Sampling Signers in Providence Island: Reflections on a Small-Scale Documentation Project

El muestreo de señantes en la Isla de Providencia: reflexiones sobre un proyecto de documentación a pequeña escala

Rehana Omardeen
University of Göttingen – Germany

ABSTRACT

Providence Island Sign Language is an indigenous sign language of Colombia, used in the Caribbean Island of Providence (Spanish: Providencia). First described in the 1960s and 1970s, the sign language has received little follow up research. In this paper I introduce a small-scale documentation project of PISL that began in 2019 and is ongoing. I describe the aim, methods and output of this project and reflect on data collection, particularly with respect to sampling members of the language community. This paper contributes to the ongoing discussions around best practices in sign language documentation, by adding a case study focused on sampling in a micro-community sign language.

Keywords: Providence Island Sign Language; methodology; sampling; micro-community sign language; meta-documentation.

RESUMEN

La lengua de señas de la isla de Providencia (PISL) es una lengua de señas indígena colombiana usada en la isla caribeña de Providencia. Descrita por primera vez en los años sesenta y setenta, esta lengua ha sido poco investigada. En este artículo presento un proyecto de documentación a pequeña escala de la PISL que comenzó en 2019 y aún está vigente. Describo el objetivo, los métodos y los resultados del proyecto y hago una reflexión sobre la recopilación de datos, en particular con respecto al muestreo de los miembros de la comunidad señante. Este documento contribuye a las discusiones en curso sobre la implementación de mejores prácticas en la documentación lenguas de señas, al adicionar un estudio de caso centrado en el muestreo en una micro comunidad de señantes.

Palabras clave: Lengua de Señas de la isla de Providencia; metodología; muestreo; lengua de señas micro comunitaria; meta documentación.

1 Master of Arts Radboud University (The Netherlands). PhD candidate. rehana.omardeen@uni-goettingen.de
1. Introduction

There is a bias in sign language research: the field is dominated by languages used at the national level across large communities, also referred to as macro-community sign languages. On the other hand, languages used on a smaller scale, among specific culturally or geographically bound groups, also referred to as micro-community sign languages, are much less-studied and documented. As linguists become more interested in documenting and describing micro-community sign languages, however, it becomes increasingly clear that data collection methods designed for macro-community settings are not one-size-fit-all. The widely divergent sociolinguistic settings found across different signing communities have resulted in quite different methods being used in different projects.

One of the clearest examples of this is in the domain of sampling: the little literature that does exist that discusses who to sample highlights the diversity found across different projects. This is because the core question of sampling - who forms a representative portion of the signing community? - is addressed differently across settings. Take the example of audiological status. Deaf people who use sign language as their primary means of communication are at the centre of any signing community. In the case of many macro-community corpus projects, the sampling begins and ends with these deaf signers. Projects even often concentrate solely on deaf signers with deaf parents, who are considered the most fluent language users (Fenlon et al., 2015). Other members of any signing community are hearing signers: friends, family members and other people who use sign language to communicate with deaf people. For micro-community sign languages, these signers are also often included in data collections (De Vos, 2016; Nyst, 2015), while they are rarely, if ever included in macro-community samples.

Zooming in on individual communities further reveals the complexity of sampling. This is particularly noticeable in micro-community settings where the small size of the community allows a clear view of the tremendous diversity contained within. A good example is the case of Providence Island Sign Language, a sign language isolate used on a small Caribbean Island by its 13 deaf inhabitants and their fellow hearing islanders. The deaf residents of Providence range from very old to very young, some prefer to sign and some prefer to speak, some have lived in the island their whole life and others have migrated from or spent considerable time abroad, for some PISL is their only sign language and for others it is one of multiple. Hearing islanders sign with varying levels of confidence and proficiency and are most familiar with the lexical sign variants used by their closest deaf contact. These various people all belong to some extent to the signing community, but settling on and filming a representative sample of this community is no straightforward task.

This paper provides a meta-documentation (Austin, 2013) of a small sign language documentation project in Providence Island. I discuss the theoretical and practical concerns encountered during the creation of a corpus of Providence Island Sign Language (PISL), presented as a case study of sampling in a micro-community sign language. Section 2 introduces the signing community and outlines previous research. Section 3 introduces the aims, methods and output of the PISL Documentation project including a brief overview of the resulting dataset. Section 4 provides in depth reflections about the challenges of sampling and reflects on how
these challenges relate to other sign language documentation projects. Section 5 concludes and offers future directions for PISL documentation.

2. Providence Island Sign Language

2.1 Demographic Sketch

Providence is part of the department of San Andres, Providence and Santa Catalina (see Figure 1), a culturally, ethnically and geographically distinct region of Colombia. The population of Providence is relatively small, currently estimated at 4500 people in 2018 census (DANE, 2018), despite significant increase in the last 50 years. They mostly comprise of Raizales, a minority ethnic group descended from settlers and slaves brought to the islands in the 1730s (Bartens, 2011; Washabaugh, 1983). The island is 17km² and is mainly populated on the coast, with the mountainous interior used for farming. Historically, villages were relatively isolated, however travel within the island was transformed in 1961 with the creation of a ring road connecting the coastal villages (Washabaugh, 1979). Inter-island travel is common and residents of the two smaller islands often visit San Andres to see family, receive medical care and do shopping.

![Figure 1: The location of the archipelago of San Andres, Providence & Santa Catalina.](image)

Providence is a multilingual place where several languages are used in day-to-day life. There are three major spoken languages used in the island. An English-lexifier creole (termed Islander Creole English in the literature) is the primary language of everyday use in Providence (Bartens, 2013). Spanish is commonly used in official settings, as a result of pressures for national assimilation from the central government (Ramírez-Cruz, 2017). English is found in religious settings, especially in some churches (García León & García León, 2019; Ramírez-Cruz, 2017). In this multilingual setting, the creole is the dominant language, however there is frequent borrowing and code mixing with Spanish (Bartens, 2013), especially among younger generations.

In addition to these spoken languages, there is also a sign language in use that has emerged on the island due to sustained high prevalence of deafness. The island’s history of deafness traces back to the late 1800s with the birth of the first deaf islander (Washabaugh, 1986: 18). Since then, the number and proportion of deaf islanders have fluctuated: a high of 20 deaf people comprising 0.66% of the
population was recorded in the 1980s (Washabaugh, 1986: 9) and followed by a gradual decrease to 17 deaf people comprising 0.35% of the population in the 1990s (Lattig et al., 2008) and today there are 13 deaf islanders, who make up 0.28% of the island’s population and living distributed around the villages of the island (see Figure 2). There are multiple causes of deafness in Providence, including sporadic non-genetic deafness, as well as at least two distinct genetic sources (Lattig et al., 2008). This may make the future prevalence of deafness, and in turn the continued existence of the sign language, particularly sensitive to demographic and cultural changes (Braithwaite, 2019).

The 13 deaf islanders range from 8 to 81 years old and have diverse backgrounds and lived experiences. Some have lived in Providence their whole lives, others have lived in San Andres for an extended period of time, and at least three younger deaf people have lived on the mainland (in Panama and Colombia). Deaf people live in villages with both historically many and historically few deaf inhabitants, and there are three living members of a multigenerational deaf family. Preferred communication practices also vary from person to person: some deaf people prefer speaking to communicate with hearing people, whereas others primarily rely on signing. Some signers have broad communication networks, spanning across the whole island (and sometimes the neighbouring island) due to their occupation or family ties, others interact mainly with their immediate family.

One unifying experience is that deaf people live and interact mostly with hearing relatives and neighbours, a feature that shapes the signing landscape of the island. Unlike in macro-community sign settings, in Providence there are no clubs or schools specialized for deaf people to learn and use PISL with one another. There are some strong individual relationships among deaf people due to age, family relationships and geographic proximity. However, deaf people sign most often with
hearing people, and signing is used across multiple communication networks that centre on deaf-hearing interactions (Washabaugh, 1979). Hearing signers have varying degrees of signing proficiency, and those who share a household with a deaf person have been described as the most skilled signers (Washabaugh, 1979). Their signing is often most closely related to their deaf family member: they use the same lexical variants (Washabaugh, 1986: 52) and sometimes report less success in communication with other deaf people. Notably, there is smooth communication among most deaf people themselves, despite variation across communication networks.

2.2. Linguistic Research

Providence Island Sign Language was first described in the 1960s and 1970s in a series of works by creolist William Washabaugh and colleagues (Washabaugh, 1979, 1979, 1980b, 1980a, 1981, 1986; Washabaugh et al., 1978, 1978; Woodward, 1979). PISL was one of the first micro-community sign languages to ever be researched, and writings from this period established several typological differences between PISL and well-studied macro-community sign languages of the time. For example, Washabaugh found PISL to use considerably more non-manual elements in sign formation than the macro-community sign language American Sign Language (1986: 56), and even to have some purely nonmanual signs (1978: 98). Woodward (1978) found PISL to have among the fewest signs in the domain of kinship in a comparison of 20, mostly macro-community sign languages. PISL signers were also shown to often point to things in the real world to refer to associated concepts (Washabaugh et al., 1978). Signers used this strategy, known as metonymic pointing, for various types of referents including colours, places and people – a notably different referential system than those that had been documented in macro-community sign languages.

PISL has received a great deal of academic attention for such a small language, yet, most observations about its uniqueness were made at a time when there was little typological research on sign languages. Many of these claims arose from comparing the structural and sociolinguistic features of PISL with languages from very different ecological niches, that is national, institutionalized macro-community sign languages. Since then, many more sign languages around the world have been described by linguists. As more micro-community languages are documented, they have been shown to share various typological features thought to be unique to PISL, such as metonymic pointing and small sets of kinship signs (de Vos & Pfau, 2015).

Furthermore, recent research has also revealed that many hallmark features of PISL are also frequently used and play an important role in both macro- and micro-community sign language alike; for example non-manual elements are shown to be important features of sign formation in German Sign Language, a macro-community sign language (Pendzich, 2020) and Kata Kolok, a micro-community sign language (Lutzenberger, 2018; Marsaja, 2008).

Against the backdrop of a growing interest in documenting non-institutionalized micro-community signing settings (Braithwaite, 2020a; Kusters, 2010), new research on the signing situation in Providence has emerged. This work has focused largely on documenting and updating the sociolinguistic situation. Two recent MA theses...
explore the sociolinguistic situation: Hooker O’Neil (2016) focuses on family life and the communicative practices of individuals, while Cortés Bello (2016; 2019) discusses social networks of deaf signers. Most recently, Braithwaite (2020c) describes the evolving sociolinguistic situation on the island. Within these projects, there has also been some basic lexical elicitation conducted among deaf islanders (Braithwaite, 2020b; Cortés Bello, 2016). In addition to the strong sociolinguistic focus, all studies so far have been based on small, elicited datasets.

The existing body of work on PISL presents a rich potential for in-depth follow-up research, and much remains to be explored. Follow-up studies can evaluate typological claims by collecting a well-balanced diverse set of data from signers, allowing for a comparison of elicited and spontaneous signing. Research so far has largely been based on elicitation and field observations; through this many claims about PISL have been generated which can be empirically examined with more spontaneous data. Furthermore, new data from signers who participated in previous research can provide a diachronic snapshot that can be compared to original recordings from the 1970s. This is a unique opportunity to examine language change over the lifetime: as most micro-community sign languages have a relatively short history of documentation, whereas in the case of Providence it is possible to look at an individual’s language use at two points 40 years apart.

Finally, with recent social and demographic changes in the island, the future of Providence’s signing tradition is uncertain. The number of deaf people on the island has shrunk considerably both in number and in proportion of the population over the past 50 years and it is possible that changing demographics on the island combined with genetic counselling may radically alter the continued incidence of deafness on the island (Braithwaite, 2020c). Furthermore, there is mounting pressure among islanders (deaf and hearing) to receive education in the mainland: as a result, three young deaf people have already spent significant time in Colombia and Panama to seek formal education.

Increase both in signed multilingualism among young deaf people and awareness of national sign languages among hearing people appear to be contributing to a shift in attitude towards the PISL. Given these factors, there is a clear and time sensitive motivation to collect data while the signing tradition is still alive, in order to capture a unique aspect of the island’s cultural and linguistic heritage.

3 The PISL Documentation Project

In 2019, supported by the Endangered Languages Documentation Program, a small research team conducted a three-month field-based data collection as part of a project to document Providence Island Sign Language. The core research team was made up of three members: (i) myself (Rehana Omardeen), a hearing sign linguist; (ii) Ian Dhanooolal, a deaf sign language teacher and activist and (iii) Carlos Newball, a local deaf islander. In addition to this core team, other individuals were involved in data collection primarily through conducting interviews. They were: (i) Ben Braithwaite, a hearing linguist at the University of the West Indies who had previously visited the island, (ii) Maureen Hooker O’Neill, a hearing researcher and audiologist from San Andres, and (iii) Carmelina Newball, the hearing mother of Carlos who runs the local...
cultural centre, the Casa de la Cultura. I include additional biographical information for each of the core team members below.

I, Rehana Omardeen, am a hearing Trinidadian and researcher trained in sign language linguistics. At the time of the project, I had just begun working on my PhD at the University of Göttingen, Germany. I was introduced to sign languages via studying linguistics and had basic signing skills in American Sign Language with some exposure to Sign Language of the Netherlands and International Sign.

Ian Dhanoolal is a deaf Trinidadian researcher and community activist working as a sign language teacher in Trinidad at the University of the West Indies. Ian had experience in Providence from a previous short research trip in 2016 alongside researcher Ben Braithwaite. Ian signs Trinidad and Tobago Sign Language and American Sign Language and has a great deal of multilingual competence from research within the Caribbean region.

Carlos Newball is a deaf islander from Providence who was recruited as a collaborator in the field. Carlos was involved in the previous research trip in 2016 and thus had some experience with language documentation. At the time of the project, Carlos signed mostly PISL, and had some exposure to Colombian Sign Language from friends and the internet.

The primary aim of the project was to create and publish a rich and varied dataset of signing practices in Providence that could be used for linguistic research. In particular, we wanted to collect data that could form the basis for both describing aspects of PISL and examining previous claims about the language. We also wanted to collect data that could serve as a basis for lexical database of Providence signs. In addition, we hoped that the resulting dataset would function as a basis for community resources such as dictionaries, and in doing so help to preserve a critical aspect of the island’s cultural and linguistic heritage.

We collected 15h 31m of data, from 5 hearing and 11 deaf signers, as well as 1h 39m of spoken interviews with 4 hearing family members of deceased deaf islanders. When communicating with informants, we used a combination of languages to be best understood.

With hearing people, we drew on knowledge of Spanish, English, and Trinidadian English Creole and Islander Creole English. With deaf people we negotiated communication using a combination of cross-signing (Zeshan, 2015), gesture and our (growing) knowledge of PISL.

The data collected include several different genres from elicitations, to guided interviews to spontaneous conversation. We prioritised both (semi-)spontaneous data and lexical elicitation, as they would provide complementary sources for collecting a lexicon, and would be useful sources to describe and evaluate typological aspects of the language. Table 1 summarises the data collected and archived in the project. All video data can be found in the Endangered Languages Archive (ELAR) at https://www.elararchive.org/dk0562. Work on the lexical database and annotation are still underway and these files will be added to the collection gradually.
Table 1. Overview of data collected

<table>
<thead>
<tr>
<th>Genre</th>
<th>Sub-Genre</th>
<th># Recordings</th>
<th>Time (hh:mm:ss)</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elicitation</td>
<td>Lexical Elicitation</td>
<td>46</td>
<td>04:27:00</td>
<td>Signers are shown photos/objects and name the items to Carlos</td>
</tr>
<tr>
<td></td>
<td>(Picture-based)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Lexical Elicitation</td>
<td>3</td>
<td>01:42:03</td>
<td>Object- and word-based elicitation (with deaf and hearing signers)</td>
</tr>
<tr>
<td></td>
<td>(Other)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Story Re-telling task</td>
<td>23</td>
<td>00:20:56</td>
<td>Signers summarise a movie clip to Carlos</td>
</tr>
<tr>
<td>Discourse</td>
<td>Conversation</td>
<td>11</td>
<td>04:42:53</td>
<td>Conversation between two signers (one deaf-hearing pairing, rest deaf-deaf)</td>
</tr>
<tr>
<td>Interview</td>
<td>Signed Interview</td>
<td>4</td>
<td>02:22:38</td>
<td>Individual and group interviews conducted in sign by Carlos, Ian, Ben &amp; Rehana</td>
</tr>
<tr>
<td></td>
<td>Mediated Interview</td>
<td>4</td>
<td>01:55:34</td>
<td>Interviews with questions posed by Rehana/Ian in speech/sign and interpreted by either hearing or deaf signers into PISL</td>
</tr>
<tr>
<td></td>
<td>Spoken Interview</td>
<td>3</td>
<td>01:39:05</td>
<td>Spoken interviews conducted by Maureen and Carmelina</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>17:10:09</td>
<td></td>
</tr>
</tbody>
</table>

We elicited both lexical and narrative data. The aim of the lexical elicitation was to provide a base for creating a lexical database of PISL and to familiarise myself and Ian with local signs. We first performed a picture-naming task, showing signers pictures and asking them to name the object in the picture by signing to Carlos. We showed groups of pictures that belonged to several themes, such as animals, foods, and household items. This data was collected for 11 of the island’s 13 deaf signers. In addition to the picture naming, we performed object-based lexical elicitation with two deaf signers, where we showed them real-world objects and asked for the signs. We also performed one session with one deaf woman and her hearing relative in which we used a mixture of spoken words and photos for elicitation. For narrative data, we collected data from 4 deaf signers. We performed a story retelling task, aimed at collecting directly comparable narrative data from different participants. We presented signers with clips from a locally filmed movie *Bad Lucky Goat* (Oliveros et al., 2017), and asked the signer to summarise what they saw in the clip to Carlos.

We also collected spontaneous and semi-spontaneous data from deaf and hearing signers in the form of conversations. Because these recordings were aimed at capturing naturalistic conversation between PISL signers, neither Ian or myself were present during the recording in order to minimise the influence of community outsiders. We recorded 11 dyadic conversations (10 deaf-deaf, 1 deaf-hearing) between 7 deaf and 1 hearing signer, totalling 4h 42m. In these sessions, we prepared participants by suggesting themes, but signers largely directed the conversation themselves. These recordings were a total of 4h 42m from 7 deaf signers (including Carlos) and one hearing signer.
In addition, we also collected semi-spontaneous signing in the form of semi-structured interviews. During interviews, one or more team member was present in the recording session and guided the conversation around deaf islanders’ life experiences and family histories. We collected interviews using different modalities. For deaf islanders, we conducted interviews both completely in sign but also using a combination of speech and sign. The latter were mediated interviews in which I would ask questions in spoken language and family members would translate them into sign. In one instance, we also performed a mediated interview in sign where a deaf family member translated questions signed by myself and Ian, and interpreted the answers from their deaf relative. We recorded 2h 22m of signed interviews from 4 deaf islanders including individual and group settings and 1h 55m of mediated interviews with 3 deaf signers and their hearing relatives. Finally, we collected spoken language interviews, in which Maureen Hooker O’Neill and Carmelina Newball interviewed hearing people who had close family members who were deaf but were already deceased, a total of 1h 39m from 4 hearing participants.

At several stages throughout the project, we were met with the challenge of first understanding who uses Providence Island Sign Language, and second deciding who of these signers we should film for data collection. Far from being a straightforward question as it is treated in the literature, we encountered many challenges, especially because the signing situation in Providence is very different from that of well-studied macro-community sign languages. The deaf people are a small, highly heterogeneous and loosely connected group. They are spread across many villages and their main communication partners are mainly hearing people. They use different signs from each other, but seem to understand each other well. Nevertheless, signed settings like Providence are not rare. Several been described recently, for example, in rural Mali (Nyst et al., 2012), in the highlands of Papua New Guinea (Reed, 2019) and in the Mexican Yucatan (Safar, 2020). Nevertheless, this particular ecological niche is highly underrepresented in discussions on methodology of documentation and corpus methods (see however Hou, 2017; Nyst et al., 2012; Safar, 2021). I take the rest of this paper to reflect on how we performed sampling in Providence and how the lessons learned here can extend to other similar and dissimilar signing communities.

4 Sampling PISL signers

4.1 Sampling in sign language documentation

A major aim of documentary linguistics is to collect a diverse and representative sample of a language and, by extension, language users (Woodbury, 2003). However, the way in which different sign language documentation projects achieve this is quite variable. Sampling strategies vary considerably and often need to be adapted based on community size and structure. Despite the substantial differences across documentation projects, explicit discussion of sampling has been scarce in the literature. However, as the range of sign languages being documented diversifies, some key considerations have emerged with regard to how to create a representative sample of signers.

While most sign language documentation projects aim to capture sociolinguistic diversity, the sociolinguistic factors that are relevant may differ across communities.
For example, in macro-community sign languages used on a national scale, signers’ lexical choices often vary by region (e.g. Lucas et al., 2001; Stamp et al., 2014). As a result large corpora have made great efforts to sample across regions (e.g. Schembri et al., 2013). In micro-community sign languages, (extended) family or clan membership has been linked to lexical variation (e.g. Hou, 2016; Mudd et al., 2020; Safar, 2021; Sandler et al., 2011), and thus, efforts are made to sample across these groups. At present, we know much more about sociolinguistic factors that govern variation in macro-community sign languages used in Global North settings, yet given the cultural diversity of other settings in which sign languages are found, there is a lot of community specific work to be done to understand variation (Lutzenberger et al., in press).

Furthermore, considerations of nativeness, fluency and participation in the language community can lead some groups to be systematically excluded from some corpora and systematically included in others. Take, for example, the inclusion of hearing signers. In macro-community settings, while signing communities are made up of both hearing and deaf people, most corpora are structured around deaf signers often in an effort to centre deaf people and deaf epistemologies in sign language research. However in micro-community settings, hearing singers are thought to make up a larger proportion of the language community and as a result have often been systematically included in data collection (De Vos, 2016; Neveu, 2019; Nonaka, 2009; Nyst, 2015). Another example is the case of native signers. Learning a sign language from birth or early childhood has been linked to language fluency in deaf signers (Cormier et al., 2012; Emmorey, 2001). As a result, early deaf signers are often prioritised when planning data collection (Fenlon et al., 2015), despite this group making up an estimated 5-10% of all deaf signers in macro-community settings in the Global North (Kyle & Woll, 1985; Mitchell & Karchmer, 2004). In fact because most deaf people acquire sign later in life, and in some communities there are no ‘native’ signers (Costello et al., 2008; Nyst, 2008), many macro-community documentation projects are forced to relax or adjust their criteria for native/early signers (Crasborn & Zwitserlood, 2008; Schembri et al., 2013).

Pre-set sampling criteria can also present challenges where multiple factors intersect. For example, the BSL corpus project aimed to sample 10% of non-white signers to reflect the demographic composition of the UK, however they were unable to recruit enough non-white signers who also met other criteria (Schembri et al., 2013). Unequal distribution of certain sociolinguistic groups may also pose similar sampling challenges in both large and small communities. For example, the BSL project was unable to find non-white signers in the city of Belfast (Schembri et al., 2013). Similarly, in an investigation of sociolinguistic variation in the micro-community of Kata Kolok signers, Mudd and colleagues (2020) were unable to find old deaf people who belonged to a specific family clan.

Overall, there is very little literature that explicitly discusses sampling issues in sign language documentation, and fewer still that focus on micro-community settings. In the rest of this section, I focus on the PISL documentation project as a case study, evaluating the process of creating a representative sample of a signing community. I discuss how and why we delineated our aims including our limitations, as well as what worked and what did not work with respect to sampling. Finally, I discuss our project in the broader context of sign language documentation and corpus creation.
4.2 Designing a sampling strategy

Upon arrival in Providence, it became clear that there was a highly diverse group of signers, each with quite variable signing styles. Given this high level of variation, we were wary of over-representing one deaf person’s signs/signing as those used in the whole island. Following Nyst’s suggestion, “[i]n case a sign language has a relatively small number of users, it may be preferable to document the signing of all its signers, rather than of a sample” (2015: 113), we decided in order to document the range of sign language in use on the island, we would take a highly inclusive approach.

A result of this decision was that we did not exclude any signer on the basis having lived abroad or knowing other sign languages. Thus, deaf people who were born on San Andres and moved to Providence, and deaf people who spent time in the mainland were all included in the data collection. We performed a baseline of lexical elicitation with these signers, to fully capture the range of signs in use on the island for a single concept. We did not discourage signers from using other sign languages, but instead we recorded multilingual signers in spontaneous discourse with other islanders, so that we could capture what natural language use looks like in these settings. We also took measures to understand these individuals’ language backgrounds and repertoires. Given that some deaf signers spent considerable time in both mainland Colombia and San Andres, we consulted with signers of Colombian Sign Language, and signing varieties in San Andres to identify specific signs or variants that were used in these places.

Given the backgrounds of the signers, we also chose not to apply any criteria of nativeness or fluency as we felt this was not useful for the aims of our project. Deaf people in Providence develop and use signing skills from an early age to communicate with a range of mostly hearing interlocutors of varying proficiency, and local signing is not taught in an institutionalised setting. All deaf people are born to hearing parents and, with only one exception, have not had deaf adult signing models in their immediate family when growing up. This is quite different from situations in which age of acquisition is often considered, namely acquisition of national sign languages in schools or from deaf signing parents (Lillo-Martin & Henner, 2021). Given these considerations, we considered all deaf people in Providence as appropriate participants for data collection.

While we aimed to include all signers in the corpus, we still made decisions to prioritise some groups or sampling distributions, based on previous research, and our early observations. Particularly given the limited time, we found this useful to guide our data collection.

One such factor was age of deaf islanders. For the purposes of this paper, we divide signers into two groups, those aged 40 or under (younger) at the time of data collection and those over 40 (older). Given that several deaf signers over the age of 40 had been included in Washabaugh’s initial documentation, we aimed to record at least elicited and semi-spontaneous data from these signers for potential diachronic comparisons. We also suspected age-based variation as younger signers have more contact with Colombian Sign Language and other national, macro-community sign languages through education, travel and the internet. We thus decided to prioritise older signers who generally used more distinctively local signing. Finally, given that
we were unsure of the possibility to return to Providence in the near future, we decided to prioritise filming older signers.

Another factor was village membership. Since village membership has been described to influence lexical choices in Providence, we aimed to include signers from various villages in our data collection. In the past, where a signer lived had a greater impact on their social network: given limited mobility around the island, signers interacted more with their fellow villagers. During this time, Washabaugh describes characteristic lexical differences across different villages in the island, and also notes that some villages have a rich history of signing and deafness and others do not (Washabaugh, 1986: 51). Today, despite a much greater freedom of movement, recent lexical elicitations have highlighted that those lexical differences documented by Washabaugh persist across different villages (Braithwaite, 2020b; Cortés Bello, 2016). To document this variation, we aimed to collect data from each of the seven villages in which deaf signers currently live on the island.

We also considered audiological status while sampling. In line with other sign language documentation projects, we focused on deaf signers, as we consider them the core of language users. However, we aimed to also sample across hearing signers because they make up a significant portion of language users. Furthermore, work by Washabaugh (1979) documented differences in word order among deaf and hearing signers on the island, leading us to suspect this factor as relevant for other types of linguistic variation. Given that this variation was observed in word order and hearing signers were observed to use similar lexicon as deaf signers (Washabaugh, 1986), we did not focus on lexical elicitation with hearing signers but tried include them into semi-spontaneous data collection.

4.3 Implementing the sampling strategy

Overall, we were relatively successful in achieving our sampling goals: we collected data across six out of seven villages where deaf people lived, from 11 out of 13 deaf islanders from ages 19 to 81 years and from five hearing islanders. The distribution of data collected from different informants can be found in Table 2, highlighting the three sampling factors considered. Despite general success, our priorities for sampling were met with practical limitations in the field, that cut across different factors. In what follows, I describe unforeseen issues we encountered in collecting a diverse and representative sample.

<table>
<thead>
<tr>
<th>Signer Code</th>
<th>Audiological Status</th>
<th>Age Group</th>
<th>Village</th>
<th>Type of Data</th>
<th>Total Data (hh:mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AB</td>
<td>Deaf</td>
<td>older</td>
<td>Southwest Bay</td>
<td>Elicitation, Interview, Discourse</td>
<td>03:40:45</td>
</tr>
<tr>
<td>AO</td>
<td>Deaf</td>
<td>younger</td>
<td>Bottom House</td>
<td>Elicitation, Interview</td>
<td>00:51:33</td>
</tr>
<tr>
<td>BlaT</td>
<td>Hearing</td>
<td>older</td>
<td>Rocky Point</td>
<td>Interview</td>
<td>00:26:28</td>
</tr>
<tr>
<td>BT</td>
<td>Deaf</td>
<td>older</td>
<td>Rocky Point</td>
<td>Elicitation, Interview, Discourse</td>
<td>02:15:30</td>
</tr>
<tr>
<td>CN</td>
<td>Deaf</td>
<td>younger</td>
<td>Town</td>
<td>Discourse</td>
<td>00:49:41</td>
</tr>
<tr>
<td>DB</td>
<td>Deaf</td>
<td>older</td>
<td>Bailey</td>
<td>Elicitation, Interview</td>
<td>01:36:44</td>
</tr>
</tbody>
</table>
Village membership

By recording data from all available deaf adults on the island, we were able to meet the goal of recording across several villages. However, it became clear that variation was not always based on village membership, but instead centred around individual deaf people and their communication and family networks. For example, two unrelated signers in one village may have very different signs for two concepts, even if they know and communicate with each other relatively. However, the two deaf sisters and their deaf uncle who lived in the same village shared much more lexical overlap. Given these observations, we aimed to collect lexical data from all available deaf signers, including multiple signers who lived within the same village, to capture the range of signs used on the island.

Age

The two youngest deaf people were not included in the current corpus due to age-related restrictions: the youngest was underage and therefore unable to give consent, and the second youngest was attending school in Medellín and as a result not available during the majority of the research stay. Other than that, we found no major challenges with sampling signers across different ages and focusing on older signers over 40.

Audiological status

When attempting to record both deaf and hearing signers, we found at times that the mixed audiological status of our team had an effect on participants language choices. Our team comprised two deaf members (Carlos, Ian) and one hearing member (Rehana). While most deaf participants preferred to sign with us when all together, it became apparent that one deaf woman tended to use far more mouthings when I was present in recording settings. For this reason, we tried to record data from this signer both with and without me present. A related issue arose when trying to record hearing people in signing mode: once I was present, some hearing signers became more interested in chatting in spoken language instead of sign. In one instance, we visited the house of a deaf man and his sister to record a group conversation among all of us. However, the group quickly divided: we ended
up with the hearing sister talking to me in spoken language and Carlos, Ian and the deaf man signing to each other.

These situations may reflect cultural norms in Providence. Washabaugh describes that while hearing people use signing for 1:1 interactions with deaf people that often when more hearing people are around, they do not sign at all (Washabaugh, 1986: 100). Similarly, he describes that some deaf people prefer to interact with hearing over deaf people (1986: 135). In deaf-hearing communication, mouthing and speaking is a widespread communication strategy used, and thus unsurprising that it would be used more actively in the presence of a hearing researcher. Having a hearing team member allowed us to observe the way that, in some settings, local ideologies and practices privilege spoken language.

Hearing participants also presented different challenges to data collection in terms of familiarity and comfort with the research project. While deaf people mostly had clear ideas and understandings of the project goals and quickly became accustomed to equipment and recording settings, hearing people were often more reluctant to be filmed. This is likely because given the limited time of the field trip, we prioritised spending time and developing individual relationships with deaf people and instead interacted with hearing people through other avenues, such as workshops and events in collaboration with the government. Spending time with deaf people made us familiar to their hearing family members and neighbours, however these hearing people did not often attempt to join our signed conversations during visits. Altogether, we spent far less time cultivating relationships with hearing islanders, and as a result collected far less data from them.

Finally, identifying and recruiting hearing signers to film was a challenge. Our first aim was to identify and record ‘good’ signers, however we quickly realised that determining the degree of signing proficiency among hearing people is quite a complex task in micro-community settings, especially for community outsiders (Mudd et al., 2020; Nonaka, 2009). In Providence, Washabaugh (1979) created a schematic representation of signing networks, with the closest relatives and household members of deaf signers as the most fluent. In practice however, we found that even those hearing people who were close family members were not necessarily judged to be the highly proficient or the most desirable interlocutors by their deaf family members, despite their own beliefs about their signing skills. After some difficulties identifying appropriate hearing signers, we finally found success by asking a deaf signer to invite her hearing friend to be recorded together.

During this project, we did not prioritise filming hearing people with less contact with deaf people. However, observations in the field made it clear that filming the many micro-interactions between deaf signers and their extended network, such as local shopkeepers, customers, employers, co-workers, would be an important next step in documenting the signing community. Given hearing signers’ reluctance to be filmed, collecting such data would require different methods, for example following singers throughout their day and filming their interactions (Nyst, 2015). Drawing on methods from linguistic ethnography may be particularly useful in this case as they allow flexibility in capturing everyday interactions (Kusters & Hou, 2020). By moving away from the idea of sampling ‘good’ signers, we are likely to deepen our understanding of how communication is negotiated in places like
Providence where there are many hearing people who use sign to communicate in specific domains.

**Individual differences**

Among the small population of deaf people, there were large differences in availability and willingness to participate in data collection. Some deaf people, while quite interested in the project, were limited in their participation due to busy work schedules, family commitments, or travel abroad. Despite ample free time, another deaf person seemed largely uninterested in the project and as a result was rarely approached for data collection. We responded to participants’ personal circumstances by focusing time on those who were willing and available to participate, and collected more data from them. Despite various limitations, we also made a concerted effort to collect basic lexical elicitation from all deaf people for whom this was a logistical and ethical possibility leaving us with data from the majority of deaf islanders.

Differences in experience with linguistic research also played a role early on, particularly as the research team developed relationships with individual deaf people. Some deaf people had multilingual signed competence, or experience with other linguistic research and as a result had some degree of metalinguistic knowledge. This facilitated communication and understanding of the project aims between researcher and participant, and these individuals quickly became important informants. In particular, one deaf person who had been involved in previous language documentation by Washabaugh was extremely enthusiastic and interested, devoting a great deal of time to the data collection. Others had less experience with linguistic research but were happy to participate, and through getting to know participants they became more acquainted with the aims and methods of the project more gradually.

In addition, interpersonal relationships between participants restricted multiparty data in two ways. First, to capture naturalistic signing, we aimed to record the two deaf sisters signing to each other. However, we quickly realised that signers who were in close daily contact were not particularly interested in talking to each other: when left alone to chat on camera, these signers very quickly signalled that their conversation was over. Similar observations have been made in other signing communities for deaf people in the same household, for example married couples (Schembri et al., 2013), and siblings (Ali, 2021). Second, we had to adjust flexibly to the (at times fragile) interpersonal dynamics between signers.

For example, one deaf woman was scheduled to be recorded with another and a few days before the recording event they argued and we had to cancel. Another deaf man, despite being a member of a deaf multigenerational family, was largely considered unintelligible by everyone around him. One signer was particularly chatty and dominated group conversations. We took different angles to solve these problems: we switched conversation partners for the arguing pair, recruited a family member to mediate an interview for the hard-to-interpret signer, and included Ian and/or Carlos to moderate group conversations.
4.4 The result: a representative sample of PISL?

To briefly recap, we had a great deal of success recording deaf people and found that concentrating on them gave a good understanding of the range of lexical variation in the island’s signing community. Given the small diverse population, we aimed to be as inclusive as possible, particularly because some deaf people couldn’t participate due to ethical and practical limitations. Even so, we became aware of the critical importance of managing individual personalities and schedules in order to make sure we could get data from as many deaf people as possible. We found hearing people more challenging to record, in part because we did not prioritise developing relationships with them. We also made observations that are important for future research. In terms of recording hearing signers, we found the most success in employing a ‘bring-a-friend’ set up for recording discourse. However, we observe that that focusing on ‘good’ hearing signers might be unnecessarily limiting, and expanding the method and scope of recording hearing signers may be useful to understand how signing is used in Providence.

Instead of opting for a maximally structured data collection, we were opportunistic, recording more data with those who were more flexible and willing to do so. In doing so, we adopted a ‘emergency butterfly collection’ approach (Chelliah & de Reuse, 2011), which resulted in a lot of data from some signers and less data from others. This opportunistic approach might have its merits in terms of representative sampling. In our case, those who contributed no or very little data appeared to be overall less integrated into the signing landscape for various reasons: they were too young, they were reclusive, they preferred to speak than sign. Similarly, the hearing signers we recorded all used sign in daily interactions with deaf people. While this is not an exhaustive sample of PISL signers, it seems likely that the group represented in our sample are active language users. Nyst (2008) discusses another example of opportunistic sampling of Malian Sign Language, used in the capital of Mali, which emerged in the public spaces where deaf people meet and gather. Nyst comments that while they found difficulty in collecting data from female signers, it is likely that the signing of the majority-male sample is more representative of the active signing community, as the public spaces in which the language arose are outside the traditional sphere of women. One critical consideration however is that in micro-communities, sign language documentation is often done by community outsiders and it is possible that without close collaboration with community members that they may miss settings in which sign language is used.

An opportunistic strategy is quite different from the structured collection of macro-community corpora, which often aim to collect data that is balanced across multiple variables, such as age, region and gender. Nevertheless, many issues of representability are still unaddressed in these more structured approaches. In recent times, the preferential sampling of native and near native signers has been problematised, particularly because in many cases they represent only a small fraction of sign language users (Costello et al., 2008). However, achieving a true representative sample in macro-community sign languages may be challenging because in larger signing communities it is much more difficult to ascertain the full

---

2 See for example the discussion following a recent tweet by Julie Hochgesang https://twitter.com/jahochem/status/112034835983915392 (Hochgesang, 2019)
degree of demographic and sociolinguistic variation of deaf signers. For example, in
the context of British Sign Language Corpus, Schembri and colleagues (2013: 140)
ote that “[n]ot enough is known about the population of deaf sign language users
in the United Kingdom in order to recruit a representative group whose
characteristics we can confidently say reflect those of the larger population.”

Sign language corpora and datasets have widely different aims which guide the
inclusion/prioritisation of certain groups. All sign languages, including those used
in both macro- and micro-communities are minority languages and the very first
large scale documentation projects have approached this endeavour form the
perspective of endangered language documentation (e.g., Johnston, 2008). However,
the major divergence in sampling between macro- and micro-community
focused projects is likely not just due to differing community size and network
structure, but also to differences in project aims, methods and ideologies. One
element of this is criteria about fluency. In macro-community sign languages,
corpus research projects often target both linguistic research but also usability for
teaching, learning and interpreting (Konrad, 2012). On the other hand,
micro-community sign language corpora have often more explicitly focused on
documentation and description of the signing community. These diverging project
aims may underly differing motivations in micro- and macro-community corpora to
include signers considered to be ‘less fluent’.

Nevertheless, the differences observed in sampling across signing communities
suggests that we should not take for granted that all corpus/documentation projects
(i) do, or should, include the same groups and (ii) are, or should be, comparable.
This has important implications for the growing field of sign language typology, and
particularly for typological studies that compare sign languages with very different
community size and structure. Providing comprehensive meta-documented
of corpus creation and sampling will help, if not to refine ideas about what a
representative sample is (or should be), at least to better understand whose language
use is being represented in a given dataset or corpus.

5. Conclusion

In this paper, I have presented a small-scale documentation project of Providence
Island Sign Language as a case study to critically reflect on how to create a
representative sample of sign language users. In doing so, I highlighted the flexibility
required to carry out sampling in a small and highly diverse community, and discuss
key unforeseen issues that we encountered with respect to managing individuals and
interpersonal relationships. This reflection also highlights the differences in
sampling between our corpus and similar small-scale corpora of micro-community
sign languages and large-scale corpus projects of macro-community sign languages
in the Global North. I hope that this makes the PISL dataset, and any research that
builds on it, more transparent, accessible and informative to others.

Finally, the question remains, what is next for the Providence Island Sign Language
corpus? At present, planned follow up data collection has been halted due to the
COVID-19 pandemic and the devastation of the island by hurricane Iota in 2019.
While future data collection is on pause, one immediate opportunity to expand the
corpus is through the digitisation, archiving and annotation of historical data, which
has been generously offered for use by William Washabaugh. Collecting these materials alongside the current data will allow the novel possibility of comparing data from the same signer 40 years apart.

In its current state, the corpus already presents a great deal of potential for linguistic research. The combination of elicited and spontaneous data invites in-depth studies of linguistic features of PISL. At present I am currently using conversational data to examine features of PISL language in interaction including person reference (Omardeen et al., 2021) and other initiated repair strategies (Omardeen & Manrique, in prep) as part of my PhD dissertation. The corpus also opens up the possibility for cross-linguistic comparisons between PISL and other sign languages. For cross-linguistic comparisons however, I encourage researchers to reflect on the methodology and context of data collection particularly when comparing PISL to languages from very different signing communities.

References


Hochgesang, J. A. (2019, April 22). 📝 #SignLanguageResearchTwitter do you know of any refs that mention (or even better, discuss) native signers as a potentially overstudied population in signed language research? @AdamCSchembri @mdemeulder @ErinMoriartyH @kemmorey1 @DLilloMartin @jmhenner @linasigns @ryanlepic [Tweet]. https://twitter.com/jahochcam/status/1120348359833915392
Sampling Signers in Providence Island: Reflections on a Small-Scale Documentation Project

Rehana Omardeen


Omardeen, R., & Manrique, E. (in prep). Other-initiated Repair in Providence Island Sign Language.


