(2000: 81–104). More detailed accounts are given in specialized textbooks on ethnolinguistics, register, and genre such as Saville-Troike (2003) and Biber and Conrad (2009).

With regard to morphosyntactic analysis, Samarin's requirements for a good corpus imply the following recommendations for gathering textual data. First, for a varied corpus one needs to make recordings of several types of text spoken by different kinds of people, such as traditional narratives (epics, legends, etc.), spontaneous narratives (anecdotes, personal histories, etc.), descriptions of activities, descriptions of objects, conversations, etc. (see §3.5.3.2), and as the corpus should also be repetitious, each genre should be represented several times. Since this ideal text corpus, which also has to include transcriptions and translations of all recordings, cannot be gathered in just a few years, the researchers and the speech community have to be selective and set priorities which are not determined by linguistic criteria, but by the practical necessity of recording first what the community considers as most important. If this, for example, is the traditional oral literature or the description of traditional rituals, the corpus may in the end lack casual conversations. On the other hand, the community may want to prioritize the recording of typical everyday communication as the basis for language revitalization measures. Accordingly, a grammar based on such corpora would not cover the full spectrum of verbal interaction, but nevertheless it can be an excellent grammar as long as it is made clear on what kind of data it is based.

Since the speakers' selection of linguistic forms depends on the kind of speech situation, the sources of all texts have to be described by metadata (Austin 2006: 92–4; Bowern 2008: 56–8; Caprile, Rivierre, and Thomas 1992; Himmelmann 2006a: 11–15; Samarin 1967: 102–4; and Thieberger and Berez, Chapter 4 below). Furthermore, in order to allow future researchers to scrutinize the morphosyntactic analysis, all text examples given in the grammar should be retrievable in the text corpus, and the text corpus itself should be accessible. To date these requirements are only met by a very few grammars, for example Thieberger (2006) and Wegener (2008).

Second, to 'get people talking', some fieldworkers use picture prompts like the frog stories as recommended
p. 87 by Bowern (2008: 116); others are more critical, ↳ because what people say when looking at such picture books is neither 'natural' nor 'interesting'. As Foley (2003) demonstrates, it may also differ structurally quite considerably from authentic narratives, and consequently may lead to false generalizations about the grammar of the language.

Third, translated texts of any kind should not be used, unless the researcher is fully aware of translational problems and wants to conduct a specialized investigation on translational interferences from the contact language. Recommendations to use translated material otherwise should not be taken seriously, although they are found in the fieldwork guides. Vaux et al. (2007: 105–7) even suggest inventing texts for an 'informant' to translate and 'tailor the text to fit your own interests as an investigator'. For a critical assessment of the use of Bible translations in morphosyntactic research, see De Vries (2007).

3.5.3.2 Text types

Rivierre (1992) presents a classification of text types that are relevant for linguistic fieldwork. He starts with the distinction between texts of the oral tradition 'with their careful, affected and often even archaic style', including the major traditional genres of historical narratives, myths, poems, etc., and 'more spontaneous texts, such as explanations of techniques, biographical accounts and anecdotes, or conversations in quite a different style which is often neglected' (Rivierre 1992: 56). He recommends collecting texts 'while developing the lexicon, proceeding by means of categories' and distinguishes six thematic categories (pp. 59–61):

1. locations and geographical and social environment;