articles
Presence of explicit methodology counted as yes/no

2. Tracked mention of:
- participants
- data collection equipment
- data analysis tools/software
- time in field
- genres collected
- archiving process

3. Data citation:
Rated on 1-5 scale (1=None, 5=Fully resolvable to underlying data) and conventions noted
Data distribution

Red - published grammars
Blue - dissertations
Methodology comprehensiveness

PhDs contain more of the categories that make a good methodology

Independent t-test
$t (98)=3.65, p=.0001$
Methodological features discussed

Some features more common than others

PhDs generally outperforming published grammars
Example referencing

Low rates of good practice in referencing back to original data.

Red - published grammars
Blue - dissertations

Number of publications

Score (1-5)
Still many improvement to make

A minimally good descriptive grammar should report on:

- participants
- data collection equipment
- data analysis tools/software
- time in field
- genres collected
- archiving process

And

It should link description to underlying data

e.g. Ring (2015: 424)

(1065)  

\[\text{tæ pan map utæ u hadem}\]

\[\text{teæ} \quad \text{pan} \quad \text{map} \quad \text{u}\text=eæ \quad \text{u}\text=hadæm}\]

\[\text{NVIS ask mercy M=NVIS M=Hadem}\]

‘then the Hadem (Daloi) asked forgiveness’ [PP05KO_049]
Encouraging good practice

We can also encourage good practice by others as well
Encouraging students

We can introduce students to good practice (cf. Pawley 2014)

At University of Hawaii major change to PhD Handbook of Requirements (since Fall 2013):

- Students whose theses are based on fieldwork are required to properly archive their data
- Archiving plans part of the Dissertation Proposal. Only accepted DELAMAN archives may be used.
- Students required to submit proof of deposit to the committee before the dissertation can be approved.
- Descriptive theses must cite resolvable resources.
Encouraging colleagues

Peer review process provides one opportunity to encourage colleagues to give more information about their research.

Informal gatherings like Linguistics in the Pub provide spaces for talking about documentation methods outside of ICLDC (www.rnld.org/LIP)
Conclusion: Put your good practice into words

Clear and transparent methodologies and examples linked to underlying data help make language documentation more reliable and replicable.
References

Berez, Andrea. 2014. “Reproducible research in descriptive linguistics: Integrating archiving and citation into the postgraduate curriculum at the University of Hawai‘i at Mānoa.” In Amanda Harris, Nick Thieberger & Linda Barwick (eds.), Research, records, and responsibility: Ten years of the Pacific and Regional Archive for Digital Sources in Endangered Cultures. Sydney: University of Sydney Press.

Berez, Andrea, Lauren Gawne, Barbara Kelly and Tyler Heston. In prep. Citation and transparency in descriptive linguistics.


Thank you.

These slides can be viewed at bit.ly/GoodMethods

Special thanks to the University of Hawaii Manoa University Research Council and the Dean’s Office of the College of Languages, Linguistics & Literature

Thanks to The University of Melbourne library staff