

Linguistic diversity in first language acquisition research: Moving beyond the challenges

First Language
2015, Vol. 35(4-5) 286–304
© The Author(s) 2015
Reprints and permissions:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/0142723715602350
fla.sagepub.com


**Barbara F. Kelly, William Forshaw,
Rachel Nordlinger and Gillian Wigglesworth**

University of Melbourne, Australia

Abstract

The field of first language acquisition (FLA) needs to take into account data from the broadest typological array of languages and language-learning environments if it is to identify potential universals in child language development, and how these interact with socio-cultural mechanisms of acquisition. Yet undertaking FLA research in remote field-based situations, where the majority of the world's languages are spoken and acquired, poses challenges for best-practice methodologies assumed in lab-based FLA research. This article discusses the challenges of child language acquisition research in fieldwork contexts with lesser-known, under-described languages with small communities of speakers. The authors suggest some modified approaches to methodology for child language research appropriate to challenging fieldwork situations, in the hope of encouraging more cross-linguistic acquisition research.

Keywords

Best practice, endangered languages, fieldwork, language acquisition, methodology, minority languages, remote communities

Introduction

The intersection of language, culture and cognition have long been a research focus in the unfolding process of children's language acquisition and in the speech directed to children (Altinkamış, Kern, & Sofu, 2014; Kulick & Schieffelin, 2004; Luo, Snow, & Chang,

Corresponding author:

Gillian Wigglesworth, School of Languages and Linguistics, University of Melbourne, Parkville, Victoria 3010, Australia.

Email: gillianw@unimelb.edu.au

2011; Ochs, 1982; Schieffelin & Ochs, 1986; Tomasello, 2000). The overarching aims of language acquisition research are to determine the biological dispositions for language, and to identify how these interact with socio-cultural mechanisms of acquisition. The seminal English-based work of Brown (1973) and colleagues has been foundational to the strong psychology-based underpinnings of language acquisition research, which focuses heavily on cognitive influences on language. However, as Bowerman (2011) has argued, the field also needs to take into account data from the broadest typological array of languages and language-learning environments. It is only by documenting acquisition in different languages and contexts that potential universals in child language and child directed speech can be identified, by comparing sequences of the acquisition process at the level of grammatical, lexical, conversational and narrative skill development. While there is a substantial body of cross-linguistic research in first language acquisition (FLA), the majority of this focuses on languages that are widespread (e.g. English, French, German, Japanese). There is no doubt that the field of FLA needs to substantially broaden its attention cross-linguistically if it is to reflect even a small proportion of the learning processes for the world's 6000–7000 languages (Bowerman, 2011).

While FLA researchers might welcome investigation of a broader array of languages, this is not without its challenges. Conducting child language research on diverse languages, particularly those where the adult language is also under-documented, typically entails data collection in fieldwork conditions. These conditions often make it difficult to follow the best-practice approaches to data collection which are commonly assumed in lab-based FLA research. In this article we begin with a brief discussion of best practice in data collection and analysis in FLA (Blom & Unsworth, 2010 and papers therein; Demuth, 1996; Ingram, 1989). We then consider issues that arise in the investigation of the acquisition of small languages spoken in so-called non-technological communities (Lieven & Stoll, 2013) and within the framework of language documentation (Eisenbeiss, 2006), and consider the modifications they require for best-practice methodology. We go on to discuss the modified approach we have taken in our research in Indigenous Australia as part of the Language Acquisition in Murrinhpatha (LAMP) project. A goal of this article is to contribute to an ongoing discussion about how to enable high quality FLA research that moves beyond traditional lab-based and relatively structured contexts; to encourage more researchers to do FLA research in fieldwork contexts, and help the field to account for a greater proportion of the world's linguistic diversity.

Typological diversity in FLA research

Linguistic typology and the documentation of languages that have hitherto not been examined in depth are burgeoning areas in linguistic research, but acquisition studies of such languages are relatively rare. This is despite the efforts of several researchers working on smaller and typologically different languages to develop guidelines or provide detailed discussion about their data collection (e.g. Lieven & Stoll, 2013) and the efforts of child language researchers to encourage typologists to expand their data collection foci to include children (e.g. Eisenbeiss, 2006).

Over many years child language researchers have demonstrated the range of markedly different language and socio-cultural child-raising practices, and a brief glance into the

past illustrates the impact such research has had on the field. Slobin (1982) laid out a range of typological comparisons in acquisition, focusing on what is universal and what is language-specific, resulting in the now classic series *The cross-linguistic study of language, Volumes I–V* (Slobin, 1985–1997). Ochs's (1982) and Schieffelin and Ochs's (1986) studies of child directed speech in lesser-known communities of Papua New Guinea and Western Samoa brought excitement and a new perspective to the field. These studies to a very large extent influenced developmental theory, for example by showing that western middle-class cultural practices such as modified input forms (then known as 'motherese') are neither universal nor necessary for children to learn their first language.

These works represented a foray into developmental research beyond the European and (to a lesser extent) Asian languages that had hitherto been the focus of the field. They highlighted the possibility and value of exploring smaller and less-studied languages. However, despite this, FLA research on lesser-known languages has remained relatively rare. There are exceptions, such as Demuth's (1989, 1992, 1998) extensive collection and research on Sesotho; Brown's longitudinal work on Tzeltal (Brown, 1993, 1998, 2014); Allen's extensive work on children's acquisition of Inuktitut (Allen, 1996, 1998, 2013); and Bavin's documentation of children's Warlpiri acquisition (Bavin, 1987, 1992, 2000, 2013). Overwhelmingly, however, the majority of FLA research has focused on the larger and more widespread languages.

More recently, a growing number of researchers are working on FLA in remote or relatively remote communities. These include: Papua New Guinea (e.g. Rumsey and Merlan's project on children's language learning and intersubjectivity in Ku Waru [Rumsey, 2013]); Alcock and colleagues' work on early Bantu acquisition, particularly passives in two languages of Kenya (Alcock, Rimba, & Newton, 2012); Stoll and Lieven and colleagues' work on the acquisition of Chintang in Nepal (Lieven & Stoll, 2013; Stoll & Lieven, 2014); Rose and Brittain's (2011) Cree language project; as well as projects elsewhere in the world as evidenced in this Special Issue. Such projects signify a shift in the field, and a growing awareness of the importance of FLA research covering the full range of the world's linguistic and cultural diversity. To capitalize on and help build this momentum we raise here the methodological challenges of such research, especially when it comes to meeting the established best-practice standards of lab-based FLA research. We argue that it is imperative that the field moves towards more methodological flexibility in order to enable FLA research to be conducted on a greater proportion of the world's languages.

Best-practice approaches in FLA research

Multiple factors underpin the collection and analysis of reliable and valid first language data using benchmarks considered to be best practice, which are essential to understanding how children learn language. However, what actually constitutes best practice can be difficult to identify since the field is varied in focus and comprises research across the humanities and social sciences, including linguistics, psychology, anthropology and sociology. Despite this, general guides have been published which provide insight into basic assumptions regarding data validity. Ingram's (1989) volume focuses on method, description and explanation in language development studies, and critiques a range of different approaches in terms of their methodological validity and subsequent

contribution to the development of child language theory. More recently, Blom and Unsworth's (2010) edited volume investigated experimental studies in child language across a variety of different research areas with each paper concluding with a list of 'do's and don'ts' related to the chapter's focus. These provide a useful summary of assumptions around FLA research methodology. Researchers have also touched on different methods of data collection, such as naturalistic and semi-structured research methods (e.g. Blom & Unsworth, 2010; Eisenbeiss, 2010) or investigated variation across different populations (e.g. Baker & van den Bogaerde, 2010; Paradis, 2010). While covering different topics, these papers, together with published research in the broader field, suggest some key expectations about methodology and analysis. These are discussed in the following sections.

Representative sample

An adequate sample size is necessary for comparability. Demuth (1996) recommends commencing a study with at least four children. A brief examination of cross-linguistic studies in FLA journals indicates that for experimental data and cross-sectional sampling, most current studies comparing development across ages have a minimum of around 12 participants per age group. In longitudinal sampling studies there are generally four to six children per age group. In the 1970s and early 1980s researchers were generally looking at two to three children per age group (e.g. Bowerman, 1973; Clancy, 1986), but more recent research appears to have moved towards slightly higher numbers (see for example, the Aboriginal Child Language Acquisition project where between eight and ten children were selected per community (Simpson & Wigglesworth, 2008)).

High demands on data requirements

Child language data collection differs in relation to the research questions being addressed or the hypotheses being tested. Data may be longitudinal, cross-sectional or both, and may be collected across a range of genre, interactional contexts and tasks. Underpinning any data collection is the expectation of high quality recordings with low variability in recording contexts (Eisenbeiss, 2010). It is also expected that data are collected at regular and frequent intervals, in particular in longitudinal studies, but also in cross-sectional studies where there is follow-up (Behrens, 2008).

Furthermore, ideally inter-coder reliability on transcription and coding is assessed (see Singleton and Supalla [2011] for discussion of inter-rater language coding) which means that multiple people need to have access to the children's language and to have the capacity to code the data.

Replicable methods and findings

Stemming primarily from the field of psychology, an underlying expectation of FLA research is the possibility of replicability. This is considered crucial both within a language, and across languages, since as a field we are focused on how children learn language with the ultimate goal of identifying what is universal versus specific, as well as

what is culturally mediated. Studying as large a range of languages as possible is important to understanding the relationship between the acquisition of linguistic forms, the development of cognition and the semantic notions encoded in the language (see Gathercole, 2006). However, in order for researchers to embark on comparative research they need to be able to replicate what others have done previously, and for this reason, detailed and replicable methodologies are required (De Houwer, 2009).

FLA research in the field

Researchers working in remote communities often struggle to meet the expectations of best practice in FLA research (see Kelly & Nordlinger [2014] for discussion). In these contexts, approaches taken to data collection, transcription and coding may require modifications to the protocols (discussed in the previous section) that are typical in either lab-based or more structured data collection contexts. Researchers such as Lieven and Stoll (2013) and Eisenbeiss (2006) have engaged with these concerns in some detail and have made modifications in their fieldwork situations.

Eisenbeiss (2006), writing primarily for fieldworkers in remote contexts, highlights the benefits of integrating the collection of child language data into any sort of language documentation project, arguing that child language data should minimally include samples of speech and experimental data, as well as staged communicative events such as songs or games or narratives. She also (Eisenbeiss, 2006, 2010) suggests that data should be collected from at least three children per age group because of individual variation, and be supplemented with chronological age-specific testing and a general measure of linguistic development, such as mean length utterance (Brown, 1973).

Lieven and Stoll (2013) also consider the position of FLA researchers who carry out research in very different linguistic and cultural environments. They propose a data collection scenario that may involve monitoring language use of several people beyond the target child across a variety of contexts and in spontaneous varied interactions.

In the remainder of this section we discuss a number of issues that arise for FLA researchers in fieldwork situations. We then consider modifications that different situations require in light of best-practice FLA research as outlined earlier. Then in the following section we discuss the methodological adjustments we have made in our own research in Indigenous Australia as part of the Language Acquisition in Murrinhpatha (LAMP) project.

Ethics and consent

Researchers working with human participants, especially children, are required to undergo ethics screening at an institutional level and to comply with the ethics guidelines generally based on medical or health protocols. These guidelines generally entail an expectation that researchers obtain informed consent. While this may be relatively straightforward in technological communities, FLA researchers report a range of different approaches for ensuring participants who may be unfamiliar with such practices have an understanding of what the study entails and what purposes it may be used for. These include obtaining written consent statements via local research assistants as in research

on Ngas in Nigeria (Childers, Vaughan, & Burquest, 2007) and Walmajarri in Australia (Moses, 2009). As with adult-oriented fieldwork, it is also possible to use time and date stamped verbal consent statements (Thieberger & Musgrave, 2007).

While expectations may be that consent for child participants is obtained from their parent, this may not always be the case, as cultural assumptions vary across communities. Community consent, generally given by a community elder or elders, may be required. This may be for individuals to participate in research and/or for researchers to undertake research more broadly in the community.

A further consideration in explaining consent to adult guardians may also relate to permissions with respect to who can see the data; this is particularly important where video is used, or in cultures where there are naming prohibitions in the case of deaths. It is therefore crucial that consent procedures for FLA research incorporate some discussion of these issues in culturally appropriate ways.

Participants and recruitment

In any FLA research the recruitment of participants can be difficult. In field-based research the challenges can be compounded by a number of factors. First, communities speaking lesser-known languages are often small. This can make it difficult to get the sample sizes that are considered best practice in cross-linguistic comparative work, especially across multiple age groupings (see earlier section on Best-practice approaches). And while there may only be a small pool of possible child candidates to study, there may also only be a small pool of adults able to assist with data transcription.

Furthermore, the nature of acquisition research in a remote community requires researchers to have close working relationships with a number of individuals, particularly young mothers and their children. This is due to the fact that in such a small community the researcher will interact with these people on a regular basis beyond the data collection process. Moreover, researchers will most likely need to work together with mothers and other adult language consultants on transcription and translation work. How these relationships are managed is central to the success of a project, and this may require considerable additional time spent on relationship-building before participants can be recruited, limiting the number of possible participants.

The field-based literature shows diversity in the number of participants, the groupings of ages studied, as well as the length of time over which children are studied. Some studies investigate language use in one child, as in Fortescue's (1984a, 1984b) study of a Greenlandic-learning child. Others have multiple children across different age cohorts, as in Rose and Brittain's (2011) study of Northeast Cree, Allen's (1996, 1998) study of Inuktitut and Stoll et al.'s (2012) Chintang language project. Others entail detailed recordings, often ethnographically situated, over larger participant numbers and over many years, such as Ochs's study of Western Samoan (1988), De León's research on Tzotzil (2007), Brown's work on Tzeltal (1998, 2014), Meek's (2010) work on Kaska and Demuth's (1989, 1992, 1998) work on Sesotho. The important contributions of all of this work to our understanding of language development cross-linguistically highlight the importance of flexibility in the FLA field around participant numbers in research.

Contexts and methods of data collection

The majority of recordings of child language data in the field focus on child–caregiver interaction (Stoll et al., 2012) and entail naturalistic data collection in which children interact in their own environment, such as Bavin’s (1987, 1992, 2000, 2013) study of Warlpiri and Stoll and Lieven’s (2014) study of cross-linguistic acquisition. They may also entail semi-structured data collection, such as Courtney’s (2006) study of Quechua, which sometimes utilized photographs and picture cards, and O’Shannessy’s (2008, 2012, 2013) study of Light Warlpiri, which employed context-appropriate picture stimulus book prompts. The rise in use of widely available wordless picture books such as *Frog, where are you?* (Mayer, 1969) allowed for a range of cross-linguistic comparisons of language development based on the same stimuli (Berman & Slobin, 1994; Stromqvist & Verhoven, 2004), although this entails some familiarity with western traditions of book use and story-telling. Beyond this, the advent of new technologies for development of experimental research materials and tools, as well as the public availability of these materials, has allowed researchers even greater scope to run cross-culturally and cross-linguistically appropriate experiments with relative ease. One such example is the suite of materials available through the Max Planck Institute (MPI) *Language and Cognition Field Manuals and Stimulus Materials* repository. These stimuli, for testing targeted language use, have been employed to explore several language development questions in fieldwork settings. For example, the Cut and Break Clips (Bohnemeyer, Bowerman, & Brown, 2001) have been used to directly compare the acquisition of aspects of argument structure cross-linguistically, as in K’iche’ Mayan and Mandarin (Pye, Loeb, & Pao, 1995).

However, researchers working in fieldwork situations need to be flexible as to the method of data collection employed, as different communities with different cultural and practical norms will mean that some well-established methods of data collection may be unsuccessful. We return to this issue in our discussion of the LAMP project later in the article.

Regularity of data collection and amounts of data

The widest diversity in field-based child language data is in the frequency and regularity of recording. This is not surprising given the challenges in recruiting participants in small communities together with the difficulty of access to some of these communities. Where the community is large enough, and it is possible to train and employ local language users to record, transcribe and annotate the data, then FLA fieldwork can yield large amounts of data not unlike those collected in urban societies, with regular recording intervals (e.g. Allen, 1996 for Inuktitut; Stoll et al., 2012 for Chintang; Demuth, 1992 for Sesotho). However, when communities are small and remote or otherwise difficult to access, datasets may be much smaller, and may entail variable recordings, such as Mithun’s (1989) reports for Mohawk for which she made fortuitous observations of several children while collecting adult data. Alternatively, recordings may be somewhat sporadic, as reported by Courtney and Saville-Troike (2002) for Navajo.

Furthermore, for researchers working in remote communities, cultural practices may lead to population movement. A researcher may start working with children, lay out a plan with parents for a longitudinal study, and then return to the community in which the children reside and find them absent at that time. While families are interested and in some cases eager to participate, cultural practices and traditions take precedence and dictate their movement in ways that may be in conflict with the needs of the research project.

Geographical and practical challenges

Communities in which lesser-known languages are spoken can often be remote, expensive and difficult to get to regularly, and challenging to live in long-term. Even in countries like Australia, where access should be relatively easy, particularly for researchers from the region, many researchers work in isolated communities or those which are cut off from main roads at certain times of the year, and therefore isolated much of the time (Bowern, 2008). This makes it very difficult for frequency and regularity in data collection and the type of longitudinal sampling required for acquisition studies requiring statistical measures. The expense and difficulty of getting to such field sites, even with relatively large populations, can mean that it is very difficult to have any sort of replicability of the research.

Furthermore, such remote communities often also present resource-related challenges such as a lack of electricity. Researchers working in language description and documentation often work in communities where there is no electricity, or no regular, stable electricity. Such environments generally lack state-of-the-art facilities, resulting in unstable recordings with lots of signal and background noise, which can be particularly difficult in child language recordings.

It is this set of challenges for FLA research in the field that provides the biggest impediment to the execution of best-practice principles as outlined in the earlier section, and it is also the area in which the modification options are most limited. In the next section we discuss some of the ways in which we have attempted to balance these issues as part of the LAMP project, but it is here where flexibility from the field at large is most needed to ensure that FLA research is encouraged in remote contexts despite the obstacles.

Endangered language contexts and language shift

With endangered languages, language may be undergoing rapid language shift so that earlier descriptions are no longer relevant, and there is no current description of the children's target (Simpson, 2008). This can mean that it is difficult to determine which behaviours are developmental and which result from the process of language shift (Meakins, 2011; O'Shannessy, 2008).

Where the target language is endangered, and the community under pressure of language shift, researchers often need to document the language change and analyse the language while the language is in a process of rapid transition. Alongside this, researchers need to be mindful that adult users of a language may have strong opinions regarding

which variety of the language is being examined and being transmitted. It can often be unclear whether the variety of the language that is considered in some way prestigious, either overtly or covertly, is the one that is being reported and documented. These are things that may not be known when going into a community, and although they are all general issues for linguistic field researchers, they can often confound the findings in developmental work if researchers are not aware of the need to mitigate against them.

A further factor that is often a challenge in remote communities is limited literacy, or whether the target language even has an established orthography. This means that it can be very difficult to undertake transcription and train assistants to develop the skills to assist in transcription. However, such helpers are crucial since children are difficult to understand even when adults share the target language, and transcription adds further complexity to the task of documenting language development.

Acquisition of Murrinhpatha (LAMP)

The previous sections have addressed what constitutes best practice in FLA research as well as highlighting key challenges for FLA research conducted in remote communities. Now we consider, at a practical level, how these benchmarks and challenges play out by focusing on a specific remote Australian Indigenous community context.

The LAMP project, begun in 2011, is investigating the acquisition of Murrinhpatha, a polysynthetic language spoken predominantly in the remote town of Wadeye (formerly Port Keats) in the Daly River region of Australia's Northern Territory. With approximately 2500 speakers, Murrinhpatha is the main language of day-to-day communication in Wadeye. For a traditional Australian Indigenous language, Murrinhpatha is unusual in that it is still being acquired by children as a first language. Today, fewer than 20 of the original 250–300 languages of Australia continue to be acquired by children as first languages (Kelly, Nordlinger, & Wigglesworth, 2010). While Murrinhpatha has been the focus of a growing amount of linguistic research (e.g. Blythe, 2009; Mansfield, 2014; Nordlinger, 2010; Street, 1987; Walsh, 1976), there is no comprehensive grammatical description and it is still best categorized as under-described. This brings obvious challenges for FLA research, since there is no comprehensive description of the target language that children are acquiring.

In general, there are low levels of literacy across the community, with high unemployment. People tend to live in large family groups, and the responsibility for childcare is spread across the extended family rather than seen only as the responsibility of one or two primary caregivers. With the high birth rate, 60% of speakers are aged under 24 years, and 40% aged under 14 years (Kelly et al., 2010) and older children routinely care for younger siblings. The local Catholic school operates a bilingual programme, transitioning to English after year 3. However, attendance is generally low and many of the enrolled children do not attend regularly.

Murrinhpatha is a highly polysynthetic language with complex verbal predicates that pose a number of interesting challenges for the child-learner (Forshaw et al., in press). Furthermore, it offers a rare (and diminishing) opportunity to study the acquisition of a polysynthetic Australian language. Thus, despite the fact that the adult language is still undergoing description itself, beginning in July 2012, a longitudinal corpus of

'semi-naturalistic' interactions focusing on the language development of six focus children was begun.

The LAMP project in Wadeye has faced many of the challenges outlined in the previous section. The town itself is very remote, lying at the end of a 200 km long dirt road that is impassable during the wet season (generally December to April). During this time the town is serviced by daily plane flights and a barge freight service, but it is not possible to drive in and out of the town, or into surrounding bush areas. Accommodation in the town is limited, and what exists is charged at exorbitant rates. This impacts on the time that researchers can spend in the field, with an obvious impact on the frequency and regularity with which participants can be recorded. Furthermore, the low literacy levels and the lack of opportunities for researchers to acquire the language outside of the field location mean that the transcription of the child language data can be enormously time-consuming as neither the researcher nor the adult native speaker participants have the requisite skills to undertake it alone.

Nonetheless, despite these challenges, the project (while still underway) has been very successful and has brought to light a number of interesting findings regarding the acquisition of complex verbal systems (Forshaw, forthcoming; Forshaw et al., in press) and language socialization in Wadeye (Kelly, Mansfield, Forshaw, Nordlinger, & Wigglesworth, 2014). In the following we outline the methodology we have adopted for the project and the various ways in which we have accommodated the challenges associated with FLA research in this remote location.

Building relationships

Some members of the research team had been working in the Wadeye community for a number of years, and so already had well-established relationships with local participants interested in linguistic research. We engaged one of these people as our primary Murrinhpatha research assistant (RA), who then facilitated discussions with community stakeholders about the study, assisted in the recruitment of the child participants and their families, explained the project in Murrinhpatha, and assisted in collecting ethical consent in the same way as reported by Moses (2009) for Walmajarri.

A particular challenge arose due to the fact that the student researcher (William) is male, and it is not culturally appropriate for young men and young women to work together one-on-one in the Murrinhpatha community. This made it potentially difficult for William to work alone with the young mothers of the recruited children. This was managed by ensuring that during recording and transcription sessions there was always at least one additional adult present. Initially this was the RA, but later, as participants became more relaxed, William was able to work together with two young women of a similar age to him.

In addition to the data collection, William spent a great deal of time interacting with participant families in a non-work environment. The more familiar the families became with William in the non-work environment, the more productive the work environment became. These interactions also provided valuable opportunities for William to learn more about the language and culture of the community, and for the community and participants to better understand the aims of the research.

In other studies in remote communities outside of Australia this relationship-building might be achieved by the researcher living with one of the participant study families (e.g. De León, 2007). However, this was not possible in this context because limited housing options in Wadeye mean that all available rooms tend to be used by local families, and also, more importantly, it would be culturally inappropriate for a researcher to live with a family. Houses are very private family domains used primarily for sleeping and it is only on rare occasions that non-family members enter these spaces. Consequently, recordings, whether for documentation or for acquisition purposes, are made either on the veranda, in public spaces, or on bush trips outside the community (see below).

Representative sampling

This study followed the language development of six children, which is currently a fairly standard number of children for a longitudinal study. A key difference, however, was that the children were aged from 1;9 to 4;3 at the beginning of data collection, and were followed over three years. Following Allen and Crago's (1996) Inuktitut recording method, this wider age range was selected to provide a broader picture of language development over several years given that this was the first study of its type in this community. While the drawback to this approach is that data at either end of the age scale, particularly at the beginning, are less abundant, the advantage is that it provides developmental information from 1;9 to around 7 years of age by the time data collection is completed.

Remote Indigenous communities in Australia tend to be very mobile for cultural reasons. The strong kinship system involves multiple commitments to family members and may mean travelling frequently to attend funerals, and other social and ceremonial occasions. In order to guard against anticipated high rates of attrition, 15 children were recorded during the first trip to ensure that a larger pool of children would be available to participate in the study if necessary. In this case, the anticipated high rate of attrition did not eventuate, and with the exception of one child, all children who participated in initial recordings have resided predominantly in Wadeye for the majority of the data collection period. We should note the main caregiver of each child was paid for the child's participation and bonuses were provided for the child being present for the next field trip as an incentive for children to remain in the study across multiple field trips.

Data collection

Recording environments are important. The participants need to feel comfortable in the recording location, and the location needs to allow for high quality video and audio recording. Furthermore, the makeup of participants needs to be a 'natural' but manageable combination.

In this study the majority of recordings involved small groups of between three and ten people. Two of the participants were focus children with the rest a mixture of family relations including parents, aunts, uncles, grandparents, siblings and cousins. Similarly to the recordings reported in Alcock et al.'s (2012) study of Kiswahili and Kirigiana, and Rabain-Jamin (2001) in Senegal, one-on-one interactions are not typical in everyday life in Wadeye, so such interactions between a child and a sole caregiver were deliberately

avoided. Recordings were made in the bushland surrounding Wadeye. Participants selected a location and travelled there by car with William. William provided participants with materials to have a picnic, go fishing or collect bush foods. Interactions were recorded using two Sony HXR-NX30P video cameras fixed on tripods, with one occasionally handheld. The focus children wore Sennheiser EW 112-p G3 wireless Lavalier microphones in backpacks, a successful method of data collection, as reported in Demuth (1996). This ensured they could move about freely but also achieved high quality recordings. It also makes it easier to identify the speech of focus children in group recordings.

The bush-based recording environment was selected for a number of reasons. First, families often undertake these sorts of trips on weekends. This environment was therefore well known to participants and involved an enjoyable activity that they commonly did together. Second, it meant that the number of participants in a recording was restricted (i.e. to the number that could fit in the car), which made both recording and transcription easier. When groups became too large they would often split into more than one conversation resulting in lots of overlap, which is difficult to transcribe. In addition, young children have fewer chances to produce language in the presence of large groups. Given that the amount of recordings made was limited it was important to ensure that children produced as much language as possible during that time.

The variety of bush locations meant that a variety of different yet comparable recording contexts were used, as suggested by Eisenbeiss (2006). For example, if a recording was made by a river, participants would fish, and if there were yams in the area these would be dug up, prepared and cooked. The main drawback to this type of recording was that it was susceptible to weather. This meant that this type of recording was not undertaken during the monsoon season, and was avoided on windy days since this greatly impacted audio quality, despite the a wind-deadening furry cover on the microphone. Recordings were typically done in the morning so as to avoid high temperatures.

Two other recording environments were trialled during this study, although neither of these proved to be as successful as the bush recordings, highlighting the importance of flexibility in data collection (as discussed earlier). The first was to construct a portable baby lab which was designed to promote semi-structured play. This was a large tent constructed in a room in town and lined with carpet to improve recording quality. Small groups of participants were given a selection of toys to play with. These included toys relating to scenarios with which the children were familiar, such as going to the shop or visiting the doctor. Although this environment was useful in introducing participants to the project, it was ultimately not feasible. The play scenario presented to the participants was quite foreign since families do not generally play with toys at home in the same way as western middle-class families do. A further drawback was that the baby lab was constructed in a shared room meaning that the lab had to be constructed and deconstructed for each recording.

The other option trialled was recording outside participants' homes, as has been done in other FLA cross-linguistic and cross-cultural studies (e.g. Lieven & Stoll, 2013). This was quickly abandoned because passers-by often engaged in multi-party conversations, which made transcription of the sessions difficult and, as mentioned above, meant the children and their caregivers were often not the primary speakers.

In addition to longitudinal data collection, some experimental data were collected. This involved a picture stimulus task which targeted a specific morphological structure that was largely absent in the naturalistic longitudinal data. This task was run with seven participants and showed that children even as young as 2;7 could indeed use these rarer forms when provided with the relevant context (Forshaw, 2014), highlighting the complementary nature of longitudinal and experimental data, as discussed by Eisenbeiss (2010).

Frequency and regularity of recording

While best practice, as described in the earlier section, requires recording at frequent and regular intervals (e.g. weekly, or monthly), this was not feasible in the LAMP project for a number of reasons. First, the remoteness of the community and the restrictions on accommodation for researchers limited the time that could be spent in the field to two trips of two to three months each, per year. Second, the bush-based recording method (outlined earlier), although highly successful for data collection, is also very time-consuming and therefore restricts the number of recording sessions that can be managed in each field trip. Furthermore, since all data transcription also needed to be completed while in the field (since this required both the researcher and the adult participants to work together, as outlined below), the researcher's time had to be carefully balanced between data collection and transcription during each field trip.

Thus, we modified best practice to ensure that each of the six focus children was recorded at least twice for between half an hour and an hour during each field trip (with the exception of one child who was recorded only once during one field trip). Ideally each child was recorded at least once near the beginning of a trip and once during the middle to allow time for transcription to be done while in the field and to ensure that all transcription was completed during the field trip while assistance was available. Subsequent recordings tended to be organized spontaneously in much the same way Mithun (1989) reports for her study of Mohawk.

Transcription and language familiarity

It is preferable that researchers have good familiarity with the target language being studied so as not to rely entirely on language consultants for translations, although some studies have relied on teams of literate native speaker language consultants to complete transcriptions (e.g. Rose & Brittain, 2011; Rumsey, 2013; Stoll et al., 2012). In the Murrinhpatha-learning context neither of these options was available. William had limited familiarity with the language but not enough language knowledge to transcribe conversation independently. Additionally, no highly literate bilingual Murrinhpatha/English speakers were available to do transcription work.¹

Instead, transcription relied on 'team transcription' common in language documentation contexts in which a linguist and native speaker work together to transcribe language data. This can be a time-consuming approach which tends to result in a transcription bottleneck where there is much more recorded language collected than is ever likely to be transcribed. For this reason, we employed Allen and Crago's (1996) model of

selecting sections for transcription which had a high density of potentially relevant productions. Although this process may result in a skewed corpus because it excludes sparser interactions, it was a necessary sacrifice in order to transcribe as much language data as possible.

The team of language consultants consisted predominantly of the mothers and grandmothers of the focus children. The mothers tended to be better at identifying what younger children were saying whereas the grandmothers, who had higher levels of literacy and English proficiency, were better at providing explanations and translations of more complicated constructions. Consequently these groups of consultants were employed in ways that best suited their skills, with mothers helping to transcribe younger children's language and grandmothers the older children's language. This involved having multiple recordings ready for transcription at all times, as it was difficult to predict who would be available for work, to ensure that transcription could progress regardless of who was available.

Reflexivity and flexibility

This case study highlights the essential aspects of fieldwork research required to achieve the best scientific outcomes when doing language acquisition research in remote communities with under-researched languages. Data collection methodologies need to be sensitive and reflective of the contexts in which they are being undertaken. Cultures, languages and communities vary, and FLA methodologies designed for specific contexts may at times result in data that are less comparable across studies. However, the alternative of not collecting such data risks not learning anything about how children acquire languages in remote and under-described communities and so the field needs to be flexible to ensure that research into these communities is not discouraged.

Fundamental to the success of this project was to build and maintain positive relationships with both study participants and the wider community, because this type of research is not possible without the participation of community members. The research methodology had to accommodate to people's lives and abilities, which dictated how recording and transcription was undertaken.

In order to accommodate the field situation and meet the requirements of flexibility and cultural sensitivity, our research has had to compromise on most of the methodological principles outlined in the section on best-practice approaches. The small number of literate participants has limited our chances of ensuring inter-coder reliability in transcription and coding. The bush-based recording sessions reduced the chances of high quality recordings with limited variation across the participants and the recording sessions, but have been vastly more successful in allowing for naturalistic language data from young children than the portable baby lab sessions trialled early in the research. The practicalities of the field site have reduced the frequency and regularity of recordings, yet we have longitudinal data for all focus children across a three-year period with at least two sessions recorded on each field trip (spaced evenly throughout the year). These compromises have been an unavoidable consequence of undertaking child language research in this community and, while not ideal, do not significantly undermine the quality of the data collected, nor their value to the field of FLA.

Moving beyond the challenges

For people working in language documentation and description the above discussion is neither new nor elucidating, but in the field of FLA, these issues have received only limited discussion. FLA researchers have the opportunity to approach remote communities in different ways, keeping a best-practice model in mind as the ideal, while acknowledging that what is achievable in one context will not necessarily be the case in another. Best practice is important, but the ideal of adhering only to best practice risks rejecting the sorts of findings that were so crucial in the 1970s in the ground-breaking work of Ochs and Schieffelin (1984), as well as those of many other more recent scholars undertaking FLA research with small and remote communities. Drawing on our own experiences and those of others working in similar contexts, we suggest that there is value in modified datasets. Even with minimal data, it is still possible to both inform and build upon our current models of acquisition and to make valuable contributions to cross-linguistic theories of language acquisition, while forming a basis for future work.

Language acquisition research across the broadest possible range of languages is crucial to the goal of understanding what potential universals might exist, and how they interact with acquisition. The study and dissemination of research into lesser-known, under-described languages with small communities goes some way towards this aim and it is encouraging that such research is being undertaken while we have the opportunity and while such languages are still being transmitted. We hope that this article will encourage more researchers to embark on this task, despite the challenges, and encourage the FLA field more broadly to bear in mind that '[b]etween the best-laid schemes of a researcher and their execution falls the shadow of what is possible in the field' (Moses, 2009, p. 58).

Acknowledgements

For helpful feedback on earlier versions of this article that have led to substantial revision and improvement, we are grateful to Evan Kidd, Kevin Durkin and two anonymous reviewers.

Funding

Our work on language acquisition in Murrinhpatha has been supported by the Australian Research Council (DP110100961 'From little things big things grow') and the ARC Centre of Excellence for the Dynamics of Language.

Note

1. This is not because there are no people who could complete this task, but because such people are in high demand in Wadeye and are consequently not regularly available for language work.

References

- Alcock, K., Rimba, K., & Newton, C. (2012). Early production of the *passive* in two Eastern Bantu languages. *First Language*, 32, 459–478.
- Allen, S. (1996). *Aspects of argument structure acquisition in Inuktitut*. Amsterdam, The Netherlands: John Benjamins.

- Allen, S. (1998). Categories within the verb category: Learning the causative in Inuktitut. *Linguistics*, 36, 633–677.
- Allen, S. (2013). The acquisition of ergativity in Inuktitut. In E. L. Bavin & S. Stoll (Eds.), *The acquisition of ergativity* (pp. 71–105). Amsterdam, The Netherlands: John Benjamins.
- Allen, S., & Crago, M. (1996). Early passive acquisition in Inuktitut. *Journal of Child Language*, 23, 129–156.
- Altunkamiş, N. F., Kern, S., & Sofu, H. (2014). When context matters more than language: Verb or noun in French and Turkish caregiver speech. *First Language*, 34, 537–550.
- Baker, A., & van den Bogaerde, B. (2010). Measuring the linguistic development of deaf learners. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 245–268). Amsterdam, The Netherlands: John Benjamins.
- Bavin, E. (1987). Anaphora in children's Warlpiri. *Australian Review of Applied Linguistics*, 10(2), 1–11.
- Bavin, E. (1992). The acquisition of Warlpiri. In D. I. Slobin (Ed.), *The cross-linguistic study of language acquisition* (Vol. 3, pp. 309–371). Hillsdale, NJ: Lawrence Erlbaum.
- Bavin, E. (2000). Ellipsis in Warlpiri: An analysis of frog stories. *Linguistics*, 38, 568–589.
- Bavin, E. (2013). The acquisition of ergative case in Warlpiri. In E. Bavin & S. Stoll (Eds.), *The acquisition of ergativity* (pp. 107–131). Amsterdam, The Netherlands: John Benjamins.
- Behrens, H. (Ed.). (2008). *Corpora in language acquisition research: History, methods, perspectives* (Vol. 6). Amsterdam, The Netherlands: John Benjamins.
- Berman, R., & Slobin, D. (1994). *Relating events in narrative*. Hillsdale, NJ: Lawrence Erlbaum.
- Blom, E., & Unsworth, S. (Eds.). (2010). *Experimental methods in language acquisition research* (Vol. 27). Amsterdam, The Netherlands: John Benjamins.
- Blythe, J. (2009). *Doing referring in Murriny Patha conversation* (Unpublished PhD thesis). University of Sydney, Sydney, New South Wales, Australia.
- Bohnenmeyer, J., Bowerman, M., & Brown, P. (2001). Cut and break clips. In S. Levinson & N. Enfield (Eds.), *Manual for the field season 2001* (pp. 90–96). Nijmegen, The Netherlands: Max Planck Institute for Psycholinguistics.
- Bowerman, M. (1973). *Early syntactic development: A cross linguistic study with special reference to Finnish*. Cambridge, UK: Cambridge University Press.
- Bowerman, M. (2011). Linguistic typology and first language acquisition. In J. J. Song (Ed.), *The Oxford handbook of linguistic typology* (pp. 591–617). Oxford, UK: Oxford University Press.
- Bowern, C. (2008). *Linguistic fieldwork: A practical guide*. Basingstoke, UK: Palgrave Macmillan.
- Brown, P. (1993). The role of shape in the acquisition of Tzeltal (Mayan) locatives. In E. V. Clark (Ed.), *Proceedings of the 25th annual Child Language Research Forum* (pp. 211–220). Stanford, CA: CSLI/University of Chicago Press.
- Brown, P. (1998). Children's first verbs in Tzeltal: Evidence for an early verb category. *Linguistics*, 36, 713–753.
- Brown, P. (2014). The interactional context of language learning in Tzeltal. In I. Arnon, M. Casillas, C. Kurumada & B. Estigarriba (Eds.), *Language in interaction: Studies in honor of Eve V. Clark* (pp. 51–82). Amsterdam, The Netherlands: John Benjamins.
- Brown, R. (1973). *A first language: The early stages*. Cambridge, MA: Harvard University Press.
- Childers, J., Vaughan, J., & Burquest, D. (2007). Joint attention and word learning in Ngas-speaking toddlers in Nigeria. *Journal of Child Language*, 34, 199–225.
- Clancy, P. (1986). The acquisition of Japanese communicative style. In E. Ochs & B. Schieffelin (Eds.), *Language acquisition and socialization across cultures* (pp. 213–250). New York, NY: Cambridge University Press.
- Courtney, E. (2006). Adult and child production of Quechua relative clauses. *First Language*, 26, 317–338.

- Courtney, E., & Saville-Troike, M. (2002). Learning to construct verbs in Navajo and Quechua. *Journal of Child Language*, 29, 623–654.
- De Houwer, A. (2009). *Bilingual first language acquisition*. Bristol, UK: Multilingual Matters.
- De León, L. (2007). Parallelism, metalinguistic play, and the interactive emergence of Zinacantan Mayan siblings' culture. *Research on Language and Social Interaction*, 40, 405–436.
- Demuth, K. (1989). Maturation and the acquisition of the Sesotho passive. *Language*, 65, 56–80.
- Demuth, K. (1992). Acquisition of Sesotho. In D. Slobin (Ed.), *The cross-linguistic study of language acquisition* (Vol. 3, pp. 557–638). Hillsdale, NJ: Lawrence Erlbaum.
- Demuth, K. (1996). Collecting spontaneous production data. In D. McDaniel, C. McKee & H. S. Cairns (Eds.), *Methods for assessing children's syntax* (pp. 3–22). Cambridge, MA: MIT Press.
- Demuth, K. (1998). Argument structure and the acquisition of Sesotho applicatives. *Linguistics*, 36, 781–806.
- Eisenbeiss, S. (2006). Documenting child language. In P. K. Austin (Ed.), *Language documentation and description* (Vol. 3, pp. 106–140). London, England: SOAS.
- Eisenbeiss, S. (2010). Production methods in language acquisition research. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 11–34). Amsterdam, The Netherlands: John Benjamins.
- Forshaw, W. (2014, July). *Little kids big paradigms*. Poster presented at the International Association for the Study of Child Language Conference, Amsterdam, The Netherlands.
- Forshaw, W. (forthcoming). *Little kids, big verbs: Acquisition of Murrinhpatha bipartite stem verbs* (Unpublished PhD thesis). University of Melbourne, Parkville, Victoria, Australia.
- Forshaw, W., Davidson, L., Kelly, B., Nordlinger, R., Wigglesworth, G., & Blythe, J. (in press). The acquisition of Murrinhpatha (Northern Australia). In M. Fortescue, M. Mithun & N. Evans (Eds.), *Oxford handbook of polysynthesis*. Oxford, UK: Oxford University Press.
- Fortescue, M. (1984a). *West Greenlandic*. Beckenham, UK: Croom Helm.
- Fortescue, M. (1984b). Learning to speak Greenlandic: A case study of a two-year-old's morphology in a polysynthetic language. *First Language*, 5, 101–112.
- Gathercole, V. (2006). Introduction to Special Issue: Language-specific influences on acquisition and cognition. *First Language*, 26, 5–17.
- Ingram, D. (1989). *First language acquisition: Method, description and explanation*. Cambridge, UK: Cambridge University Press.
- Kelly, B., Mansfield, J., Forshaw, W., Nordlinger, R., & Wigglesworth, G. (2014, December 10). *The use of speech acts in Murrinhpatha language socialisation*. Paper presented at the 2014 Australian Linguistics Society Conference, Newcastle, New South Wales, Australia.
- Kelly, B., & Nordlinger, R. (2014). Fieldwork and first language acquisition. In L. Gawne & J. Vaughan (Eds.), *Selected papers from the 44th Conference of the Australian Linguistic Society, 2013*. Retrieved from <http://bit.ly/ALS2013Proceedings>
- Kelly, B., Nordlinger, R., & Wigglesworth, G. (2010). Indigenous perspectives on the vitality of Murrinh-Patha. In Y. Treis & R. De Busser (Eds.), *Selected papers from the 2009 Conference of the Australian Linguistic Society*. Retrieved from <http://www.als.asn.au>
- Kulick, D., & Schieffelin, B. (2004). Language socialization. In A. Duranti (Ed.), *A companion to linguistic anthropology* (pp. 349–368). Malden, MA: Blackwell.
- Lieven, E., & Stoll, S. (2013). Early communicative development in two cultures: A comparison of the communicative environments of children from two cultures. *Human Development*, 56, 178–206.
- Luo, Y. H., Snow, C. E., & Chang, C. J. (2011). Mother–child talk during joint book reading in low-income American and Taiwanese families. *First Language*, 32, 494–511.

- Mansfield, J. (2014). *Polysynthetic sociolinguistics: The language and culture of Murrinh Patha youth* (Unpublished PhD thesis). Australian National University, Canberra, Australian Capital Territory, Australia.
- Mayer, M. (1969). *Frog, where are you?* New York, NY: Dial Press.
- Meakins, F. (2011). *Case marking in contact: The development and function of case morphology in Gurindji Kriol*. Amsterdam, The Netherlands: John Benjamins.
- Meek, B. (2010). *We are our language: An ethnography of language revitalization in a Northern Athabascan community*. Tucson, AZ: University of Arizona Press.
- Mithun, M. (1989). The acquisition of polysynthesis. *Journal of Child Language*, 16, 285–312.
- Moses, K. (2009). *How do dinosaurs hug in the Kimberley?: The use of questions by Aboriginal caregivers and children in a Walmajarra community* (Unpublished PhD thesis). University of Melbourne, Parkville, Victoria, Australia.
- Nordlinger, R. (2010). Verbal morphology in Murrinh-Patha: Evidence for templates. *Morphology*, 20, 321–341.
- Ochs, E. (1982). Talking to children in Western Samoa. *Language in Society*, 11, 77–104.
- Ochs, E. (1988). *Culture and language development: Language acquisition and language socialization in a Samoan village*. Cambridge, UK: Cambridge University Press.
- Ochs, E., & Schieffelin, B. (1984). Language acquisition and socialization: Three developmental stories. In R. Shweder & R. LeVine (Eds.), *Culture theory: Mind, self, and emotion* (pp. 276–320). Cambridge, UK: Cambridge University Press.
- O'Shannessy, C. (2008). Children's production of their heritage language and a new mixed language. In J. Simpson & G. Wigglesworth (Eds.), *Children's language and multilingualism* (pp. 261–282). London, England: Continuum International Press.
- O'Shannessy, C. (2012). The role of code-switched input to children in the origins of a new mixed language. *Linguistics*, 50, 305–340.
- O'Shannessy, C. (2013). The role of multiple sources in the formation of an innovative auxiliary category in Light Warlpiri, a new Australian mixed language. *Language*, 89, 328–353.
- Paradis, J. (2010). Comparing typically-developing children and children with specific language impairment. In E. Blom & S. Unsworth (Eds.), *Experimental methods in language acquisition research* (pp. 223–244). Amsterdam, The Netherlands: John Benjamins.
- Pye, C., Loeb, D., & Pao, Y.-Y. (1995). The acquisition of breaking and cutting. In E. Clark (Ed.), *Proceedings of the Twenty-Seventh Annual Child Language Research Forum* (pp. 227–236). Stanford, CA: CSLI Publications.
- Rabain-Jamin, J. (2001). Language use in mother–child and young sibling interactions in Senegal. *First Language*, 21, 357–385.
- Rose, Y., & Brittain, J. (2011). Grammar matters: Evidence from phonological and morphological development in Northern East Cree. In M. Pirvulescu, M. C. Cuervo, A. Pérez-Leroux, J. Steele & N. Strik (Eds.), *Selected proceedings of the 4th Conference on Generative Approaches to Language Acquisition North America* (pp. 193–208). Somerville, MA: Cascadilla Press.
- Rumsey, A. (2013). Intersubjectivity, deception and the 'opacity of other minds': Perspectives from Highland New Guinea and beyond. *Language and Communication*, 33, 326–343.
- Schieffelin, B., & Ochs, E. (Eds.). (1986). *Language socialization across cultures*. Cambridge, UK: Cambridge University Press.
- Simpson, J. (2008). Language landscapes of children in remote Australia. In R. Amery & J. Nash (Eds.), *Inaugural Indigenous Languages Conference* (pp. 51–60). Adelaide, South Australia, Australia: University of Adelaide.
- Simpson, J., & Wigglesworth, G. (Eds.). (2008). *Children's language and multilingualism: Indigenous language use at home and school*. London, England: Continuum International.

- Singleton, J. L., & Supalla, S. (2011). Assessing children's proficiency of natural signed languages. In M. Marschark & P. Spencer (Eds.), *Oxford handbook of deaf studies, language, and education* (2nd ed., pp. 306–321). New York, NY: Oxford University Press.
- Slobin, D. (1982). Universal and particular in the acquisition of language. In L. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art* (pp. 128–170). Cambridge, UK: Cambridge University Press.
- Slobin, D. (1985–1997). *The cross-linguistic study of language* (5 Vols.). Hillsdale, NJ: Lawrence Erlbaum.
- Stoll, S., Bickel, B., Lieven, E., Paudyal, N., Banjade, G., Bhatta, T., & Rai, N. (2012). Nouns and verbs in Chintang: Children's usage and surrounding adult speech. *Journal of Child Language*, 39, 284–321.
- Stoll, S., & Lieven, E. (2014). Studying language acquisition cross-linguistically. In H. Winkler & P. Pradakannaya (Eds.), *South and Southeast Asian psycholinguistics* (pp. 19–35). Cambridge, UK: Cambridge University Press.
- Street, C. S. (1987). *Introduction to the language and culture of the Murrinh-Patha*. Darwin, Northern Territory, Australia: Summer Institute of Linguistics.
- Stromqvist, S., & Verhove, L. (2004). *Relating events in narrative, Vol. 2: Typological and contextual perspectives*. Hillsdale, NJ: Lawrence Erlbaum.
- Thieberger, N., & Musgrave, S. (2007). Documentary linguistics and ethical issues. In P. Austin (Ed.), *Language documentation and description* (Vol. 4, pp. 26–36). London, England: SOAS.
- Tomasello, M. (2000). Culture and cognitive development. *Current Directions in Psychological Science*, 9, 37–40.
- Walsh, M. (1976). *The Murinypata language* (Unpublished PhD thesis). Australian National University, Canberra, Australian Capital Territory, Australia.