# A grammatical description of Shiwiar

Kohlberger, M.

Cover Page



The handle http://hdl.handle.net/1887/123115 holds various files of this Leiden University dissertation.

**Author**: Kohlberger, M.
**Title**: A grammatical description of Shiwiar
**Issue Date**: 2020-06-29

Chapter 4

# Sources and Data

This grammatical description is largely based on primary data collected during 12 months of fieldwork, but there are occasional references to secondary sources as well. This chapter lists all the data sources used in this dissertation and explains the methodology used to collect and compile the primary data. Section **4.1** is a review of all the linguistic work that has previously been written on the Shiwiar language. In §**4.2** I describe the fieldwork I undertook to document spoken Shiwiar and in §**4.3** I acknowledge all the Shiwiar people who worked with me on this project. The process through which the data was annotated and analysed is described in §**4.4**. Finally, in §**4.5**, all the recordings used as a basis for this grammar are listed, as well as information about where they are archived and how they can be accessed.

## 4.1. Previous work on Shiwiar

At the time of writing this work, the existing literature on the Shiwiar language was minimal. Previous work was done by SIL missionary linguists (Fast Mowitz, Warkentin de Fast and Fast Warkentin), community members (Vargas Canus and Tsetsekip) and a BA linguistics student (Oranen). **Table 4.1** is an exhaustive list of publications about the Shiwiar language.

| Reference | Topic |
|---|---|
| Fast Mowitz, Warkentin de Fast & Fast Warkentin 2008 | Dictionary and grammar sketch; mostly based on Achuar data |
| Oranen 2013 | Possession marking |
| Vargas Canus & Tsetsekip 2002 | Small dictionary |

**Table 4.1.** Previous work on the Shiwiar language

It should be noted that because SIL considers Achuar and Shiwiar to be one language ("Achuar-Shiwiar"), it is unclear how much of their data and generalisations refer only to Achuar or Shiwiar, or both.

## 4.2. Fieldwork

This work is almost entirely based on data I collected between 2011 and 2016 in three main fieldwork sites: Juyuintsa, Puyo and Shell. Juyuintsa is the easternmost Shiwiar village in Ecuador, located on the Conambo River (see **Table 3.1**). It has 59 inhabitants and Shiwiar is the primary language spoken there (although most people also speak Kichwa and Spanish to varying degrees; see §**3.6**). Puyo, the capital of the Pastaza province, is the nearest administrative centre and is located over 200km to the northwest of Juyuintsa. It has a population of 36,659 people (INEC 2010) and is a largely Spanish-speaking town, but it is also the home of Kichwa, Achuar, Shuar, Shiwiar and Wao communities which continue to use their traditional language. Most of the work in Puyo was conducted at the NASHIE headquarters, which are visited by Shiwiar community members on a daily basis. Shell is a small town of 8,752 people (INEC 2010), located 8km to the west of Puyo. It is the busiest transport hub in the province because of its airport: flights from Shell are the only access route to many isolated jungle communities, including all Shiwiar settlements. Whenever Shiwiar people need to come to Puyo to settle financial or legal matters, they often find temporary residence in Shell. **Table 4.2** is a list of the aforementioned fieldwork sites including their geographic coordinates.

| Location name | Geographic coordinates | Population | Main language (Other languages) |
|---|---|---|---|
| Juyuintsa | 2° 07' S, 76° 11' W | 59 | Shiwiar (Spanish, Kichwa) |
| Puyo | 1° 29' S, 77° 59' W | 36,659 | Spanish (Kichwa, Achuar, Shuar, Shiwiar, Wao) |
| Shell | 1° 30' S, 78° 03' W | 8,752 | Spanish (Kichwa, Achuar, Shuar, Shiwiar, Wao) |

**Table 4.2.** Fieldwork sites

Between 2011 and 2016, I visited the field six times. During each trip I collected data by recording speakers of Shiwiar in a variety of interactive scenarios. All of the recordings made in 2011 and 2012 only have an audio component, whereas most of the recordings made from 2014 onwards have a video component as well. **Table 4.3** summarises the start and end dates of

each field trip, as well as the locations visited.  It also lists the number of audio and video recordings collected during each trip and the combined duration of the recordings.  In total, the corpus includes 30 hours of spoken Shiwiar.[22]

| Fieldtrip Code | Start Date | End date | Field Sites | Recordings Collected (Hours:Minutes) |
|---|---|---|---|---|
| T01 | 2 Nov 2011 | 29 Dec 2011 | Juyuintsa; Puyo | 21 audio (4:30) |
| T02 | 1 Jul 2012 | 6 Nov 2012 | Juyuintsa; Puyo | 37 audio (6:55) |
| T03 | 15 May 2014 | 11 Jun 2014 | Puyo; Shell | 28 audio (6:04); 29 video (5:52) |
| T04 | 15 May 2015 | 9 Jul 2015 | Puyo | 15 audio (7:33); 3 video (1:02) |
| T05 | 1 Nov 2015 | 2 Dec 2015 | Puyo | 2 audio (00:17); 1 video (0:15)[23] |
| T06 | 9 May 2016 | 23 May 2016 | Puyo; Shell | 7 audio (3:03) |

**Table 4.3.** Fieldtrips

The audio data was recorded using a Zoom H4N solid state recorder and a selection of three microphones: a head-worn Beyerdynamic Opus 55.09 Mk II SC omnidirectional condenser microphone, a head-worn Shure SM10 unidirectional dynamic microphone and the inbuilt unidirectional stereo condenser microphone of the Zoom H4N recorder.  The video recordings were done with a Canon Legria HF G25 wide angle high-definition camcorder.

Most of the data is natural speech, i.e. unconstrained connected discourse by one or multiple speakers.  Many of the recordings were initiated in somewhat staged environments: often a recording session would be prearranged, and speakers would assemble with the specific aim to record

---

[22] The cumulative total hours of audio and video recordings in **Table 4.3** appears to be more than 30.  However, some of the audio and video recordings were of the same speech event, so I only count them once when I report the total hours of speech documented in the corpus.

[23] During this trip, most of my time was spent transcribing, translating and annotating earlier recordings, which is why only a minimal amount of data was added to the corpus.

speech. Despite this – although occasionally after some initial awkwardness – speakers became comfortable with the situation and spoke naturally. Some recordings resulted from completely spontaneous situations. The discourse can also be classified into a number of genres, including traditional stories, biographical accounts, procedural texts and conversations.

Elicited speech was also recorded, but it makes up a much smaller subset of the total data. Elicitation tasks were mostly used in order to collect specialised lexicon that is infrequent in spontaneous speech (such as words for specific flora and fauna). I would like to stress that analytic observations or generalisations based solely on elicited data are avoided in this work in order to minimise all of the potential problems that are known to arise when using this methodology, and to remain true to the way that the Shiwiar language is actually used by its speakers (cf. Mithun 2001). However, elicited examples are sometimes used in this dissertation in order to convey simple phonological or morphological points.

## 4.3. Speakers and consultants

Over thirty Shiwiar speakers and consultants were involved in the creation of the documentary corpus. One person in particular, Verónica Suquilanda, played a central role throughout the process. She introduced me to many people who ended up collaborating with the project, and she often volunteered to be part of the recorded speech event in order to ease in other interlocutors and distract them from the possible awkwardness of being recorded.

**Table 4.4** is a list of most of the Shiwiar people who participated in the project; but it does not include the participants who explicitly asked not to be acknowledged. It is important to note that the participants come from ten different villages and have a wide range of ages, making the corpus representative of the Shiwiar Nation as a whole.

| Name | Village | Gender | Age at time of recording |
|---|---|---|---|
| Francisco Timias | Juyuintsa | Male | 21-23 |
| Jorge Santi | Kurintsa | Male | 36 |
| Verónica Suquilanda | Juyuintsa | Female | 26-30 |
| (Anonymous) | Kurintsa | Female | 35 |

**Table 4.4.** Speakers

| Name | Village | Gender | Age at time of recording |
|---|---|---|---|
| Florentina Cuji | Tunguintsa | Female | 45-48 |
| (Anonymous) | Panintsa | Male | |
| (Anonymous) | | Male | |
| Nelly Aguasanta | Juyuintsa | Female | 20 |
| Rosenda Tsetsekip | Juyuintsa | Female | 40 |
| Fanny Timias | Juyuintsa | Female | 50 |
| Jerónimo Timias | Juyuintsa | Male | 20 |
| (Anonymous) | Kurintsa | Female | 43-45 |
| Julia Pantam | Tunguintsa | Female | 65 |
| Nelson Cuji | Tunguintsa | Male | 20 |
| Samuel Mayancha | Panintsa | Male | 25 |
| Agusto Dahua | Kapirna | Male | |
| Yapanda Cuji | Kapirna | Female | |
| Yambia | Tanguntsa | Male | 65 |
| Daisy Cuji | Juyuintsa | Female | 27-28 |
| Laura Aguasanta | Wiririma | Female | |
| (Anonymous) | Wiririma | Female | |
| Efraín Andi | Panintsa | Male | 30 |
| Eduardo Aranda | | Male | |
| Alfredo Gualinga | Yandanaentsa | Male | |
| (Anonymous) | | Male | |
| Maruja Gualinga | Juyuintsa | Female | |
| Guadalupe Timias | Juyuintsa | Female | |
| Beatriz Nango | Kanai | Female | |
| Andrés Paantam | Bufeo | Male | |
| Efrén Nango | Kanai | Male | |
| Dixon Timias | Juyuintsa | Male | |

**Table 4.4. (cont.)** Speakers

All the participants were told in detail about the nature of the project and gave informed consent to being recorded. At the end of the recording, speakers were given another opportunity to decide whether the speech event should be added to the corpus. Although some recordings were deemed to be too private and therefore discarded, most recordings were approved to be included.

## 4.4. Corpus compilation and annotation

All of the data collected was annotated and compiled into a corpus for the purpose of analysis. The recordings were first grouped into so-called bundles referring to a single recording session. For example, if the recording method for a session was audio-visual, two files (one audio and one video recording) would be bundled together. If a recording had to be interrupted and then continued within a single thematic session, multiple sequential audio and video files would be grouped into a bundle. Each bundle was then organised and, if necessary, synchronised in ELAN (The Language Archive 2014; Wittenburg et al. 2006).

Each of the recordings was transcribed orthographically and translated into Spanish with the help of Shiwiar consultants. I am particularly grateful to Verónica Suquilanda, Francisco Timias, Maruja Gualinga, Florentina Cuji, Daisy Cuji and Eduardo Aranda for their patience and insights during that part of the documentary process. The transcriptions and translations were imported into the ELAN bundles, divided up into prosodic units and time aligned to the media files. I then added a phonetic transcription and an English translation to the bundle. The four tiers (phonetic transcription, orthographic transcription, Spanish translation and English translation) were then imported into FLEx (SIL 2014) for interlinearisation and morphological glossing. The parsed and glossed texts were finally reimported back to ELAN. Consequently, the entire annotated corpus can be browsed in ELAN, and all the texts, along with a full lexicon, are saved in a FLEx file.

The entire corpus is in the process of being deposited into the Archive of the Indigenous Languages of Latin America (AILLA; https://ailla.utexas.org/) under the collection name "Corpus of Spoken Shiwiar (Chicham, Ecuador)". At the time of writing, the full corpus is not yet archived, but the process is expected to be completed in 2020. Each of the recordings in the corpus will have its own access conditions as stipulated by the speakers who were recorded: some will be publicly accessible, some will require a simple access request procedure, and a few recordings with sensitive information will be kept under embargo during the lifetime of the speakers involved.

## 4.5. List of recordings

Although the grammatical insights that underlie this work came from a much larger and more diverse set of recordings, the data and examples cited in this dissertation were only taken from a subset of recordings for which speakers agreed open access to. A list of recordings cited in this dissertation can be found in **Table 4.5**. These recordings represent a wide range of speech genres, including traditional narratives, autobiographical accounts, anecdotes, day-to-day conversations amongst neighbours, political speeches and procedural texts. Many of the recordings involve dialogue between multiple speakers, but it should be noted that even monologic narratives were recorded in an interactive environment with native speaker listeners present, so as to ensure that the discourse was as natural as possible.

| Recording Name | Title | Genre |
|---|---|---|
| T01-S01-01.wav | Words and phrases 1 | Elicitation |
| T01-S01-02.wav | Words and phrases 2 | Elicitation |
| T01-S02-01.wav | Words and phrases 3 | Elicitation |
| T01-S02-02.wav | Words and phrases 4 | Elicitation |
| T01-S02-03.wav | Origin of the white-lipped peccary | Narration |
| T01-S02-04.wav | Amasank | Narration |
| T01-S02-06.wav | The sun and the wind | Narration |
| T01-S02-07.wav | The runway in Kurintsa | Anecdote |
| T01-S02-09.wav | Origin of the boa (interrupted) | Narration |
| T01-S02-10.wav | Origin of the boa (full) | Narration |
| T01-S03-01.wav | Origin of the potoo | Narration |
| T01-S03-03.wav | Growing up in a Shiwiar village 1 | Anecdote |
| T01-S03-04.wav | Growing up in a Shiwiar village 2 | Anecdote |
| T01-S03-05.wav | Growing up in a Shiwiar village 3 | Anecdote |
| T01-S03-06.wav | The boy and the old woman 1 | Narration |
| T01-S03-07.wav | The boy and the old woman 2 | Narration |
| T01-S04-01.wav | Cinnamon and genip | Narration |
| T01-S05.01.wav | The heron | Narration |

**Table 4.5.** Recordings

| Recording Name | Title | Genre |
|---|---|---|
| T02-S03S05-01.wav | Future plans | Conversation |
| T02-S03S12-01.wav | Community work | Conversation |
| T02-S03S16-01.wav | Recent events 1 | Conversation |
| T02-S03S17-01.wav | Family | Conversation |
| T02-S06-01.wav | Planning a workshop | Political speech |
| T02-S07-01.wav | Introduction to a workshop | Political speech |
| T03-S01-01.wav | The story of the star woman | Narration |
| T03-S01-03.wav | Mythical cannibals | Narration |
| T03-S01-04.wav | The foundation of Juyuintsa | Anecdote |
| T03-S01-05.wav | The lost children | Narration |
| T03-S01-06.wav | Recapping the events of the day | Anecdote |
| T03-S01-07.wav | The Spanish film crew | Anecdote |
| T03-S01S08-01.wav | Discussing which story to tell | Conversation |
| T03-S03S12-01.wav | Words and phrases 5 | Elicitation |
| T03-S03S14-01.wav | Conversation about family 1 | Conversation |
| T03-S03S14-02.wav | Conversation about family 2 | Conversation |
| T03-S03S15-01.wav | Conversation about home life | Conversation |
| T03-S08S09-02.wav | Remembering childhood | Conversation |
| T03-S08S09-03.wav | Recent events 2 | Conversation |
| T03-S09-01.wav | The origin of snakes | Narration |
| T03-S10-01.wav | Gardening | Procedural text |
| T03-S10-02.wav | The spirit Amasank | Narration |
| T03-S11-01.wav | How to build a house | Procedural text |
| T03-S11-02.wav | How to hunt | Procedural text |
| T03-S12-03.wav | The origin of birds | Narration |
| T03-S12-04.wav | Working in the Shiwiar leadership | Anecdote |

**Table 4.5. (cont.)** Recordings

The recording name itself immediately provides three pieces of information separated by hyphens. The first code (which begins with the letter T) represents the fieldtrip in which the recording was made (see **Table**

**4.3** for details about each fieldtrip). The second code (which begins with the letter S) represents the main interlocutors who were active in each recording. The third code indicates the recording number. In other words, T03-S03S14-02.wav is the second recording of Speaker 3 and Speaker 14 recorded during the third fieldtrip. The reason why speakers are not identified by their full names in the recording labels is because, although most participants of the documentation project agreed to be acknowledged by name in general, the vast majority did not want to be immediately identified in each recording or example in this text. For this reason, the recordings are anonymised in this grammar.