

## Child Language Documentation: A Pilot Project in Papua New Guinea

Birgit Hellwig

*Department of Linguistics, University of Cologne*

Child language documentation: A pilot project in Papua New Guinea.<sup>9</sup>

### Abstract

The central aim of language documentation is to comprehensively document the characteristic speech practices of a community. Such practices necessarily also include child language and child-directed speech—and yet there are only very few documentation projects that focus on language from tandem with children. This paper argues for studying first language acquisition and socialization within a language documentation context, focusing on the types of data needed for such a study and drawing on the insights from a pilot project among the Qaqet of Papua New Guinea. The aim of this pilot project was to investigate the feasibility of a comprehensive child language documentation project, and this paper discusses the central challenges to such an endeavour and shows how they were addressed in the project.

### 1. Introduction

Language documentation is famously characterized as aiming to create a “comprehensive record of the linguistic practices characteristic of a given speech community” (Himmelman 1998: 166). Over the past two decades, there have been considerable advances in addressing the challenges of identifying and recording characteristic practices, of exploring ways of making the recordings accessible for multiple purposes and audiences, and of establishing a place for language documentation within descriptive and theoretical linguistics. An important aspect of the overall program is the notion of comprehensiveness (see Foley 2003; Franchetto 2006; Himmelman 1998, 2006; Lehmann 2001; Lüpke 2010; Seifart 2008). In an ideal world, projects would aim for a comprehensive coverage of linguistic practices – an aim that is necessarily limited by practical constraints of time, personnel, funding, and opportunities. A cursory survey of projects conducted under the umbrella of the major funding agencies of DoBeS (Dokumentation bedrohter Sprachen), ELDP (Endangered Languages Documentation Programme), and DEL (Documenting Endangered Languages) testifies to an impressive variety: some projects aim for comprehensiveness, while others explicitly limit this undertaking in various

---

<sup>9</sup> Many thanks to Susan Penfield and the National Science Foundation (Grant 1246525) for organizing the speaker’s series at the 3rd International Conference on Language Documentation and Conservation (ICLDC 2013) where the ideas for this paper were first presented. I warmly thank the audience as well as two anonymous reviewers for their interesting and challenging questions.

ways and focus on specific practices; there is also considerable variety in the chosen approaches. But the documented practices have one aspect in common: they are almost always practices of the adult language. Documentation projects that make reference to child language are conspicuously absent: out of roughly 500 projects funded by the above three agencies, only 5 are explicitly concerned with documenting child language, and another handful mention child language amongst the practices to be documented. With one exception (Stoll & Lieven n.d.), these projects are all pilot projects or smaller projects that explore specific aspects of child language or language transmission.

The exclusion of child language from language documentation cannot be theoretically motivated:<sup>10</sup> characteristic linguistic practices necessarily include child language, child-directed speech, and, more generally, the processes by which children are being socialized in a given community. Assuming that there are no theoretical reasons for excluding child language, the question arises as to why there are so few instances of child language documentation. One obvious reason for its absence is, of course, that some languages are no longer acquired by children. But this cannot be the whole story: there are large numbers of minority languages that are being acquired and for which processes of language transmission could be documented. It is more likely that its absence is due to various practical and methodological reasons, and this paper sets out to explore some of them by means of a case study.

The paper is structured as follows: §2 sets the scene by identifying challenges to child language documentation, focusing on the issue of data types and arguing for why we should nevertheless make an effort to address the challenges (see also Kelly & Nordlinger 2013 for a comparable discussion). It is acknowledged that this section only touches on these issues and that each of them could – and should – be expanded into a full paper in its own right. §3 takes up some of the challenges of §2, and describes in detail how they were addressed in a pilot project on child language documentation in Papua New Guinea. §4 then summarizes and concludes this paper.

## 2. Challenges to child language documentation: Data types

Child language documentation is anchored both in anthropology (i.e., language socialization research) and psycholinguistics (i.e., language acquisition research), drawing on three distinct data types: anthropological, experimental, and longitudinal data. Documenters are most familiar with the first data type, and there is a long-standing tradition of including anthropological components within language documentation projects (e.g., Franchetto 2006). But we are less familiar with the other two data types – both of which play a central role in language acquisition research. Experimental methods (see especially Blom & Unsworth 2010; Eisenbeiß 2010; Kidd 2006) allow researchers to systematically explore the extent and limits of a child's production and comprehension, and to understand the range of variation amongst a larger number of children. For a successful experiment, appropriate contexts and test items need to be identified and created, i.e., this presupposes a good understanding of the adult or target language. For well-documented languages,

---

<sup>10</sup> There are parallel debates about the place of children in anthropological research (Hirschfeld 2002). I thank an anonymous referee for pointing out this debate to me.

researchers can draw on numerous psycholinguistic resources: on lexical databases that give information on formal and semantic properties of words and their frequencies, or on standardized guidelines for assessing the communicative development of children. There may also be an understanding of the temporal organization of language development, i.e., of which types of non-target-like utterances tend to appear at which ages. And this knowledge, in turn, makes it possible to identify appropriate age ranges for a given experiment. Such resources and knowledge are not available for under-documented languages, and we thus need to develop them in parallel to studying child language. Arguably the most important such resource is a longitudinal corpus where the same children are followed over a longer period and are recorded at regular intervals (Demuth 1996; Tomasello & Stahl 2004). This data type, again, comes with its own challenges: ensuring that recordings take place at regular intervals over at least one year, that the recordings capture natural settings and are representative of the children's typical day-to-day activities, and that the masses of resulting data are transcribed and further processed (see Ochs 1979 for an excellent discussion of issues in transcribing child language). While language documentation has gained considerable methodological expertise in collecting and analyzing different data types, we have very little expertise in the two data types most relevant for language acquisition research: experimental and longitudinal data.

Taking a slightly different perspective, we can approach the issue of data types from the perspective of community involvement. Ideally, the community plays an active role in all language documentation research, contributing actively to the shaping of the project goals and research methods. There are very good reasons for this approach (Amery 1995; Czaykowska-Higgins 2009; Dobrin 2005; Grenoble & Whaley 2006; Grinevald 2003; Hinton & Hale 2001; Mosel 2006; Rice 2011; Wilkins 1992, 2000), but it can work against the inclusion of child language. Understandably, the speech community tends to be concerned about the disappearance of prestigious genres and specialized knowledge, rather than with the documentation of mundane everyday language, e.g., child language. In the face of scarce resources, it is thus not immediately obvious how to motivate a focus on child language. Similarly, the community tends to be interested in natural settings, reflecting their own invaluable communicative practices – something that cannot easily be reconciled with experimental setups that presuppose outside agency. These views tend to be mirrored by documenters who are often more comfortable with anthropological approaches to language socialization (Gaskins 1999; Kulick 1992; Ochs 1982; Schieffelin & Ochs 1986) and who remain cautious as to the applicability of experiments in field situations (e.g., the need to run experiments with a large number of speakers – something that is not always possible in the context of endangered languages).

While there is no question that anthropological methods are indispensable to investigating socio-cultural factors and their impact on children's language development, they cannot entirely replace experimental methods. Within language documentation, there are already provisions for including staged data: to supplement natural data (e.g., in order to document events that cannot easily be documented in a natural setting) or to better interpret natural data (e.g., to generate negative evidence and test the limits of productivity). Our normal types of staged data tend to be fairly unconstrained: speakers are presented with a stimulus, and they can talk naturally within the context created by the stimulus (Lüpke 2010). Experiments, by contrast, are typically more constrained, and they put the participants in the awkward role of passively reacting to a stimulus. The

situation is further complicated by the nature of experimental tasks: in order to provide meaningful results, they have to be challenging, i.e., they have to be just difficult enough for children to produce non-target-like utterances. In the worst case scenario, they create a feeling in the community that children are being set up to produce errors. An experimental approach is thus not easy to reconcile with the ideals of investing agency in the speakers of endangered languages and of fostering community engagement. More generally, this issue ties in with the overall question of fieldwork ethics: much effort has been directed towards issues of ethics and responsible fieldwork (Dwyer 2006; Musgrave & Thieberger 2006; Rice 2006), but we have little experience when it comes to working with children and to appraising the ethical consequences of experimental methods.

Supposing we have overcome the challenges of data collection, we are then faced with interpretative challenges. An important consideration in this respect is our limited knowledge of the target language. We focus on under-documented languages, and we usually do not have a good (let alone comprehensive) understanding of the adult language. Moreover, it is not self-evident that the adult language is actually the target language: we often work with endangered languages, i.e., with languages that are spoken in multilingual settings and whose speakers are potentially shifting to other languages. Both these circumstances pose challenges for interpreting child data: to be able to identify non-target-like utterances, and to distinguish between developmental effects (where a child has not yet acquired the target) and the effects of language shift (where a child is shifting towards a different target) (see McConvell & Meakins 2005; Meakins & Wigglesworth 2013; and O'Shannessy 2008, 2012 on the emergence of mixed languages in such contexts).

In the light of such challenges, it is not surprising that we know very little about the transmission and acquisition of small and endangered languages. In the CHILDES database (MacWhinney 2000), the vast majority of corpora are of Indo-European languages (58% of monolingual corpora; 83% of bilingual corpora). Conversely, there are estimates that language acquisition data exist for only 1-2% of the world's languages (Lieven & Stoll 2009: 144).

There is a need to remedy this situation. From the perspective of language acquisition, cross-linguistic corpora are urgently needed to test claims about universals of child language and child-directed speech. The current sample is biased towards Indo-European languages, i.e., towards languages that are related, that are typologically similar, and that are acquired in similar contexts (of monolingualism or family bilingualism). From this perspective (see Bates and MacWhinney 1989; Bavin 1995; Bowerman 2010; Eisenbeiß 2005; Slobin 1982, 1985–1997; Stoll 2009), there is thus considerable interest in expanding the current sample. This interest is reflected in a more general concern that the exclusion of the majority of the world's languages, peoples, and societies from consideration is untenable (Evans & Levinson 2009; Henrich et al. 2010). From the perspective of language documentation, child language and child-directed speech undoubtedly constitute characteristic linguistic practices (see §1) and thus need to be part of any comprehensive documentation. And finally from the perspective of language endangerment, we can hope to gain new insights into the causes of language endangerment and death. It is often argued (Fishman 1991; Grenoble & Whaley 2006) that a central cause is the interrupted transmission of language to the next generation, but we have little information on the processes of how languages are transmitted (or not transmitted) in such communities. An excellent study in this respect is Kulick (1992) who studied language socialization among

the multilingual Gapun in Papua New Guinea, identifying reasons for their beginning shift to the *lingua franca* Tok Pisin. From all three perspectives, there is thus an interest and a need to make the considerable effort to address the practical challenges posed by the documentation of child language.

### 3. Towards a child language documentation project

This section reports in detail on a pilot project on documenting child language among the Qaqet in Papua New Guinea. The intention of this pilot project was to investigate the feasibility of, and prepare for, a comprehensive child language documentation project. Given the methodological challenges outlined in §2, we considered it infeasible to start with either experimental or longitudinal studies. Instead we took the opportunity to discuss and engage community involvement, gain first anthropological insights into aspects of language socialization and typical children's activities (in preparation for future detailed research into language socialization), explore the collection of staged data (in preparation for future experiments), and try out possible recording scenarios (with a view to identifying natural contexts in preparation for future longitudinal studies).<sup>11</sup>

This project is set in the Qaqet communities of Raunsepna and Lamarain, located in the remote interior mountains of the Gazelle Peninsula (East New Britain Province, Papua New Guinea). Qaqet [ISO 639-3: byx] belongs to the Baining languages, a group of six related non-Austronesian languages, and it is spoken by an estimated 10,000 speakers. We have access to older descriptions of Qaqet (Bley 1914; Parker & Parker 1974, 1977; Rascher 1900, 1904) as well as to anthropological and historical sources (Dickhardt 2008, 2009, 2012; Fajans 1983, 1985, 1993, 1997, 1998; Hesse & Aerts 1996; Hesse 2007; Hiery 2007; Laufer 1946/1949, 1959; G. Pool 2015; J. Pool 1984; Rohatynskij 2000, 2001). In more recent times, a substantial body of descriptive work has been published on the related Baining language Mali [gcc] (see especially Stebbins 2011), but there is only a little information available on the other Baining languages. One of them, Makolkol [znh], is probably already extinct. Most Baining communities are multilingual, and people additionally speak the national *lingua franca* Tok Pisin, and often also a neighboring language (especially the Oceanic language Kuanua). In remote areas, the Baining languages are still strong, but in more accessible areas, Tok Pisin and other languages are becoming dominant and are starting to replace the Baining languages.

In 2011, I started researching the adult Qaqet language in Raunsepna. In the course of this research, the idea of documenting child language arose. This idea was triggered partly by my own developing research interests, and partly by the community's interest in education and literacy. We developed a pilot project together in order to test the feasibility and scope of such a project, and we were joined by two researchers with complementary interests and expertise: Evan Kidd (language acquisition) and Alexandra

---

<sup>11</sup> This pilot project would never have been possible without the generosity, hospitality, and enthusiasm of the people of Raunsepna and Lamarain, both adults and children. My thanks go to the entire community, and especially to John Landi, Tony Alin, Bruno Lalem, and their families for all their help throughout the project. The support of ELDP (Endangered Languages Documentation Programme, Grant SG110) for funding this pilot project and of the ARC (Australian Research Council, Grant FT0991412) for funding initial research on the adult language is gratefully acknowledged, and so is the endorsement of the National Research Institute of Papua New Guinea. This pilot project has since resulted in a comprehensive child language documentation project, and I thank the Volkswagen Foundation for their generous support of, and trust in, this project.

Marley (an MA student in sociolinguistics). This section outlines some of the key activities from the pilot phase: it describes views from the community (§3.1) and it outlines two case studies on narrative development (§3.2) and children’s activities (§3.3).

**3.1 Views from the community** There have been many discussions (both informal and recorded) within the community and between the community and the research team about the motivations for setting up a child language documentation project. The two most prominent views are illustrated in the quotes below. First, the view that young children do not talk very much (in 1). And second, the view that children do not talk correctly (in 2), especially if they are born into interethnic marriages (in 3).

- (1) Ee, kuasiq ai, luqa ip, [x] nani nyit, prama siitka amaigulka saa qutka.  
[...] Nyitagen praum ip taquarl, ip katika dip kadrlem iaum, taqurla.<sup>12</sup>  
‘Yes, there is no way that you can go with a long story to a baby. [...] You will say a short (word) just like (the baby), because he will only know the short (word), that’s all.’ [I12ABLAJLATASocio2 581.635 591.890]
- (2) Deiv iani ngerenarli angama rluimini ngamraqen taqurla, de maika dip kuasiq ini ngereseserl vriana rluimini  
‘And when someone hears a little child talk like this (incorrectly), they will not always correct their little child.’ [I12AANACLADNSocio2 1220.160 1224.640]
- (3) Kuasiq i rataqa tdadem. [...] Amanepna. [...] Bequrini itika quasiq ikias ka uranaik.  
‘They do not follow (Qaqet) straight. [...] (Because) they are mixed. [...] I assume it’s because we are not amongst ourselves.’ [I12ABLAJLATASocio3 92.012 150.699]

Such views were often followed by veiled skepticism: since children do not speak very much, and since they do not speak correctly, it would be better to focus attention on the adult language. Such skepticism, however, was countered by the equally prominent view that efforts needed to be made to strengthen the children’s Qaqet skills. This included, on the one hand, the speaking skills of children of interethnic marriages, as there is a strong sense that such children do not speak Qaqet well.<sup>13</sup> And it included, on the other hand, the literacy skills of all children. The national language policy at that time (Devette-Chee 2012; Litteral 2004; Siegel 1996) envisaged that children should be taught in either a vernacular language or in Tok Pisin in the first two years of elementary school (Prep and Year 1), and should then transition to English in the last year (Year 2). Following that, English should become the language of instruction throughout primary school (Year 3 to Year 8). In Raunsepna, too, teachers attempt to teach literacy in the vernacular language, but they are faced with numerous seemingly unsurmountable difficulties. The Qaqet are a

12 All examples are presented in the Qaqet orthography and followed by an English free translation; they do not contain any interlinear glossing. The translation is followed by a bracket containing an identifier for the example (file name and time index), linking to the Qaqet corpus archived with the Endangered Languages Archive (ELAR) (Hellwig 2010–2013).

13 This sense was partly confirmed by the sociolinguistic part of our pilot project (see Marley 2013 for details). While Qaqet is still strong in this remote community, there are certain demographic factors that trigger shifts away from Qaqet and towards the lingua franca Tok Pisin—including growing up in an interethnic marriage. Such marriages are becoming more common, thus supporting the community’s view that Qaqet is losing ground under such circumstances.

marginalized group, and the school is under-resourced in every respect, including a lack of Qaqet literacy materials. The general level of literacy is low, and children from this community do not perform well on the national exams. The community attributes these facts to the early school years: the challenges of teaching literacy in the absence of any Qaqet materials (let alone age-appropriate materials) as well as the absence of any materials that would help students transition from Qaqet to English. At the moment, it is entirely up to the teachers to create their own materials for teaching literacy and transitioning to English.

Against this background, the community expressed an interest in studying child language with the explicit goal of developing age-appropriate materials to be used in the elementary school (to teach Qaqet literacy and to transition to Tok Pisin and English literacy). We therefore set up the project in close co-operation with the teachers, especially with Betty Dargas and Chris Mitparlingi, and with additional support from Patrick Lemingel and Paul Liosi. The scope of the project was shaped in discussions with these teachers, necessitating numerous modifications of our earlier views on what is possible and desirable. At the same time, the teachers acted as intermediaries to the community, and discussed the project with the children and parents. In the end, we were able to jointly identify a number of acceptable recording contexts, and the teachers took care of the logistics: organizing the recordings, acting as interlocutors for the children, and helping with the transcriptions and translations. We also updated the community on the progress of the project in the form of regular announcements after Sunday church. And parents and relatives were invited to view the recordings of their children at any time. Many of them took this opportunity, and these meetings often gave new impulses for further recordings of both children and adults. In the end, we had 70 participants in the project (20 adults and 50 children), testifying to the enormous interest and involvement of the entire community.

**3.2 Case study 1: Narrative development** In one study, we made use of the so-called Frog Story, i.e., the illustrated wordless picture book *Frog, where are you?* (Mayer 1969) that depicts the adventures of a boy, his dog, and their frog. We had 41 children (between 5 and 13 years of age) look at the book, and retell the story to an adult interlocutor, either in Qaqet (25 children) or in Tok Pisin (16 children). This part of our study was partly intended to explore possible pitfalls when collecting staged data with children (in preparation for future experimental studies) and was partly linked to the community goal of developing age-appropriate literacy materials. The idea here was to create a stock of picture books of local and new stories, to ask children to tell these stories (thereby enabling us to investigate their narrative development), and to then create a written version of these stories by writing down the expressions used by the children in their oral narratives. We trialed this approach with the Frog Story, and each participating child in the end received their own version of this story.

The Frog Story has been used successfully to investigate narrative development in many different languages, and the methods for data collection are well defined (Berman & Slobin 1994; Strömquist & Verhoeven 2004). In particular, the adult interlocutor is asked to take as non-active a role as possible (only offering non-committal back-channels that do not interfere with the child's narrative), and the child is not supposed to know the story beforehand. In our pilot study, we ran into numerous challenges trying to follow this methodology. This section recounts the two most prominent such challenges:

the children's narratives reflected 1) normal classroom behavior and 2) normal story-telling practices within the community. We assume that these challenges are not unique to our fieldwork setting, but rather characteristic of conducting such studies under fieldwork conditions.

The first challenge is directly linked to our decision to work together with the local elementary school, pursuing a collaborative research model (e.g., Czaykowska-Higgins 2009; Rice 2011; see also references in §2). As summarized in §3.1, this decision worked exceedingly well on one level. It ensured the involvement of the community as well as the usefulness of the project to the community, and it facilitated the logistics of data collection. But on a different level, it worked less well, as it created a classroom atmosphere. Both teachers and children worried that we had a hidden agenda of evaluating their performances. This worry showed most obviously in numerous instances where children hesitated, and their teachers then whispered the 'right' words for them to repeat, so that they would tell the story 'correctly'. In the end, it was fairly straightforward to deal with this particular issue (through further discussions with the teachers about the purposes of the study). But the classroom atmosphere also showed in another issue: a tendency to co-construct the story. As outlined below, this issue was more complicated to deal with, since it reflected normal classroom practice.

At the beginning of this study, many children were shy and clearly uncomfortable. Added to their very understandable apprehension at performing for strangers, it is not usual in this community for children to be singled out and to answer questions individually. Instead, it is more common for them to repeat a teacher's words or to answer questions in a chorus. The teachers were thus concerned that our normal research procedure would only increase their shyness and discomfort. We addressed this problem by asking the children to narrate the story in pairs. Example (4) illustrates a typical story recorded with this procedure: the two children (both 9 years old) often talk at almost (but not quite) the same time, with almost (but not quite) the same words.<sup>14</sup> That is, one child or other would take the initiative, and the other would chime in, either repeating what was said or continuing on from what was said.

- |     |      |                                     |   |
|-----|------|-------------------------------------|---|
| (4) | ZGM: | <i>iva de..</i>                     | and then..  |
|     | ZGD: | <i>iane<b>brlany</b></i>            | the two [sleep                                    |
|     | ZGM: | <i>[iane<b>brlanya</b></i>          | [ <b>the two</b> sleep now                        |
|     | ZGM: | <i>dequasiq ian<b>tliana</b>..</i>  | and the two don't [see <b>their</b> ..            |
|     | ZGD: | <i>[ian<b>tliana</b> [quldit</i> .. | [ <b>the two</b> (don't) see their [ <u>it</u> .. |
|     | ZGM: | <i>[iana<b>qulditki</b></i>         | [ <u>their</u> frog                               |
|     | ZGD: | <i>be [qiamit</i>                   | and [ <b>it has left</b>                          |
|     | ZGM: | <i>[<b>be qiamit</b></i>            | [ <b>and it has left</b>                          |
- [R12ZGDZGMFrog 19.310 25.535]

Alternatively, the children negotiated a 'spokesperson' (whose identity often shifted in the course of the narrative). This spokesperson would not narrate the story

<sup>14</sup> In this excerpt, overlaps between adjacent lines are marked by "[", and boldface or underlining.



independently, but check with their friend, or the friend would offer help or corrections. These exchanges almost always took place secretly and in whispers, and were not always picked up by the microphone. Example (5) illustrates such typical interactions. The children in this example are both 10 years old.

- (5) ZAL: *kemnyim bit* he looks up  
 ZAL: [[whispers]] *da?* [[whispers]] right?  
 ZJD: [[shakes head]] [[yes]]  
 [R12ZALZJDFrog 113.645 116.805]  
 ZAL: *dap ma.. aadangga mara* and the.. his dog here  
 ZJD: [[whispers]] *John* [[whispers]] John  
 ZAL: *ia, John* sorry, John (not the dog)  
 [R12ZALZJDFrog 188.550 192.015]  
 ZJD: *Johna, qena aadangga..* John now, together with his dog..  
 ZAL: [[whispers]] *ianequuk nas* [[whispers]] the two swim  
 ZJD: *ianequuk nas* the two swim  
 [R12ZALZJDFrog 243.145 248.010]

The excerpts in (4) and (5) both exemplify children's ways of co-constructing a story: the normal way in this community for children to react to a teacher's question or instruction. From this perspective, our study offers excellent data. But our procedure does not conform to the recommended methods, thereby reducing the cross-linguistic comparability of our results on children's narrative development. It is valid to ask whether (or, maybe more to the point, to what extent) established methodologies should always be followed, even if they do not seem pertinent in a given socio-cultural context (see §4). But regardless of the answer, we have to be aware that varying the methods of data collection will make it difficult to compare the results to those obtained with other methods in other languages.

Over the course of the study, children and teachers became increasingly familiar with our research, and after a few weeks, children were happy to narrate the story on their own. We tried to further facilitate this shift by recording children in their home settings with family members as interlocutors. This significantly complicated the logistics (and we only managed to record 3 children in the home setting, compared to 38 children in the school setting), but reduced the feeling of being tested in a classroom setting. Example (6) is a typical recording from such a home setting, illustrating the ease of interaction between the 6-year-old child and her mother. There are still some remaining methodological issues in that adult interlocutors sometimes took too active of a role (e.g., the mother's question *nemgi iari?* 'who stays there?' in line 4), but this cannot be avoided altogether. For example, it would be inappropriate for the mother in line 2 to not answer her child's question. And generally, the adult interlocutors were comfortable with issuing only vague prompts such as *(be) nana?* '(then) what?' (in lines 6 and 8).

- (6) ZMS: *ademga, i qua maqasupka?* a hole, so is it of the rat?  
 mother: *kuasik, auaiiki* no, (of) a bird  
 ZMS: *auaiiki* a bird  
 mother: *nemgi iari?* who stays there?  
 ZMS: *aququanngi* an owl  
 mother: *be nana?* then what?  
 ZMS: *kiaseserl sa..* it (owl) startles..  
 mother: *nana?* what?  
 ZMS: *kiaseserl, sama qaqeraqa, de qaat meseng* it (owl) really startles, the boy, and he falls down  
 mother: *[[laughs]]* *[[laughs]]*

[R12ZMSFrog 140.480 159.315]

The second major challenge is linked to normal story-telling practices within the community. We were able to find out about such practices because our pilot study included anthropological and sociolinguistic components. As illustrated by the quote in (7), stories are told by adults to children, and children hear and repeat these stories to their friends, thus allowing the stories to spread. Something similar happened to the Frog Story. It was picked up and repeated, and became a topic of conversation in the community (as illustrated by the excerpt from a conversation in 8).

- (7) Ide resiit nanget de dama renngi aris. [...] Deip maget, de sa nyitlirang ngatit. De sa irang ngeresiit, naluqa amasiitka, imedu iani, de ngerenarliqa. [...] Tika amasiitka qatit. Sa qatira, mrama.. ama.. amaburlem nara. Be radrlem luqa amasiitka.

‘They (the parents) used to tell them (stories) at night. [...] Then later you will right away see the little ones (children) go around. And they will right away tell this (same) story (to their friends), the (story) that they have just heard. [...] And so the story spreads. It now spreads to many (people). And they now know the story, too.’ [I12AANACLADNSocio3 1135.065 1180.206]

- (8) Mani ngumr ama.. amarluis [...] uresiit savrama qalminngi.

‘Yesterday I took the children [...] and we told stories about a frog.’ [I12AANACLADNSocio3 1291.512 1295.322]

We thus cannot be sure that the children are constructing the story by themselves. In fact, there are indications that they are repeating a story that they have heard before. The best such indication comes from the ending of the Frog Story, where the boy and the dog find their frog among a whole group of frogs. Children initially ended their story by saying something along the lines of ‘they pick up their frog and they go’. But after a certain date, they typically added the following: they pointed to an empty space in the line-up of frogs, and said ‘and the space here, they took their frog from here’. On this particular day, we had recorded a grandmother telling the story to her two grandchildren. Example

(9) repeats the relevant part of this recording: the grandmother ends the story, and her 10-year-old grandson points out the empty space in the line-up of frogs. The grandmother later told me that she had not noticed this empty space, and I confirmed that I had not noticed it either. The story of this boy seeing something that the adults had not seen must have spread through the community, because after that day, many children made sure to highlight this empty space whenever they narrated the story.

(9)	grand-mother:	<i>deqerl ianluqia</i>	and the two have found it now
	YCS:	<i>nyilara ngilkaira, namen lura!</i>	see its (empty) space, among those ones!

We have yet to find a satisfactory way of dealing with this issue. We suspect that it will partly resolve itself through time and familiarization. As people become more familiar with this type of research, the stories will become less exciting and newsworthy. And we intend to circumvent the issue by varying the interlocutor (so that children will tell stories to different people) and by building up a pool of different stories (so that children cannot predict which stories we will ask them to narrate). This includes a number of little-known local stories: several people have volunteered such stories, and we are currently adapting them and creating picture book versions of them.

**3.3 Case study 2: Children’s activities** A second study focused on natural events: we intended to record a number of younger children (between 2 and 3 years of age) in natural settings in order to develop an initial idea of typical children’s activities, and of their typical language input and output. The ultimate goal was to explore suitable recording contexts for a future longitudinal study.

The first challenge was to determine the ages of the children. In this community, age is not important, and people usually do not know their or their children’s ages. When asked, people report on two salient transition points. The first such point is the inclusion in the national vaccination program (i.e., distinguishing children under the age of 5 who are still receiving vaccinations). And the second point is the enrollment in elementary school: children are supposed to have enrolled by the age of 7, but because of the under-resourcing of the school, they are usually much older (e.g., it is not unusual for a child to be 9 or 10 years old in their first year of school). These two transition points single out groups that cover considerable age ranges (i.e., children under 5, and children of 7+), and there are no other widely recognized transitions that would further narrow down their actual ages. Estimates turned out to be very unreliable, and even those who knew the community well tended to judge children much too young. There are, of course, ways of establishing a child’s age: through personal discussions (i.e., establishing salient events that happened around the child’s birth, or establishing their age relative to that of other children) and through tracking down official records (i.e., clinics record the day of birth on the child’s health card, and churches record the day of baptism in the church register). But both approaches are not straightforward and require considerable time and effort. We pursued all of the above options, and, in the end, we found three children in the relevant age range (two children of age 2;11 and one of age 2;5) whose families were happy to participate in this study.

Following that, the main challenge was to find a balance between the naturalness and the quality of the recordings. We discussed typical children's activities with the community and the families, and it turned out that children are exceptionally mobile. Young children would usually spend their day in the care of older siblings and cousins, and accompany them in their activities – which take them into the bush and gardens to collect firewood or food, to run errands, or to play games in the bush and the rivers. The possibilities for recording are further constrained by the climate and environment, as there are torrential rains most afternoons, and most places are covered in mud and water throughout the day. All these aspects (older siblings as care-givers, mobility of children, environmental challenges) are not unusual in many parts of the world, and they make it difficult to document children in natural settings. Our approach was to equip the focal children with backpacks with cheap audio recorders and voice-activated microphones. On the plus side, this approach allowed us to capture natural interactions. On the downside, the quality of the recordings suffered, and we especially lacked contextual information (i.e., it was not always clear what the other children were saying or doing). For the future, we intend to experiment with appointing older children as research assistants, and equipping them with their own recording devices (and with the instruction to try and record group activities whenever possible). But no matter what solution we will come up with in the end, it is clear that this kind of data needs to be supplemented with higher quality naturalistic data that is recorded in less challenging settings.<sup>15</sup> One of the promising settings turned out to be the evening meal, which is the normal time for family, for relaxation and chatting—but which also takes place in near darkness, thus interfering with the video quality. As it turned out, it was almost impossible to interpret young children's utterances without access to the visual image. Another promising setting was a staged setting: arranging for a number of adults to be engaged in a sedentary activity (e.g., weaving a rope or a netbag), with two or more children present. It turned out that during this setting, the children do not stray too far away from the adults, and they help out and interact with the adults and the other children.

By contrast, we had much less success with a staged setting that works well in Western contexts: staged play activities, where a mother or other adult engages a child in talk while playing with toys. In the Qaqet community, this kind of activity is not normal. Adults usually do not engage children in such activities, and they are not used to drawing children out and encouraging them to speak. This is of course not unique to Qaqet culture (as shown by the ethnographic literature on language socialization, e.g., Duranti et al. 2011; Ochs 1982; Schieffelin & Ochs 1986). Example (10) is an extract from our first such recording, illustrating the helplessness and increasing despair of the mother in trying to get her child (aged 2;5) to talk and perform for us under these circumstances.

---

15 We have since embarked on a longitudinal study of 2- to 4-year-old children, and have had great success with two set-ups that we had not trialed in the pilot project reported here. First, we equipped parents with easy-to-use cameras (the Zoom Q4) and discussed with them the concept of 'natural settings'. Following that, they themselves decided when to turn on the camera. This procedure resulted in a large number of recordings of children in typical everyday situations. The metadata, albeit not the actual data, can be accessed through the Language Archive Cologne (Hellwig et al. 2014–2019). And second, we have cross-checked the representativeness of the longitudinal study by means of a) participant observation and b) wiring children with audio recorders and recording them for an entire day. As a result, we know that the longitudinal study is largely representative of the children's daily activities; and that the main gap is, indeed, a scarcity of recordings of children alone in the bush.

- (10) mother: *nyitaqen* you talk  
 mother: *nyi, nyitaqen* you, you talk  
 mother: *so, uantaqen* now, you two talk  
 YMM: *oki?* [[pointing to microphone]] don't touch?  
 mother: *ai, tika nyitlak* hey, you just play  
 YMM: [[cries; reaches for microphone]]  
 mother: *de qurliqia* leave it now  
 mother: *de nyitaqen* you talk  
 YMM: [[cries; plays with stones]]  
 mother: *nyitaqen* you talk

[C12YMMZJIPlay1 7.860 26.777]

The families adapted to this unfamiliar activity essentially by ignoring the game scenario. Sooner or later, natural interactions would find their way into the staged play activities (as illustrated in example (11), which happened sometime after example (10) above during the same recording session). We also discussed possible questions and strategies for adults to draw out their children, getting them to talk, and thereby reducing the pressure and anxiety of the mothers. It turned out that the question that worked best by far was the ‘where’ question (as illustrated in (12), with a child aged 2;11).

- (11) YMM: *amana* [[points at tea]] this one  
 mother: *as kerl uilas* it is still hot  
 YMM: *amana::* this o::ne  
 mother: *auilas* it is hot  
 mother: *as nguadang* [[pretends to be burned]] I'm burned  
 YMM: [[moves over to grab tea]] [...] [...]  
 mother: *a'ai, kua nguis?* okay, shall I blow (and cool it)?  
 mother: *nyisrlup* now you drink

[C12YMMZJIPlay1 559.262 571.479]

- (12) mother: *ah? papa qua?* huh? where is daddy?  
 ZGT: *papa lu qua?* daddy is where?  
 mother: *nguluqa qua?* where is he?  
 ZGT: *qua?* where?  
 mother: *kuaridi i qamit kua?* where did he go?  
 ZGT: *mit agalip* (he) went (for) groundnuts  
 mother: *ah?* huh?  
 ZGT: *agalip* Groundnuts  
 mother: *kamit kua?* where did he go?  
 ZGT: *galip* Groundnuts

mother:	<i>da?</i>	really?
ZGT:	[[shakes head]]	[[yes]]

[C12VARPlay 416.843 430.533]

Despite the unfamiliarity of the setting, the staged play activities proved fairly successful in the end in generating a good amount of child data and child-directed speech. But without complementary natural data, we can only guess how far these data reflect the typical language learning process and the typical language input and output. Different fieldsites will pose different challenges, but it is likely that one of the major challenges will always be to find a range of suitable recording contexts that give insights into typical children’s activities, and at the same time that lend themselves to high quality recordings of child language and child-directed language.

## 1. Conclusions

Our project among the Qaqet Baining of Raunsepna has set out to explore the possibilities of conducting a comprehensive child language documentation project in this community. This section now takes a step back and draws some general lessons from our experiences.

As is the case for all language documentation projects, child language documentation can only succeed in close cooperation with the community. A large part of the pilot project was spent in discussions with the community about the motivations and purposes of such a study, possible practical outcomes, and issues of methodology and logistics. In conjunction with these discussions, we experimented with various methods and ways of recording children. We consider this time well spent. The discussions shaped and modified our original ideas, thus ensuring the feasibility of the project; they also created an atmosphere of enthusiasm and trust in the community. The trialing of methods enabled us to form a realistic view on what is and is not possible. And, just as importantly, they familiarized everyone – the outside researchers, the community as a whole, the care-givers, and children – with the research procedures. The sessions became increasingly normal, and adults and children no longer saw them as intrusions into their lives. And with their growing knowledge of the research procedures, the community members became increasingly confident to participate in the overall project design, introducing modifications and new goals, and pointing out fruitless endeavors. This phase lasted for about three months, and it would be fair to characterize it as a phase of trial and error.

It was argued in §2 that child language documentation is of interest from different perspectives, including the perspectives of language acquisition, language documentation, and language endangerment. Our pilot project generated data that spoke to all three perspectives. The corpus includes large numbers of child utterances, both target- and non-target-like, that allow for hypotheses on the acquisition trajectory of various phenomena and that point to phenomena whose acquisition seems protracted. The sociolinguistic and anthropological data give insights into the linguistic landscape and ideologies and practices of socialization that, from a practical perspective, help us identify suitable recording contexts. At the same time, the many instances of child-directed speech in the

corpus allow for initial investigations into the linguistic and anthropological properties of this register. And finally, the instances of code-switching give insights into the uses of Qaqet and Tok Pisin. However, I have deliberately phrased these results in a careful way: the purpose of the pilot project was to investigate the feasibility of a larger child language documentation project, i.e., its focus was always on methods. From the perspective of language acquisition research, what is needed now is a longitudinal study where the same children are recorded at regular intervals over a longer period. Such a corpus will give insights into the temporal organization of children's language development. On its basis, it will then be possible to investigate individual phenomena more rigorously in the form of experimental studies: it will allow us to formulate specific testable hypotheses, to identify appropriate contexts and ages for experiments and such a corpus would also help interpret the results of experiments. We have since embarked on such a longitudinal study, and it would be fair to say that it would not have been possible without this pilot stage. We would therefore always recommend factoring in a pilot phase and allowing ample time for discussions and trialing of methods.

While the pilot project made it possible to determine the feasibility and scope of a comprehensive project, it also raised a more fundamental question. In the course of the study, we encountered numerous problems with the methods developed in a Western context. It is true that we overcame most of these difficulties in the end. However, it is not entirely clear whether it is always necessary or even desirable to go through this effort. This issue arose especially in the case study on narrative development (see §3.2). Given that the children's behavior reflects normal social and cultural practices (such as the practices of co-constructing stories, and of repeating known stories), the question arises as to how to deal with such practices. On the one hand, there is a need to document characteristic speech practices – in their own right, and as instances of alternative ways of socializing children into narrating stories. And, on the other hand, there is a need to follow prescribed methodologies that arose in Western contexts – methodologies that were developed for good reasons, and that make the collected data comparable cross-linguistically. Or, put differently, should we think of child language documentation as part of the psycholinguistic discipline of language acquisition (and the challenge then would be to adapt the existing methods to diverse linguistic, cultural, and geographic settings so that they yield results outside of Western contexts and laboratories)? Or is it possible to articulate a different (complementary or overlapping) agenda, and to integrate the more anthropological and community-oriented outlook of language documentation into the overall enterprise? Our study is far too limited to even attempt an answer to this question. But it highlights the need to bring data and experiences from non-Western contexts to bear on the methodological and theoretical debate. And it highlights a second factor: the need for collaboration between researchers from language documentation, language acquisition, and language socialization. In the course of our pilot project, it was an invaluable asset to be able to discuss the methodological and theoretical challenges within the research team and with outside colleagues. It enabled us to gain an understanding of the different perspectives and approaches, and to form an idea of how to collect data that are of use to all three fields.

## REFERENCES

- Amery, Rob. 1995. It's ours to keep and call our own: Reclamation of the Nunga languages in the Adelaide region, South Australia. *International Journal of the Sociology of Language* 113. 63–82.
- Bates, Elizabeth & Brian MacWhinney. 1989. *The crosslinguistic study of sentence processing*. Cambridge: Cambridge University Press.
- Bavin, Edith L. 1995. Language acquisition in crosslinguistic perspective. *Annual Review of Anthropology* 24. 373–396.
- Berman, Ruth A. & Dan I. Slobin. 1994. *Relating events in narrative*. Mahwah, NJ: Erlbaum.
- Bley, Bernhard. 1914. Sagen der Baininger auf Neupommern, Südsee. *Anthropos* 9. 196–220, 418–448.
- Blom, Elma & Sharon Unsworth (eds.). 2010. *Experimental methods in language acquisition research*. Amsterdam: John Benjamins.
- Bowerman, Melissa. 2010. Linguistic typology and first language acquisition. In Jae Jung Song (ed.), *The Oxford handbook of linguistic typology*, 591–617. Oxford: Oxford University Press.
- Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3(1). 15–50.
- Demuth, Katherine. 1996. Collecting spontaneous production data. In Dana McDaniel, Cecile McKee & Helen Smith Cairns (eds.), *Methods for assessing children's syntax*, 3–22. Cambridge: MIT Press.
- Devette-Chee, Kilala. 2012. The impact of Tok Pisin and local vernaculars on children's learning in Papua New Guinea. *Language and Linguistics in Melanesia* 30(2). 47–59.
- Dickhardt, Michael. 2008. Die Feuertänze der Baining. In Ines de Castro, Katja Lembke & Ulrich Menter (eds.), *Paradiese der Südsee. Mythos und Wirklichkeit*, 98–101. Hildesheim and Mainz: Roemer- und Pelizaeus-Museum; Philipp von Zabern.
- Dickhardt, Michael. 2009. *Schau nur, und also wirst du dich wandeln! Eine Studie zur Kulturanthropologie der Moralität unter den Qaqet-Baining von Raunsepna, Neubritannien, Gazellehalbinsel, Papua-Neuguinea*. Habilitationsschrift, Göttingen: Georg-August-Universität.
- Dickhardt, Michael. 2012. Die mit den Geistern tanzen. Maskentänze, Identität und Moral unter den Qaqet-Baining (Gazellehalbinsel, Neubritannien, Papua-Neuguinea). *Mitteilungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte* 33. 15–32.
- Dobrin, Lise M. 2005. When our values conflict with theirs: Linguists and community empowerment in Melanesia. In Peter K. Austin (ed.), *Language documentation and description*, vol. 3, 42–52. London: SOAS.
- Duranti, Alessandro, Elinor Ochs & Bambi Schieffelin (eds.). 2011. *The handbook of language socialization*. Malden, MA: Wiley-Blackwell.
- Dwyer, Arianne M. 2006. Ethics and practicalities of cooperative fieldwork and analysis. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 31–66. Berlin and New York: Mouton de Gruyter.
- Eisenbeiß, Sonja. 2005. Documenting child language. In Peter K. Austin (ed.), *Language*



- documentation and description*, vol. 3, 106–140. London: SOAS.
- Eisenbeiß, Sonja. 2010. Production methods. In Elma Blom & Sharon Unsworth (eds.), *Experimental methods in language acquisition research*, 11–34. Amsterdam: John Benjamins.
- Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals. *Behavioral and Brain Sciences* 32. 429–448.
- Fajans, Jane. 1983. Shame, social action, and the person among the Baining. *Ethos* 11(3). 166–180.
- Fajans, Jane. 1985. The person in social context: The social character of Baining ‘psychology’. In Geoffrey M. White & John Kirkpatrick (eds.), *Person, self, and experience. Exploring Pacific ethnopsychologies*, 367–397. Berkeley: University of California Press.
- Fajans, Jane. 1993. The alimentary structures of Kinship: Food and exchange among the Baining of Papua New Guinea. In Jane Fajans (ed.), *Exchanging products: Producing exchange*, 59–75. Sydney: Sydney University Press.
- Fajans, Jane. 1997. *They make themselves: Work and play among the Baining of Papua New Guinea*. Chicago: The University of Chicago Press.
- Fajans, Jane. 1998. Transforming nature, making culture: Why the Baining are not environmentalists. *Social Analysis* 42(3). 12–27.
- Fishman, Joshua A. 1991. *Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.
- Foley, William A. 2003. Genre, register and language documentation in literate and pre-literate communities. In Peter K. Austin (ed.), *Language documentation and description*, vol. 1, 85–98. London: SOAS.
- Franchetto, Bruna. 2006. Ethnography in language documentation. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 183–211. Berlin and New York: Mouton de Gruyter.
- Gaskins, Suzanne. 1999. Children’s daily lives in a Mayan village: A case study of culturally constructed roles and activities. In Artin Göncü (ed.), *Children’s engagement in the world: Sociocultural perspectives*, 25–61. Cambridge: Cambridge University Press.
- Grenoble, Leonore A. & Lindsay J. Whaley. 2006. *Saving languages: An introduction to language revitalization*. New York: Cambridge University Press.
- Grinevald, Colette. 2003. Speakers and documentation of endangered languages. In Peter K. Austin (ed.), *Language documentation and description*, vol. 1, 52–72. London: SOAS.
- Hellwig, Birgit. 2010–2013. Qaqet corpus. London: Endangered Languages Archive. <https://elar.soas.ac.uk/Collection/MPI188145>.
- Hellwig, Birgit, Carmen Dawuda, Henrike Frye & Steffen Reetz. 2014–2019. Qaqet Child Corpus. Cologne: Language Archive Cologne. <http://hdl.handle.net/11341/00-0000-0000-0000-202A-0@view>.
- Henrich, Joseph, Steven J. Heine & Ara Norenzayan. 2010. The weirdest people in the world? *Behavioral and Brain Sciences* 33. 61–83.
- Hesse, Karl. 2007. *A Jos! Die Welt, in der die Chachet-Baininger leben. Sagen, Glaube und Tänze von der Gazelle-Halbinsel Papua-Neuguineas*. Wiesbaden: Harrassowitz Verlag.

- Hesse, Karl & Theo Aerts. 1996. *Baining life and lore*. Port Moresby: University of Papua New Guinea Press.
- Hiery, Hermann Joseph. 2007. Die Baininger. Einige historische Anmerkungen zur Einführung. In Karl Hesse, *A Jos! Die Welt, in der die Chachet-Baininger leben. Sagen, Glaube und Tänze von der Gazelle-Halbinsel Papua-Neuguineas*, VII–XXIX. Wiesbaden: Harrassowitz Verlag.
- Himmelman, Nikolaus P. 1998. Documentary and descriptive linguistics. *Linguistics* 36. 161–195.
- Himmelman, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelman & Ulrike Mosel (eds.), *Essentials of language documentation*, 1–30. Berlin and New York: Mouton de Gruyter.
- Hinton, Leanne & Kenneth Hale (eds.). 2001. *The green book of language revitalization in practice*. San Diego: Academic Press.
- Hirschfeld, Lawrence. 2002. Why don't anthropologists like children? *American Anthropologist* 104(2). 611–627.
- Kelly, Barbara F. & Rachel Nordlinger. 2013. Fieldwork and first language acquisition. In Lauren Gawne & Jill Vaughn (eds.), *Selected Papers from the 44<sup>th</sup> Conference of the Australian Linguistic Society, 2013*, 178–192. Melbourne: University of Melbourne.
- Kidd, Evan. 2006. Language acquisition research methods. In *Encyclopaedia of Language and Linguistics*, 311–315. Oxford: Elsevier.
- Kulick, Don. 1992. *Language shift and cultural reproduction: Socialization, self, and syncretism in a Papua New Guinean Village*. Cambridge: Cambridge University Press.
- Laufer, Carl. 1946/1949. Rigenmucha, das Höchste Wesen der Baining (Neubritannien). *Anthropos* 41/46. 497–560.
- Laufer, Carl. 1959. Jugendinitiation und Sakraltänze der Baining. *Anthropos* 54. 905–938.
- Lehmann, Christian. 2001. Language documentation: A program. In Walter Bisang (ed.), *Aspects of typology and universals*, 84–97. Berlin: Akademie Verlag.
- Litteral, Robert. 2004. Vernacular education in Papua New Guinea. Paper commissioned for the EFA Global Monitoring Report 2005, The Quality Imperative.
- Lieven, Elena & Sabine Stoll. 2009. Language. In Marc H. Bornstein (ed.), *The handbook of cultural developmental science*, 143–160. New York and London: Psychology Press.
- Lüpke, Friederike. 2010. Research methods in language documentation. In Peter K. Austin (ed.), *Language documentation and description*, vol. 7, 55–104. London: SOAS.
- MacWhinney, Brian. 2000. *The CHILDES Project: Tools for analyzing talk*. Mahwah, NJ: Erlbaum.
- Marley, Alexandra. 2013. Language use amongst the Qaqet Baining: A sociolinguistic study of language choices in an ethnolinguistic minority in Papua New Guinea. Melbourne: unpublished La Trobe University MA thesis.
- Mayer, Mercer. 1969. *Frog, where are you?* New York: Dial Books for Young Readers.
- McConvell, Patrick & Felicity Meakins. 2005. Gurindji Kriol: A mixed language emerges from code-switching. *Australian Journal of Linguistics* 25(1). 9–30.
- Meakins, Felicity & Gillian Wigglesworth. 2013. How much input is enough? Correlating comprehension and child language input in an endangered language. *Journal of Multilingual and Multicultural Development* 34(2). 171–188.
- Mosel, Ulrike. 2006. Fieldwork and community language work. In Jost Gippert, Nikolaus

- P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 67–85. Berlin and New York: Mouton de Gruyter.
- Musgrave, Simon & Nicholas Thieberger. 2006. Ethical challenges in documentary linguistics. In Keith Allan (ed.), *Selected Papers from the 2005 Conference of the Australian Linguistic Society*. <http://www.als.asn.au>.
- Ochs, Elinor. 1979. Transcription as theory. In Elinor Ochs & Bambi B. Schieffelin (eds.), *Developmental pragmatics*, 43–72. New York: Academic Press.
- Ochs, Elinor. 1982. Talking to children in Western Samoa. *Language in Society* 11. 77–104.
- O'Shannessy, Carmel. 2008. Children's production of their heritage language and a new mixed language. In Jane Simpson & Gillian Wigglesworth (eds.), *Children's language and multilingualism*, 261–282. London and New York: Continuum International Press.
- O'Shannessy, Carmel. 2012. The role of code-switched input to children in the origin of a new mixed language. *Linguistics* 50(2). 305–340.
- Parker, Jim & Diana Parker. 1974. A tentative phonology of Baining (Kakat dialect). In Richard Loving (ed.), *Phonologies of four Papua New Guinea languages*, 5–43. Ukarumpa: Summer Institute of Linguistics.
- Parker, Jim & Diana Parker. 1977. Baining grammar essentials. Ukarumpa: Summer Institute of Linguistics. Unpublished manuscript.
- Pool, Gail. 2015. *Lost among the Baining: Adventure, marriage and other fieldwork*. Columbia: University of Missouri Press.
- Pool, Jeremy. 1984. Objet insaisissable ou anthropologie sans objet? Field Research among the Northern Baining, New Britain (1969–1970). *Journal de la Société des Océanistes* 40(79). 219–233.
- Rascher, Matthäus. 1900. Versuch zu einer Grammatik des Bainingischen. Unpublished manuscript.
- Rascher, Matthäus. 1904. Grundregeln der Bainingssprache. *Mitteilungen des Seminars für Orientalische Sprachen zu Berlin* 7(1). 31–85.
- Rice, Keren. 2006. Ethical issues in linguistic fieldwork: An overview. *Journal of Academic Ethics* 4(1–4). 123–155.
- Rice, Keren. 2011. Documentary linguistics and community relations. *Language Documentation & Conservation* 5. 187–207.
- Rohatynskyj, Marta A. 2000. The enigmatic Baining: The breaking of an ethnographer's heart. In Sjoerd R. Jaarsma & Marta A. Rohatynskyj (eds.), *Ethnographic artifacts: Challenges to a reflexive anthropology*, 174–194. Honolulu: University of Hawai'i Press.
- Rohatynskyj, Marta A. 2001. On knowing the Baining and other minor ethnic groups of East New Britain. *Social Analysis* 45(2). 23–40.
- Schieffelin, Bambi & Eleanor Ochs (eds.) 1986. *Language socialization across cultures*. Cambridge: Cambridge University Press.
- Seifart, Frank. 2008. On the representativeness of language documentations. In Peter K. Austin (ed.), *Language documentation and description*, vol. 5, 60–76. London: SOAS.
- Siegel, Jeff. 1996. *Vernacular education in the South Pacific*. Canberra: Australian Agency for International Development.

- Slobin, Dan I. 1982. Universal and particular in the acquisition of language. In Lila Gleitman & Eric Wanner (eds.), *Language acquisition: The state of the art*, 128–170. Cambridge: Cambridge University Press.
- Slobin, Dan I. 1985–1997. *The crosslinguistic study of language acquisition*, vol. 1–4. Mahwah, NJ: Erlbaum.
- Stoll, Sabine. 2009. Crosslinguistic approaches to language acquisition. In Edith L. Bavin (ed.), *The handbook of child language*, 89–104. Cambridge: Cambridge University Press.
- Stoll, Sabine & Elena Lieven. n.d. Child language (within the Chintang and Puma documentation project). Nijmegen: The Language Archive. <https://hdl.handle.net/1839/00-0000-0000-6F49-B@view>.
- Stebbins, Tonya N. 2011. *Mali (Baining) grammar: A language of the East New Britain Province, Papua New Guinea*. Canberra: Pacific Linguistics.
- Strömqvist, Sven & Ludo Verhoeven (eds.). 2004. *Relating events in narrative*, vol. 2. Mahwah, NJ: Erlbaum.
- Tomasello, Michael & Daniel Stahl. 2004. Sampling children's spontaneous speech: How much is enough? *Journal of Child Language* 31(1). 101–121.
- Wilkins, David. 1992. Linguistic research under aboriginal control: A personal account of fieldwork in central Australia. *Australian Journal of Linguistics* 12(1). 171–200.
- Wilkins, David. 2000. Even with the best of intentions: Some pitfalls in the fight for linguistic and cultural survival. In Francisco Queixales and Odile Renault-Lescure (eds.), *As linguas amazônicas hoje / Les langues d'Amazonie aujourd'hui*. Sao Paolo and Paris: Instituto Ambiental and IRD.