

4.5 Conclusion

p. 117 An often-voiced objection to the types of processes described in this chapter is that they take too much time. It is true that time-aligned transcriptions and interlinearized texts take some effort to construct, but we can liken it to building a house that has strong foundations. Once the initial effort has been made, the resulting house is able to withstand the elements and the passage of time, unlike a shack that is blown over by the first storm. Similarly, a well-constructed set of data will provide ongoing access, allowing you to ask new questions of it and to keep interacting with the primary material for a deeper analysis of the language. There is a great return on this initial investment of effort in creating an accessible and reusable corpus, especially if you will continue working with the same data over a long period of time. Furthermore, there is an ethical responsibility to prepare the data that we have recorded so that it can be located, accessed, and reused by those recorded or their descendants.

It is indisputable that there is more to research on a previously undescribed language than just the recorded data. The headnotes, the associated memories, and contextualization of the research that remain in the head of the researcher continue to inform their analysis. Similarly, the body of material recorded can only ever be a partial glimpse of a language that will continue to change over time. Nevertheless, the records we create will be invaluable for many purposes, and many more than we have planned for in our own research.

Creating good archival data can allow various derived forms that can be improved over time relatively easily. Rather than having to labour over formatting complex documents by hand, automatic processes can be applied to generate diverse outputs periodically as the underlying data is improved or as the analysis develops.

Finally, while the technologies we use will change, the underlying principles described in this chapter should provide some guidance in choosing new tools and methods from those that will appear in future. There is no question that linguistic fieldwork practice needs to change, and that future language descriptions will have to include recorded examples, cited in the analysis, with an underlying corpus that is archived and described in the course of the analysis. By adopting the methods described in this chapter, linguists will be building more detailed and accessible records of the world's languages.

4.6 Sources of Further Advice

p. 118 New technologies and methods associated with them are becoming available all the time. To keep up with options for tools or methods to use in your fieldwork, you should subscribe to appropriate mailing lists such as the Resource Network for Linguistic Diversity³⁷ or read blogs such as Endangered Languages and Cultures.³⁸ The SOAS Online Resources for Endangered Languages³⁹ (OREL) links to many useful sites. Another key source of information is the EMELD⁴⁰ site. There are also mailing lists specific to each of the tools discussed in this chapter, and this is likely to be the case for any new tools that appear in future. Useful references on topics around fieldwork and data management include Chelliah and de Reuse (2010: 197–225), Austin (2006), and Bower (2008: especially ch. 4 on data management and archiving, and ch. 13 on working with existing and historical field materials which covers issues not dealt with in this chapter).

Notes

- 1 Thanks to Christopher Cox and Toshihide Nakayama for their comments as reviewers of this chapter. Thanks also to Laura Berbusse, Tobias Bloyd, and Aaron Waldrip for helpful suggestions.
- 2 See Johnston (1995) and Antworth and Valentine (1998) and the other papers (a little dated but still useful) in Lawler and