

## Integrating Language Documentation, Language Preservation, and Linguistic Research: Working with the Kokamas from the Amazon

Rosa Vallejos  
*University of New Mexico*

This paper highlights the role of speech community members on a series of interconnected projects to document, study and maintain Kokama, a deeply endangered language from the Peruvian Amazon. The remaining fluent speakers of the language are mostly older than 60 years of age, are spread out across various small villages, and speak the language in very restricted situations. The aim of this paper is twofold. First, it demonstrates with concrete examples that outcomes of projects implemented in collaboration with speakers yield more broadly useful outcomes than those conducted by a linguist working alone. Second, it underscores the significance of documenting language interaction among different types of speakers in accordance with the view that language preservation is not only about promoting a linguistic code, but also includes documenting communicative practices. The projects reported here can contribute to the development of fieldwork methodologies to work with a range of speakers. The involvement of community members has been crucial for the design of culturally relevant strategies to assess fluency in Kokama, for the naturalness and variety within the collected data, and for the documentation of interactional patterns essential for revitalization initiatives. This paper supports the view that language documentation, language preservation, and linguistic research can be complementary endeavors.

**INTRODUCTION.**<sup>1</sup> The first time I visited the Kokama territory was in February of 1997. At that time I was hired to work in the *Programa de Formación de Maestros Bilingües de la Amazonía Peruana* (FORMABIAP). As part of a multidisciplinary team, I was the linguist responsible for working on the description and maintenance of Kokama, also known as Kukama-Kukamiria. My first trip to the Kokama territory lasted only 32 days, but that was enough time to ascertain the complexity of the situation. On the one hand, there was an obvious sense of urgency to work towards the preservation of this highly endangered language. The few surviving speakers that I met reported having no one with whom they could use Kokama on a daily basis, and lamented the disappearance of their language. On the other hand, the complex socio-political context made implementing initiatives towards

---

<sup>1</sup> My deepest gratitude goes to each and every one of the Kokamas that have shared their time and knowledge with me since 1997, especially to Victor Yuyarima Chota and Rosa Amías Murayari, my long-time friends and collaborators. This paper has benefited from detailed comments by Racquel Yamada and Keiko Beers. I am also grateful for discussion with Colette Grinevald and colleagues from the Laboratory Dynamique Du Langage, University of Lyon. Financial support from the National Science Foundation Documenting Endangered Languages program (BCS #0965604) is gratefully acknowledged. All errors and shortcomings are my own.

the preservation of the language a real challenge. In some villages, speaking about the language was almost a taboo subject.

However, the preservation of the Kokama language was only one component of a bigger movement initiated by indigenous organizations in the 1980s to address primarily land, education and health issues among indigenous Amazonian groups. Those efforts led to the corresponding spread of discourses that promote higher self-esteem and recognition of indigenous identity. As a result of a series of interconnected projects—which range from teaching the language through instructional programs in elementary schools, teaching the language in neighboring midsize towns and cities, daily radio programs in Kokama, and radio ads for health campaigns, among others—there has been a modest but significant positive shift in the last decade in attitudes towards Kokama identity and a renewed sense of community.

This paper reports on a series of interconnected projects implemented in collaboration with the Kokamas. In the last decade, we have seen a growing consensus that researchers, particularly those working in endangered language communities, have a responsibility to engage in collaborative work with the speech communities they study. Some of the major forces driving the field of documentary linguistics today are, in the first place, the need to create a lasting record of languages that can be useful to both speakers and linguists; and, in the second place, the acknowledgment that documentary projects in which community members are actively involved yield more broadly useful products than those in which the linguist works alone (see, for instance, Cameron et al. 1992; Cameron 1998; Dwyer 2006; Grinevald 2003; Himmelmann 2006; Austin 2010; Rice 2009, 2011; Yamada 2007, 2011; among many others). This paper takes Kokama as a case in point to demonstrate two things. First, collaborating with community members yields better data outcomes for linguistic studies and better support for community initiatives. Second, documentation projects must go beyond fluent speakers if the goal is to contribute to language preservation efforts. In the Kokama scenario, both fluent speakers and neo-speakers play a central role in ongoing initiatives. However, the language patterns produced by neo-speakers differ from those found in the speech of fluent speakers. This finding makes it all the more urgent to document language interaction among different types of speakers.

The paper is organized in three main sections. §1 offers an overview of the Kokama sociolinguistic landscape within the Amazonian context. §2 reports on a documentation project launched in collaboration with community members in 2003. §3 introduces ongoing revitalization initiatives and addresses the role of neo-speakers in those efforts. Finally, §4 presents the overall conclusions.

## **1. KOKAMA: PEOPLE AND LANGUAGE**

**1.1 LOCATION AND DEMOGRAPHY.** The Kokama people live in the Peruvian Amazon, approximately between 3° to 5° south latitude, and 73° to 76° west longitude. Kokama communities are located on the upper Amazon River and several of its major tributaries, including the Huallaga, Marañón, Ucayali, Nanay, and Itaya rivers (See map 1). The estimated 120 Kokama villages are located in the Peruvian department of Loreto, specifically in the provinces of Maynas, Alto Amazonas, Requena, Maquia and Ucayali. A considerable number of Kokamas now reside in cities within Loreto, such as Iquitos and Yurimaguas, and midsize towns like Lagunas, Nauta, Requena, among others. During the rubber boom,

towards the end of the nineteenth century and the beginning of the twentieth century, small groups of Kokamas migrated to Colombia and Brazil (Vallejos 2010b).



MAP 1. Location of the Kokama & Kokamilla communities within Peru.

In general, there is little agreement on the number of indigenous Amazonians in Peru, which is in turn visible in the different population estimates for the Kokama people. There are additional challenges to accurately calculating their population. First, speaking an indigenous language has for years been the defining factor to determine whether someone is indigenous or not, but the majority of Kokamas no longer speak their heritage language. Second, in Peru, there is a stigma attached to being part of an indigenous group, which generates unfavorable conditions for people to identify themselves as members of such groups. Third, a significant number of Kokamas have migrated from their villages to neighboring towns and cities where they have camouflaged themselves among the locals. If we put together the data provided by indigenous organizations themselves and by institutions working directly with Peruvian indigenous people, such as the *Instituto del Bien Común* (2010), the number of Kokamas living in 120 small villages is about 20,000. Perhaps the

most reasonable calculation for today's total Kokama population would be between 20,000 and 25,000, including people who no longer speak the language.

Based on linguistic and geographic criteria, the people identify themselves with two dialects: Kokamilla, which is primarily spoken in the upper Huallaga River, in the western side of the Kokama area; and Kokama, which is spoken along the Marañon, Samiria, Ucayali, and Amazon Rivers, towards the east. However, from a linguistic perspective, only a few phonetic and lexical differences have been found between these two dialects.<sup>2</sup> It should be noted that speakers are aware of the dialectal differences and are ready to point out how a specific word is pronounced by the other group if a difference exists. This is also reflected in explicit requests from the community to document dialectal variation. Having said that, there are no problems of intelligibility at any level. In this paper, I use 'Kokama' to refer to both varieties, and 'Kokamas' to refer to the people.

**1.2 SOCIOCULTURAL CONTEXT.** Despite the fact that the Kokama people have been interacting with non-indigenous groups for a long time, their cultural practices and their way of life are still traditional in many respects. The older generations, particularly, preserve vast knowledge of the forest and of the aquatic environment, which allows them to survive in an area of the jungle that is flooded almost half of the year. However, younger generations have been gradually abandoning some traditional practices, such as certain agricultural and fishing techniques and traditional medicine. A prominent aspect of the life in small Kokama villages is the bond of social reciprocity. For instance, a common practice is the *minga*, 'cooperative work.' Community members work together in order to support one another in the construction of a house, in the preparation of a new farm, during harvest, etc. Another instance of their cohesion is the reciprocal sharing of goods, especially fish, crops, meat, and other foodstuffs.

Shamanism, especially significant among the Kokamas, is also being slowly lost among younger generations. Shamans—*tsumi* 'wise-person'—have traditionally played a key role in the life of the community. They perform ritual ceremonies for everything from preventing overconsumption of natural resources to intimate curative events. Shamans can "advise, prescribe, encourage, point out dangers and make clear norms that prevent a man from making mistakes..." (FORMABIAP 2003:295). People seek the shaman when they are not able to solve a difficult situation by their own means. Health is the major area of concern for which people seek the shaman's help. Among the Kokamas, illnesses could be the expression of social misbehavior. According to Victor Yuyarima, a respected Kokamilla shaman and a member of the documentation team, *ikaros* or sacred songs exist for particular purposes, such as to 'cure' a person bitten by a snake, to make a newborn baby become a good hunter, to relieve a child's fear, or to make a new house endure, among others. Shamans learn the *ikaros* from their masters, and some may be chanted openly, whereas others must be kept secret. In the latter cases, the shaman only whistles while internally reciting the words in his or her mind. Possibly, *ikaros* are the deepest trait of spiritual bond among the Kokamas. During all of the years that I have worked among the Kokamas, I have not

<sup>2</sup> For instance, in the Kokamilla dialect, there is a general tendency to pronounce the [r] as a liquid [l], especially among women (i.e., *arara* / *alala* 'macaw'); diphthongs [ua] and [ui] in Kokama correspond to single vowels [a] and [i] in Kokamilla (i.e., *amui* / *ami* 'grandfather', *ramua* / *rama* 'other'); among others.

come across a young Kokama who has either chosen or has been chosen to be trained to become a shaman. This suggests that the younger generations might be gradually abandoning this cultural practice.

Nowadays, the Kokamas are inextricably intertwined with Peruvian national society. The location of Kokama villages along navigable rivers has facilitated the migration of a significant number of community members, especially youth, to big cities such as Iquitos, Yurimaguas, Nauta and Lagunas. Despite their wish to integrate into the mainstream culture, they are generally confined to the poverty belt of the cities, where they are looked upon with disdain. As a result, very often they end up ashamed of any trace of their ethnicity and they reject any association with their communities of origin. It is also true, though, that when they meet interlocutors who show signs of appreciation for other cultures, they openly talk about their communities of origin with nostalgia.

**1.3 SOCIOLINGUISTIC SITUATION.** As indicated earlier, the Kokama language is endangered. Based on the parameters provided by UNESCO (Moseley 2010), the Kokama language is severely to critically endangered. First, children are no longer learning the language, as natural processes of language transmission have been interrupted. Second, only about 5% (an estimated 1,000 individuals) of the total population speaks the language; the majority have already shifted to Spanish. Third, the remaining speakers are elderly people spread across small villages. Adults in their 40s and 50s show only passive knowledge: they are able to understand what is being said in the language, but are not able to produce even simple messages. Fourth, the remaining speakers use the language to speak amongst themselves, but only in very restricted situations, such as traditional events and intimate meetings. Fifth, there is no regular use of the language in new domains, such as the media, except for some sporadic initiatives, such as campaigns on the radio. So far, only a few materials for language education have been produced in Kokama by institutions like FORMABIAP, but this initiative is by no means part of a plan by the government to protect endangered languages.

Several historical reasons could explain the current sociolinguistic situation found in the communities. The Kokamas have a long history of contact with various linguistic groups, *mestizo* Spanish speakers being the most significant numerically, as well as in terms of social prestige. The introduction of schools in 1926, with instruction exclusively in Spanish, was one of the key factors for shifting to Spanish. Additional causes were the arrival of colonizers during the eras of the haciendas beginning in 1853, the rubber extraction initiated in 1885, and, more recently, the oil exploitation that began around 1970. The location of Kokama communities along navigable rivers with regular ship connections, plus their geographic proximity to big cities where Spanish is the dominant language have facilitated geographic mobility and migration. By the second half of the twentieth century, the Kokamas had intensified their contact with the ‘national society.’ Finally, the low social prestige of Kokama together with its use for limited communicative purposes provides little motivation for younger people to learn the language. All these factors have led Kokama people stop learning their language.

**2. THE KOKAMA DOCUMENTATION PROJECT.** This section elaborates on the successes, challenges, and lessons of a collaboration that emerged from my ongoing linguistic

research with the Kokamas. Because of the degree of endangerment of the language, it was clear from my first involvement with them that documenting and preserving Kokama were urgent tasks. To carry out the documentary work, I teamed up with two community members with the goals of: (a) documenting the language, traditional knowledge, and social practices through the collection of audio and video recordings; and (b) contributing resources to support the community's efforts to preserve their heritage language.

The activities reported here are ongoing, collaborative, and mutually beneficial. The remainder of this section describes them, starting with portraits of the community members who became leaders in the project (§2.1), the speakers from whom the data was collected (§2.2), and the strategies used to gather materials for the products of the project (§2.3).

**2.1 TEAM PARTNERS.** A basic premise of the documentation project is that each participant contributes unique knowledge and expertise. Given that from the start both description and preservation of the language were the ultimate goals of the documentation project, it was essential to build a team of community members to carry out the documentary activities for two reasons: firstly according to Rice (2011:51), “the more the native speaker is invited to shape the record, the richer the documentation of the language”, and secondly, engaging community members in the documentation project leaves the leadership where it belongs: with the speech community. Thus, if members of the community assumed the project as their own, the expected next step would be commitment towards the preservation of the language.

To launch this project, I teamed up with Victor Yuyarima Chota and Rosa Amías Murayari, who are members of Kokama communities. Both of them have played key roles at different stages of the project and will be referred to here as team partners. In addition, the project has benefited from the contributions of Pascual Aquituari Fachín, a Kokama teacher. My role in the documentation project has been that of directing logistical efforts for the project. I worked with my partners to identify potential activities, look for financial support, coordinate field trips, deal with equipment, carry out recordings, and process raw data into different formats. Below I highlight their contributions to these and other tasks of the project.

**2.1.1 ROSA AMÍAS MURAYARI.** *Doña Rosita*, also known among young Kokamas as *mama Rutsa*, is originally from the community Dos de Mayo, San Pablo de Tipishca, along the Samiria river. *Doña Rosita* speaks the Kokama dialect, but she is able to identify the cognates in the Kokamilla dialect when relevant. At the beginning of the documentation project, she was 58 years old. Rosa Amías is a great storyteller and an excellent interviewer. She is highly fluent in Kokama and Spanish, and also has some fluency in Quechua. She is one example of the many people of her generation that were punished for speaking their language at school. As a result, she did not teach any of her ten children either Kokama or Quechua. That is, all her children are now monolingual in Spanish.

*Doña Rosita* has been collaborating with me in my research since 2001, and with the Documentation Project since its inception in 2003. Rosa Amías has participated mainly in the collection and processing of the data, as well as in the creation of resources for the community. We traveled together every year from 2005 to 2009. During these trips, she showed an extraordinary ability for engaging and motivating community members who are

generally reluctant to have a discussion about their heritage language, let alone to talk in their language. She has conducted most of the interviews during the multiple field trips. In addition, Rosa Amías and I worked together in the transcription, glossing, and translation of the texts. She is the one who contributed most of the elicited data collected during the analysis of the texts. Over the years, she has become a community linguist with extraordinary intuition about the grammar of her language. Currently, she is the main collaborator of the Kokama-Spanish-English Dictionary Project.

**2.1.2 VICTOR YUYARIMA CHOTA.** Victor Yuyarima is originally from the community Ocho de Octubre, quebrada Paucaryacu, along the Huallaga River. He speaks the Kokamilla dialect. At the beginning of the documentation project, he was 64 years old. He is a respected shaman and very knowledgeable about Kokama traditional practices. But above all, Victor is a well-known leader among the Kokamillas. He is one of the founders of FEDECOCA, the first Kokama Indigenous Federation, and has represented his people in different venues, at national and international levels. I started to work with Victor Yuyarima in 1999, when both of us were part of the FORMABIAP project. From 2002–2006, we worked together occasionally during short summer visits.

In 2006, Victor Yuyarima, Rosa Amías, and I, together carried out a two-month field trip and visited six communities along the Huallaga River. Victor, being the local, took the lead in making the connections with authorities, setting up meetings for us, explaining the project to locals, and building trust in general. He was also efficient in identifying and contacting speakers of the language. Recall that in highly endangered situations, this is not an easy task. Given his great access to the social network of the area and his position in that network, he was able to motivate elders—and sometimes convince their relatives—to be interviewed. During the recordings, he carried out many of the interviews, especially with male speakers, and later helped gloss and translate a portion of these interviews. I collected many traditional stories from him, and recorded him performing several curing songs known as *ikaros* (see §1.2). The status of the *ikaros* within Kokama society is sacred. They are only known and performed by shamans who are chosen at an early age by a master and trained for almost all of their lives. Before recording the *ikaros*, Victor and I discussed the advantages and potential disadvantages of creating a permanent record of these songs. He understood the importance of leaving a legacy for the next generations, but he made it clear that access to specific pieces of data needs to be restricted from the public. In the end, he was willing to be recorded because he knew that we would follow a protocol to assign different levels of access to all the data we were recording, not only with him, but with any speaker. Within the team, Victor was the one who contributed the most to the overall discussion about data protection and levels of access.

**2.1.3 PASCUAL AQUITUARI FACHÍN.** Pascual is an elementary school teacher who in 2005 was 46 years old. He is from Lupuna, a Kokama community only 20 minutes by boat from Iquitos, the biggest city at the core of the Peruvian Amazon. He learned Kokama from his mother and grandparents when he was a child, but he stopped using it during his teen years. Later, he was invited to participate in several initiatives, mostly advanced by the Catholic Church, looking to promote higher self-esteem and recognition of their indig-

enous identity among the Kokamas. In all these activities, the Kokama language played a key role, and so Pascual's language skills in Kokama were greatly appreciated.

Once he became aware of the impending loss of Kokama, Pascual became a language activist and remained engaged in several language revitalization and language maintenance efforts.<sup>3</sup> Over the years, he worked as an instructor in charge of training future Kokama elementary school teachers, as well as developing instructional materials for teaching Kokama as a second language. During these years, I contributed to the training of Pascual Aquituari in the grammar of Kokama and on basic linguistic description. Pascual, being a teacher, provided significant input to the discussion about the relevance and appropriateness of particular language resources for the community. For instance, we worked together in the creation of a bilingual DVD (see §2.3.3). He recorded online translations to create bilingual clips. At present, he acts as local representative of the Ministry of Education and/or as a consultant for NGOs with projects in the Kokama area.

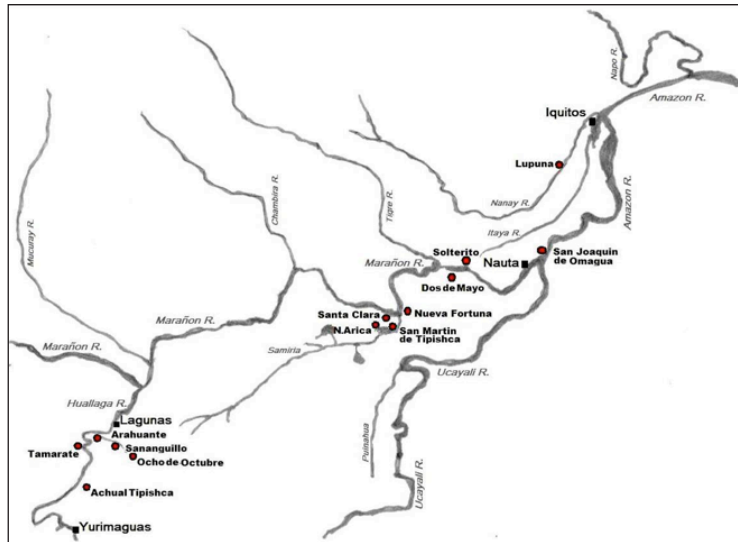
Between 2005 and 2011, the team has conducted multiple research trips. The exact locations of the communities visited as part of the documentation project are shown in Map 2. During these trips, we met fluent speakers, semi-speakers and neo-speakers. Most of the interviewed people would be classified as either highly fluent speaker or neo-speaker (defined in §2.2). But we also met a number of semi-speakers, the majority in their fifties, who gathered around us while we were doing the recordings. They would laugh at the stories being told by the elders, and respond with isolated Kokama words, or just make comments in Spanish. All in all, it seemed that listening to the language was a moment of celebration of being Kokama.

In sum, by the end of 2011 we had collected samples from fifteen villages, two midsize towns (Nauta and Lagunas), and two cities (Yurimaguas and Iquitos). Working along with community members resulted in the success of collecting this pool of samples. The leadership and connections of Victor Yuyarima, the natural documentary abilities of Rosa Amías, and the pedagogical experience of Pascual played a vital role in the many successes the project has experienced. Rosa Amías and Victor Yuyarima were extremely successful not only at finding fluent speakers but also at persuading them to talk freely in their language. Working without members of the speech community would not have resulted in the invaluable connections and positive outcomes we were able to achieve. The following sections stress the contributions of these and other community members in creating a rich catalog of recordings useful to both the linguists and the community.

---

<sup>3</sup> In 2009, Pascual suggested that we visit his community, Lupuna. This was a personal mission for him because his father, who was a fluent speaker of Kokama, had recently passed away, and no one had ever recorded him. When his father died, it became clear for him that recording his mother (85 by then) was a task that could not be delayed. During our visit to Lupuna, Pascual and Rosa Amías conducted several interviews. It was extremely moving to observe Pascual talking to his elderly mother about the life in the community, their ancestors, etc. and recalling many intimate stories. In the summer of 2013, I visited Lupuna again and learned that Pascual's mother had recently passed away.





MAP 2. Kokama &amp; Kokamilla communities of research

**2.2 DOCUMENTING LANGUAGE VARIATION.** In endangered language situations, a range of speaker types can be found (Dorian 1980, 1987; Grinevald 2003; Grinevald & Bert 2011). However, what does it mean to speak the language? What qualifies someone as a fluent speaker? What do speakers know of a language that is no longer used on a daily basis? These became recurrent questions during our multiple field trips. But most interestingly, people can have different perspectives on what it means to know an endangered language (see, for instance, Leonard & Haynes 2010). Grinevald & Bert (2011) propose four interrelated parameters for a typology of seven types of speakers:<sup>4</sup> fluent speakers, semi-speakers, terminal speakers, rememberers, ghost speakers, neo-speakers, and last speakers. Language documentation projects typically focus on the speech of fluent speakers.

But how can the fieldworker know, in the first place, that such a diversity of competence in the language exists in a particular speech community? And even if we are aware of this diversity, how can we know what type of speaker we are confronted with in a given session? In the Kokama scenario, the involvement of community members made it possible to find answers to these questions. The role of my team partners in recognizing the types of speakers among all who declared knowing the language was crucial. To date, the team has gathered primary data from 38 speakers in total. For each person, we collected data about age, gender, origin, and language(s) and dialect(s) spoken. This information was either video or audio recorded at the beginning of each interview. At the end of each interview session, Victor Yuyarima and/or Rosa Amías assigned each speaker to a category. Their judgments were based primarily on the speakers' ability to interact and engage in spontaneous conversations with them. Interestingly, 'understanding jokes' was an addition-

<sup>4</sup> These four parameters are: (i) language competence (level of acquisition attained and degree of individual loss), (ii) exposure to language versus vitality of the language at time of acquisition, (iii) use and attitude, and (iv) selfevaluation of language skills.

al test for all fluency types. Given that both Victor Yuyarima and Rosa Amías are extremely extroverted, they started conversations with jokes to first find out whether someone was able to understand Kokama or not. If community members laughed at the jokes, they were further engaged in conversations on different topics.

The categories that emerged from Victor Yuyarima and Rosa Amías' judgments were basically three: *habla bien* 's/he speaks well,' *sabe, pero necesita practicar* 's/he knows but needs to practice,' *todavía está aprendiendo* 's/he is still learning.' Additionally, Victor and Rosa Amías would make comments about other people that we met, such as *sabe, pero se averguenza* 's/he knows but feels embarrassed.' Respectively, I translate these categories as the following: fluent speaker (§2.2.1), semi-speaker (§2.2.2), neo-speaker (§2.2.3) and ghost speaker (§2.2.4). We collected data from the first three groups. This data can make crucial contributions to the discussion about language change and language attrition in endangered language situations. I elaborate on this in §3.

**2.2.1 FLUENT SPEAKERS.** *Fluent speakers* are all those able to engage in spontaneous conversations. Such speakers are able to provide narratives with very minimal use of borrowings. The average age of the highly fluent speakers we interviewed was 69 years old. It should be noted, though, that collecting information about ages was not a straightforward task because a number of the elders did not remember the year they were born.<sup>5</sup> Based on the observations provided by relatives, we were able to determine an approximate age for some of them. In this group there were two blind persons, and a third with serious vision difficulties. Also, two of them were no longer able to move around without assistance.

These fluent speakers were the type most sought after by the team. Thirty of the 38 speakers interviewed were classified as fluent speakers. Two of them had learned Kokama as a first language, and only later learned Spanish. The majority of them learned Kokama and Spanish simultaneously. Also, a few of them have some command of Quechua. During our field trips, we found an average of one or two fluent speakers per village. In most cases, they had not used the language in a very long time, and so they were thrilled to finally have interlocutors in the language. My team partners showed a profound sense of respect for these speakers, which I believe was in part because they spoke the most fluent Kokama that we had heard in a long time, but also because of the standing elders hold within the Kokama society.

As for the type of data collected from this group, traditional stories, personal narratives, and spontaneous conversations were by far the favorite genres. Note that in each interview, the elders were never left alone to talk to the camera; they always talked to either Rosa Amías or Victor Yuyarima, or both. Because Victor Yuyarima and Rosa Amías were in charge of the interviews, the collected material was quite different in nature compared to the data I collected during my first trips. For instance, when I conducted the interviews, the language use was not only prosodically more plain, but minimal use of bound morphology was the dominant pattern. Clearly, when speakers were talking to me, they were in a sense simplifying their speech. In contrast, the data collected by Victor Yuyarima and Rosa Amías from fluent speakers included more prosodically vivid and morphologically complex language. If my team partners had not conducted the interviews, we would have

---

<sup>5</sup> In the Amazon, it is common to find elderly people who do not have a birth certificate or any other type of identification document.

ended up with very different data.

**2.2.2 SEMI-SPEAKERS.** This category of speaker is representative of endangered language contexts. *Semi-speakers* are those with passive knowledge, that is, they are able to understand almost everything but with limited production. The production of these speakers is full of borrowings and code switching with Spanish. We collected data from only three semi-speakers. At the time of the recordings, their ages were 48, 54, and 59. They learned Kokama as children, growing up with their grandparents, but stopped using it at an early age. One of them reports achieving complete acquisition of Kokama, the other, only partial acquisition. In both cases, their grandparents were monolingual in Kokama, but their parents were bilingual in Kokama and Spanish. When their grandparents passed away, there was no longer a need to speak Kokama. They were very enthusiastic about contributing to the project. Their desire to respond in Kokama was obvious, as was their frustration at their inability to do so with the fluency they would have liked. Their speech was full of Spanish elements (lexicon and grammar), and at some points of the conversation they would completely switch to Spanish. They attributed this fact to not having heard the language in a very long time, since no other speakers of Kokama lived nearby.

Although the recommendation in endangered language situations is to not turn away any member of the language that expresses interest in contributing data (Grinevald & Bert 2011), collecting data from semi-speakers was not a priority at the beginning of the project. There were several reasons for this. First, there was a clear sense of urgency to make recordings of fluent speakers, given their advanced age. Second, there were constraints on time and resources. Traveling in the Amazon entails a significant investment in both, and so we decided to focus our efforts on the group that needed the most urgent attention. Third, my team partners had ambivalent feelings about semi-speakers. On the one hand, they seemed to understand the reasons why these speakers had limited abilities in the Kokama language. Yet at the same time, there was a very subtle judgment for what was perceived as a lack of effort and commitment to keep their language alive. In a few cases, my team partners decided that some volunteers did not meet their standards. For instance, people that could only remember words were turned away. After we reflected on the importance of engaging as many people as possible, things changed a little. There were a few occasions when we were requested to record particular speakers. This was the case with a neo-speaker who happened to be the grandson of one of the elders we had interviewed. The proud grandfather explained he put a lot of effort into teaching his grandson the language, and insisted in showing us how much he had learned. My team-partners were hesitant, but I was excited about the possibility of recording the elder talking to his grandson. After we agreed it was worth trying, we invited both of them. The grandson was able to partially understand what his grandfather was saying but had major difficulties producing Kokama. He was classified as a neo-speaker.

In the end, the recordings from semi-speakers ended up not only satisfying the desires of community members such as this grandfather, but also provided useful information for topics of theoretical interest in linguistics. The documentation of speech patterns produced by semi-speakers is of particular significance for language attrition research.

**2.2.3 NEO-SPEAKERS.** Neo-speakers are the product of language revitalization initiatives. They were born and raised in Kokama villages. Given that most of them have lived surrounded by relatives with some competence in the language—especially their grandparents—they have had some exposure to the language at an early age. It is important to emphasize, though, that in most cases this exposure has been minimal. For instance, some of them report hearing simple commands in Kokama from their grandparents. However, they did not acquire any speaking abilities in the language while they were kids. They are now learning Kokama as a heritage language.

We collected data from five neo-speakers. All of them had been trained as bilingual teachers by the FORMABIAP project. One component of their training was precisely to learn Kokama through second language methodologies, complemented by immersion in their communities of origin. The goal of their training is to be able to teach the language in elementary schools. The levels of fluency achieved by these neo-speakers range from minimal to somewhat competent.<sup>6</sup> Most of them are able to follow conversations, but only some of them actually produce connected speech. These neo-speakers are able to engage with elders, who are generally extremely collaborative speakers, yet none of them has been able to master (understand and produce) the nuances of the language produced by elder fluent speakers (see §3 for examples).

It should be added that elder fluent speakers, and community members in general, have positive attitudes towards neo-speakers. In fact, elders are thrilled about these young adults' desire to learn the language and praise their efforts.<sup>7</sup> They value their attempt to learn Kokama, given the social pressures to abandon this endeavor altogether. This positive attitude may also be linked to the fact that most neo-speakers are community teachers, which in itself carries relative prestige and authority. It is worth noting that this admiration and respect had an impact on neo-speakers. They seemed generally satisfied with the level of competence achieved, whatever that might be. This type of speaker is the most aware about the disappearance of the language, and language preservation is part of their discourse. Some of them are actively engaged in revitalization initiatives.

**2.2.4 GHOST SPEAKERS.** This category consists of community members who “conspicuously deny any knowledge of the endangered language in spite of evidence that they do have some level of competence” (Grinevald & Bert 2011:51). This denial is a consequence of the negative attitudes associated with speaking the language. Those community members that, according to Victor Yuyarima and Rosa Amías, *sabe, pero se averguenza* ‘s/he knows but feels embarrassed,’ belong to this category. It is important to note that my partners' assumptions might have subjective bases. As indicated before, Victor Yuyarima and Rosa Amías used a simple strategy to find out whether someone was able to understand Kokama or not: they started conversations with jokes. If community members laughed at the jokes and showed signs of understanding what they were being told but later denied knowing Kokama, they were assigned to this category. We found this situation in a number

<sup>6</sup> Besides the five teachers interviewed for the project, there are a significant number of trained Kokama bilingual teachers for whom no data on language competence is available.

<sup>7</sup> During my last field trip (Summer 2013), I met a group of elders based in the midsize town of Nauta who seem to resent the protagonist role some neo-speakers are taking in the revitalization of the language. However, they are still trying to work together.

of communities, but especially among the people that have migrated to midsize towns, such as Nauta and Lagunas, or big cities, such as Iquitos and Yurimaguas. As indicated in §1.3, in urban areas, the Kokama immigrants generally choose to deny any associations with their indigenous identity, including their heritage language.

To close this section on types of speakers, I would like to emphasize that the involvement Victor and Rosa played in assessing all of the fluency types was key. The criteria they implemented resulted in accurate judgments. A linguist working on his or her own would have had difficulty, first of all, coming up with simple and culturally relevant strategies to assess fluency in Kokama; and, secondly, making decisions as to which category particular speakers belong to.

**2.2.5 AGE VERSUS LANGUAGE COMPETENCE.** Including Victor Yuyarima, Rosa Amías, and Pascual Aquituari, we collected data from 38 speakers in total. 30 of them were categorized as fluent speakers, and their ages ranged from 50 to 86. Three speakers were categorized as semi-speakers, whose ages were 48, 54, and 59. Finally, five were classified as neo-speakers, and their age range was between 26 and 37.

If we correlate fluency level with age for all the speakers interviewed, we see a clear pattern. This is summarized in figure 1 below:

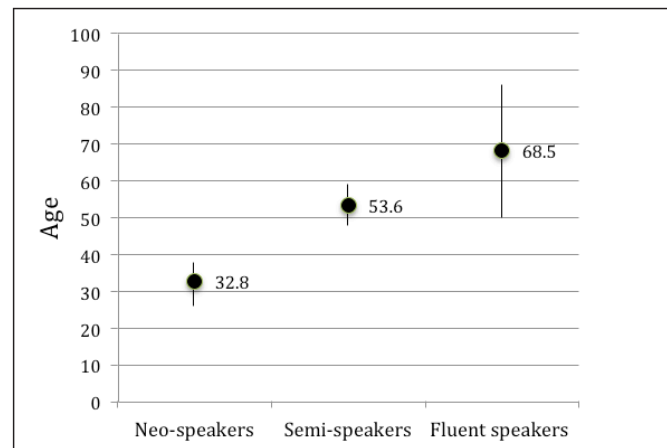


FIGURE 1. Types of speakers and their average age

While fluent speakers are advanced in age, neo-speakers are in their 30s and constitute the youngest generation of speakers. They are taking the lead in the preservation of Kokama. However, the language of neo-speakers differs in various respects from the patterns mastered by fluent speakers. The question is, then, what aspects of the language should be preserved and eventually transmitted to the next generations? Different community members answer this question differently. For instance, some elders would like to preserve the most salient features of the language (i.e., the female/male speech distinctions), while the least fluent speakers seem pleased with knowing at least highly functional language and simple constructions. Finally, community leaders tend to favor the incorporation of lexical items into their discourse to highlight their Kokama identity. However, the question of

preservation is also often rephrased into, “What will Kokama be good for in the future?” While the public discourse continues to cling to the ideal of putting Kokama back into the life of the communities, a more realistic view that includes as a goal at least a symbolic use of the language in public spheres (e.g., signs, posters, radio programs, cultural festivals, traditional celebrations, etc.) seems to be slowly emerging.

**2.3 DATA GATHERING AND OUTCOMES.** In language documentation for revitalization, the model is now moving away from seeing language as a system to assuming language as practice. Documenting social practices is crucial because language maintenance is about documenting and promoting not only the use of a linguistic code, but also patterns of communicative interaction. In that regard, documenting interactional patterns from natural settings can make crucial contributions to language maintenance because this data shows not only patterns of language use, but also is full of sociocultural knowledge.

Our aim during the visits to the Kokama villages was to find as many speakers as we could, and to record natural speech in as many contexts and genres as possible. Up to now, the main product of the documentation project is a corpus of digital recordings, which includes video, audio and photographs.

**2.3.1 BUILDING A TEXT CORPUS.** The collection of data within the Kokama-Kokamilla Documentation Project privileged video recordings. Beyond preserving Kokama traditional knowledge, the recorded material can be later developed into classroom resources, so that the sounds and patterns of the language can be reproduced in formal instructional settings. We recorded texts such as traditional stories, stories from daily life, personal experiences, spontaneous conversations, songs, etc. Speakers were interviewed individually, in pairs, and in groups. While one-on-one sessions created a stress-free environment to talk about personal experiences, sessions in groups elicited the language in real communicative situations with a true speaker-hearer relationship to ensure the most natural communicative speech. Also, since either Victor Yuyarima or Rosa Amías conducted the interviews, every recording includes at least one highly fluent speaker. Every speaker knew that if they wanted to stop the recording, they could do so at any point. Thus, when a speaker would say *aynatan* ‘this will do it,’ we knew that it was the signal to pause the recording, or stop it all together.

A point worth mentioning is the issue of genre in the collected data. The favorite topics were by far traditional stories and narratives about events in the community. The resulting data can be skewed and potentially not contain the widest range of grammatical constructions that the language has to cover a variety of discourse functions. We attempted to correct this ‘problem’ with relative success by making suggestions of topics for the interviews. A related issue is that, initially, one of the interviewers would introduce herself at length in almost every recording and with every new language consultant. However, suggesting to her what to say and how much to talk did not seem ethical. This was corrected after we worked together in the transcription of these first pieces of data. The speaker herself very soon noted that not only did we have to transcribe almost the same thing over and over, but also that in most of the tracks she had the tendency to monopolize the conversation. As a result, in subsequent interviews she introduced herself briefly, giving the interviewee more room to talk.

Because of the age of the speakers and technology-related issues while in the villages, we did not pursue recordings using stimuli to elicit controlled data.<sup>8</sup> As indicated earlier, the fact that the community members themselves conducted the interviews had a profound impact on the quality of the video recordings that we were able to collect. On the one hand, it was clear that the topics of conversation clearly mattered to all the interlocutors involved. Some of the recurrent topics of conversation were the origins of the villages, the multiple migrations throughout the years, their family connections in other villages, etc. Those were themes relevant to both the interviewers and the interviewees. On the other hand, the collected material was quite different in nature than the data I collected during exploratory trips. In Kokama, no morphology is obligatory, hence sentences full of bare words are perfectly grammatical (see §3). Clearly, when speakers were talking to me, they were in a sense simplifying their speech. The language speakers used when ‘talking’ to me consisted mostly of bare morphological units. In contrast, the data collected by the team members included more vivid and morphosyntactically complex language, rich in metaphorical figures and prosodic contours. With Victor Yuyarima and Rosa Amías in charge of the interviews, the quality of the collected data improved dramatically.

03AGO01-VY-019									
<i>tana</i>	<i>kumitsa</i>	<i>ikian,</i>	<i>ikuatawaratsui</i>						
<i>tana</i>	<i>kumitsa</i>	<i>ikian</i>	<i>ikua</i>	<i>-ta</i>	<i>-wara</i>	<i>-tsui</i>			
1PL.EX.M	say	this	know	-CAU	-NZR	-DAT			
1PL.EX.M	decir	este	saber	-CAU	-NDR	-DAT			
we talk to the teacher									
hablamos con el profesor									
03AGO01-VY-020									
<i>rutsutsen</i>			<i>yurimaka,</i>		<i>rakumitsatsen</i>		<i>maikanamuki.</i>		
<i>ra-</i>	<i>utsu</i>	<i>-tsen</i>	<i>yurimaka</i>	<i>ra-</i>	<i>kumitsa-tsen</i>	<i>mai</i>	<i>-kana</i>	<i>-muki</i>	
3SG.M-	go	-PUR3	Yurimaguas	3SG.M-say	-PUR3	spirit	-PL.M	-COM	
3SG.M-	ir	-PROP3	Yurimaguas	3SG.M-decir	-PROP3	espíritu	-PL.M	-COM	
to go to Yurimaguas, to talk to the mestizos									
para que vaya a yurimaguas, para que vaya a hablar con los mestizos.									

FIGURE 2. Text database

To date we have recorded approximately 20 hours of video and four hours of audio. It has not been a trivial undertaking to make these recordings, as they were collected in communities where the heritage language is no longer used for daily communication. Because working with video is extremely time-consuming, only about five hours of video have been fully transcribed, glossed and translated. Parsed texts are in a six-line interlinear format

<sup>8</sup> However, in the city of Iquitos, I gathered some video material from Rosa Amías and Victor Yuyarima, with focus on the lexicon and grammar associated with space. For this task, I selected a set of pictures with different elements in them. I explained to them that one speaker would describe a scene while the other would try to reproduce it on a separate sheet of paper. Next, they would compare the original picture and the drawing, and discuss the results. They found it interesting at first, but very soon thought it was a childish task and suggested moving on to document what they viewed as more important and crucial aspects of the Kokama culture.

consisting of: (1) the standard Kokama orthographic representation which expresses morphophonemic differences and the widespread final vowel epenthesis and cliticization that is characteristic of Kokama, (2) underlying representations of individual morphemes, (3) English morpheme glosses, (4) Spanish morpheme glosses, (5) English free translations, and (6) Spanish free translations. This has resulted in approximately 800 pages of text data. Note that 138 pages of this text data have been made available to the public (Vallejos 2010b). A sample is presented in figure 2.

It is important to highlight the collaborative approach to the building of this text database. The leadership of community members in data gathering has crucially contributed to the naturalness, richness, and variety of the collected material. While processing the data, the speakers provided innumerable comments and additional examples that complement the patterns found in texts. The resulting data has a tremendous value for both scholarship and language preservation initiatives. For the community, this text corpus constitutes priceless records of the language. If we recognize that language maintenance is not only about promoting the use of a linguistic code, then we can see that the documentation of interactional patterns can be essential for revitalization initiatives. For instance, the naturalness and variety within the collected data is useful to revitalization because teachers and learners will have access to more realistic data for their own learning and lesson planning. For linguistics, text corpora for a lesser-known language make it possible to account for patterns for which a good understanding of both the discourse context, and the sociocultural context in which the language is spoken, is crucial. For instance, Vallejos (2009), in a study based on a portion of the text data described here, found that Kokama has constructions that clearly distinguish focus subtypes according to narrow/broad scope and contrastive/non-contrastive information. This finding bears on theories about whether contrast simply emerges from conversational implicature, or whether it can be explicitly encoded by dedicated grammar. A more recent study (Vallejos 2014) examines three purpose constructions in Kokama. These constructions exhibit different degrees of syntactic integration with the matrix clause and entail different coreferentiality conditions between arguments. Discourse data demonstrates that in Kokama purpose clauses, syntactic integration between clauses correlates not with the semantic integration between events alone, but with discourse-pragmatic factors (temporal integration, successful outcome, and referentiality of arguments). In contrast, observations based on data collected exclusively in elicitation contexts can be problematic, but this is more so when the language is no longer used on a daily basis. I have experienced this firsthand when studying ditransitive constructions in Kokama (Vallejos 2010a). For instance, judgments of the grammaticality of particular patterns were rarely categorical. However, an examination of Kokama texts found that semantically three-place predicates are syntactically encoded by means of transitive clauses; this finding adds to the literature that questions constructions with two objects as a basic argument structure pattern. In sum, having systematically analyzed texts linked to their correspondent sound and video files is essential for quality, replicable linguistic research.

**2.3.2 FROM A LEXICAL DATABASE TO A TRILINGUAL DICTIONARY.** With the community in mind, we wanted to present the lexicon generated when glossing the texts in the form of a dictionary. In designing the Kokama-Spanish-English dictionary, consideration of potential users has been a primary concern. We expect the proposed dictionary to be of



use to (a) the Kokama speaking community, (b) community members who wish to learn their heritage language, and (c) other scholars. All decisions regarding the proposed dictionary are made with the needs of these three communities in mind. Considering the issues outlined in Frawley, Hill & Munro (2002), in addition to general design, we gave careful consideration to the entries, literacy of the potential users of the dictionary, use of graphics, among other issues.

The team has employed primarily two methods to collect lexical items: (a) collection of words via spontaneous speech, especially procedural texts, and (b) active eliciting methods (Mosel 2004). Each dictionary entry consists of a Kokama headword with its Spanish definitions, English gloss(es), word class, examples of the word in context, and semantic field. When relevant, the entries will also contain dialect variation (Kokamilla), scientific name, pragmatic and ethnographic notes, drawings or photographs, and cross reference for synonyms, antonyms, or hyponyms. It should be emphasized that the Kokamas have expressed tremendous enthusiasm about having English glosses/example translations in the dictionary. There are two reasons for this. First, English is an extremely high-prestige language in the area, even more so than Spanish. The sole idea of having English vis-à-vis Kokama is perceived as added prestige to the language. Second, an important number of Kokama villages are located within the Pacaya-Samiria National Reserve,<sup>9</sup> an area that has experienced a significant increase in ecotourism in the last five years. Having English in the dictionary responds to the community's desire to develop basic skills in English to be able to interact with their visitors. At the same time, English translations will make information about the Kokama people and the language accessible to foreign scholars interested in Amazonia. Figure 3 is an example of a dictionary entry.

**tsaichimi** *Kokamilla: saichimi. N. pampanilla, falda de algodón tradicional actualmente en desuso. Traditional skirt made out of cotton. La pampanilla consistía en un tejido de un sólo paño elaborado en forma de 'pretina' ancha a la cual luego teñían de colores rojizo, azulino o negro. Los tintes eran elaborados a partir de hojas, lianas y raíces de árboles. La pampanilla se usaba envuelta a la altura de la cintura y sujetada con un cordón. tsa nai chirura tsaichimi ukuatsuri. 'Mi abuela solía usar pampanilla.' 'My grandmother used to wear traditional skirt.'* **patapu paririkan amaniutsui inu yauki tsaichimi ukuatsuri iminua.** 'Antiguamente, ellos hacían pampanilla del algodón torcido con husos'. 'Long time ago, they made traditional skirts out of cotton spun by spindle.'  
 Categoría: **accesorios.** Ver: **chiru.**

FIGURE 3. Example of dictionary entry

We have made significant progress with a fraction of the entries; however, this task is

<sup>9</sup> For more information, see <http://pacayasamiria.org/>.

still in progress. A dictionary containing such rich information can only be possible with the decisive contributions of the speakers. This rich encyclopedia of Kokama knowledge and cultural history will be an essential tool for implementing initiatives towards preserving the language. For instance, the examples of specific words in natural contexts, the grammatical information, the ethnographic notes, etc. could be incorporated into lesson plans. For linguists, this product will make it possible to advance studies on lexical semantics, comparative-historical linguistics, and dialectology, among others.

**2.3.3 OTHER RESOURCES FOR THE COMMUNITY.** In the last two decades, there has been an emphasis on the notion of community empowerment as a guiding principle in research on endangered languages and the need for researchers to understand the contemporary ethnographic situation in the region in which they plan to work. However, as Dobrin (2008) emphasizes, what can be considered empowerment in one field situation is not necessarily so in another. It could be the case that the linguist may think the community needs are being met, but this does not mean that is truly the case (Dobrin 2008, Rice 2011). Within the team, we had discussions about the possibility of creating audiovisual material for Kokama language teaching and training speakers in documentary techniques, both strategies that have proved successful in other contexts (see, for instance, Yamada 2007). It is important to remember, however, that what could be a useful resource or strategy in one revitalization context could be less appropriate in an Amazonian village.

During the field trips, the Kokama participants often expressed their desire to have copies of the recordings. Around 1999, while working at FORMABIAP, I had the opportunity to record traditional music and lyrics and produce a cassette tape. This tape, which was distributed to some community members, made a major impact on the villages. I personally witnessed people celebrating their first (and still only) collection of Kokama songs. They would play the tape over and over until all of them—especially children—could sing the songs. That is, audio materials have proven to be powerful in connecting people with their linguistic roots. Inspired by this experience, we thought that a DVD with a set of clips would be a fine resource for the community. For instance, given that it is difficult to find fluent speakers of the language in many areas in the Kokama territory, a DVD can be used to reproduce the sounds of Kokama in formal instructional settings. Once the team selected the clips, with the collaboration of Pascual Aquituari, we created online translations into Spanish. The DVD includes five traditional songs (with subtitles in Spanish), four folk-stories (in two versions: Kokama and Spanish) and three personal narratives (in two versions: Kokama and Spanish). So far this DVD has been distributed to the community members that have participated in the project, and to some Kokama teachers. However, the impact of video materials in the Kokama scenario remains to be determined. First, schools in the villages do not have the required equipment to play CDs and DVDs. Most teachers have such equipment, but because the fuel to run electric generators is expensive, they use this equipment rather sporadically. Second, CDs and DVDs are extremely fragile media, more so in the Amazon. Thus, for now, DVDs like the one described above might be more relevant among Kokamas living in midsize towns and cities than for the people living in the villages.

A number of projects report that training native speakers in documentary linguistics is one of the most important tasks towards empowering the speech community. It turns

out that training speakers in the use of pieces of software such as Transcriber or Toolbox is not a priority among the Kokama. There are many reasons for this. First, the majority of communities do not have power; hence the lack of electronic equipment such as computers is the norm. Second, as indicated in §2.2, fluent speakers of Kokama are mostly older than 60. Although one of my collaborators has shown genuine interest in learning to use the computer for basic tasks, typing has been a challenge for her. This would mean that we need to train teachers who are interested in learning the language to work in collaboration with elder speakers. However, Kokama teachers are generally more interested in classroom strategies. Third, there is the question of how training speakers in something they will not be able to use beyond the context of the documentation project can be considered empowering the community. How would they use such training on their own, once the grant money runs out? Among the Kokamas, the most desirable training is not in documentary linguistics *per se*, but in language teaching methodologies, creation of language resources, and mobilization of data in general.

**3. LANGUAGE REVITALIZATION EFFORTS AND THE ROLE OF NEO-SPEAKERS.** As one may expect, within the Kokama community it is possible to find a range of attitudes towards the heritage language. People's positions vary from indifference, to relief, to grief, to desire to do something to prevent the disappearance of the language. Although it is still common to meet community members who are ashamed of their indigenous background, when given the opportunity to talk about their origins and personal history, they lament the disappearance of their language. Like any other ethnic group, the Kokamas want a connection to their roots and their history. Since the early 1980s, they have promoted language revitalization efforts, including annual meetings with the sole purpose of speaking their heritage language and developing local radio programs produced in Kokama. These two activities have been developed with support from the *Centro de Capacitación Campesina de la Amazonía* (CENNCA) and the Vicariato Catholic Church-Nauta, Loreto. Other initiatives supported by Catholic priests working in towns such as *Santa Rita de Castilla* (Marañón River) are campaigns for the preservation of the local environment and against the irresponsible practices of oil companies. At the center of these movements are the Kokama identity and the Kokama language.

Perhaps one of the most important initiatives has been teaching Kokama as a second language in the schools, as promoted by institutions like FORMABIAP. Given the sociolinguistic context of the Kokama communities, the heritage language had to be treated as the second language for the children. But the reality was that the elementary school teachers did not speak enough Kokama to be able to implement bilingual education in the schools. Thus, having a generation of Kokama teachers was a priority (Trapnell & Neira 2009). As a product of the efforts of indigenous organizations, educational programs, community members and language activists, there are already a group of neo-speakers of Kokama. The overall results of these revitalization efforts are still modest. However, the fact that a number of learners have gained some command of the language cannot be underestimated. This is especially noteworthy given the social pressures on Kokama young adults to abandon their traditional practices, including their heritage language. At present, there are approximately 55 bilingual elementary schools within the Kokama territory. In a number of those schools, neo-speakers are currently implementing Kokama as a second

language, particularly in the first levels of the school system.

As indicated in §2.2, at the beginning of the documentation project, the team considered that, given the limited resources, documenting the speech of elder, fluent speakers should take precedence over other types of speakers, including neo-speakers. However, we understood the importance of having data from learners of Kokama as a second language. First of all, these are the speakers who are expected to play a key role in the preservation of the language, and if their speech is not documented it will be impossible to set realistic goals regarding language teaching and learning in the school contexts. Furthermore, the documentation of speech patterns produced by semi-speakers and neo-speakers is of particular significance for linguistic theory, as language decay provides scenarios for accelerated language change. Data from Kokama can contribute to this discussion.<sup>10</sup>

Neo-speakers exhibit a range of proficiency levels, and most of those we interviewed did not exhibit the intricate constructions produced by fluent speakers of Kokama. Preliminary observations of the data show interesting patterns. Neo-speakers are struggling to acquire certain core areas of the grammar, and their overall production differs from both elder speakers and semi-speakers. Here I provide two kinds of patterns found in neo-speakers' speech: lack of bound morphology and a simplified gender system.

Kokama is a relatively isolating language. It has about 12 suffixes but none of them is obligatory. Major grammatical categories like person, number, tense, and modality are conveyed by positionally fixed clitics. Six tense clitics encode three degrees of distance into the past and two into the future. There are also seven epistemic modal clitics, which include five second-position clitics to encode speaker-oriented modality and two phrasal verb enclitics to encode event-oriented modality. Although none of these morphemes (suffixes or clitics) are obligatory, in spontaneous language use as many as four can occur together on a single verb, followed by up to two clitics. This is illustrated in examples (1)-(2), which come from narratives from elder speakers.<sup>11</sup>

- (1) *uruputini upaka-ta-ka-pa=ura=mía*  
 condor wake.up-CAU-INC-CPL=3SG.M=HYP  
 'The condor may wake him up completely'

- (2) *tapia=tua=nu umanu-ta=y=tsuri, y=ikua mukuika-n=chasu=nu*  
 savage=AUG=PL.F die-CAU=3SG.F=PAS3 3SG.F=why two-NMLZ=AFE=PL.F

*yapana*

run

'The savages killed him, that's why the other poor two men ran'

<sup>10</sup> Structural changes and processes of simplification have been investigated in many different languages undergoing death (see, for instance, Campbell & Muntzel 1989; Dorian 1973, 1978, 1981; Mithun 1984, 1989).

<sup>11</sup> Abbreviations: ABL=ablative, AFE=affective, AUG=augmentative, CAU=causative, CPL=completive, DEM=demonstrative, FOC=focus, HYP=hypothetical modality, INE=inessive, LOC=locative, MID=middle, NMLZ=nominalizer, PAS=past, PL.F= plural female speaker, PL.M=plural male speaker, PROG=progressive, RSN=reason, UNC=VBLZ=verbalizer. 1SG.F = first person singular female speaker, 1SG.M=first person singular male speaker, 3SG.M = third person singular male speaker, 3SG.F=third person singular female speaker.

In contrast to elder fluent speakers, neo-speakers do not use any bound morphology. Below is an extract from a heritage speaker that has achieved one of the highest levels in conversational proficiency in Kokama. Note that in this particular extract, the speaker does not make use of suffixes and clitics, except for one locative postposition in (3a) and (3d).

(3) a. *Yatsi-tsui ta kakāri uri nauta-ka, tīma emete kuriki*  
 month-ABL 1SG.M live come Nauta-LOC NEG exist money  
 ‘I came to live to Nauta a month ago, (but) there is not money’

b. *upi kuriki purepeta panara, ipira*  
 all money buy banana fish  
 ‘All is money, (to) buy banana, fish

c. *tīma emete kuriki, tīma eyu [...]*  
 NEG exist money, NEG eat  
 ‘(if) there is no money, don’t eat.

d. *tīma emete kuriki, tīma kakāri nauta-ka*  
 NEG exist buy NEG live Nauta-LOC  
 ‘(If) there is not money, don’t live in Nauta’

Given that no suffix or clitic is obligatory, the structures in (3) would not necessarily be considered ungrammatical, except perhaps for the utterance (3b). Beyond bound morphology, note that neo-speakers show no signs of clause combining strategies. To link simple clauses, they make vast use of prosody: sustained pitch contour at the end of a clause that is linked to the subsequent one, and falling pitch to end a unit (3c, d). While fluent speakers make frequent use of complementation, clause-nominalization and adverbial subordination strategies to combine simple clauses into larger units, heritage speakers leave clausal relationships to be inferred from context. For instance, (3b) calls for a particular purpose clause construction, and (3c, d) call for conditional clause constructions. Although fluent speakers assert they understand the propositional meaning conveyed by the above expressions, when I asked one of them for the meaning of (3c), she provided (4):

(4) *temente kuriki-ra, tīma=maka ene amatsika eyu=tsu*  
 there.is.not money-COND NEG=MOD 2SG can eat=FUT  
 ‘If there is no money, perhaps you won’t be able to eat’

Note in (4) that this fluent speaker basically transformed example (3c) by adding several pieces of grammar, including: the negative existential *temente*, the conditional subordinator *-ra*, the uncertainty modal *-maka*, and the future *-tsu*. Neo-speakers’ use of sequences of simple clauses reveals that perhaps they have not yet acquired complex constructions.

Another example of the patterns found among neo-speakers comes from the gender system. The existence of grammatical paradigms that differ depending on the speaker’s gender is one of the most interesting typological features of Kokama. That is, the language

distinguishes the gender of ‘who is speaking,’ and not the gender of ‘who one is speaking to,’ or ‘what one is talking about.’ This includes number particles, personal pronouns, demonstratives and connectors. For a sentence such as ‘She saw these armadillos in our farm’, a man would say (5), whereas a woman would say (6). The relevant pieces are in bold.

(5) *uri*      *tsumi*      *ikian*      *tatu=kana=uy*      *tana*      *ku=kuara*  
 3SG.L.M    see      DEM.M    armadillo=PL.M=PAS1    1PL.M    farm=INE

(6) *ay*      *tsumi*      *ajan*      *tatu=nu=uy*      *penu*      *ku=kuara*  
 3SG.L.F    see      DEM.F    armadillo=PL.F=PAS1    1PL.F    farm=INE

The male/female distinction is a highly salient feature in Kokama discourse, and speakers are very aware of it. From the elder fluent speaker’s perspective, the use of the correct set of pronouns signals whether someone is a good speaker or not; it also indicates whether someone is in the process of learning the language. Community members often laugh at people—especially men—if he/she uses the incorrect set. This does not mean that men never use the female speech forms. In fact, spontaneous speech is full of direct quotations, and if speakers are quoting someone from the opposite gender, they have to switch to the appropriate set of forms. The following fragment from a traditional story illustrates this. The speaker telling the narrative is a man, who at a certain point quotes a female character. During the narrative the speaker consistently uses male speech forms (a–c), but when he quotes the female tiger, he switches to female speech forms (d), including the demonstrative *ajan* (instead of *ikian*) and the first person pronoun *etse* (instead of *ta*). Immediately after the quote, the speaker goes back to the male form (*ra* ‘third singular’) (e). The relevant pieces are in bold.

(7) a. *rana*                      *chikari=ura,*                      *ikian*      *mima=kana* [...]   
       3PL.M                      look.for=3SG.M                      this.M    puppy=PL.M  
       ‘They look for her, these puppies [tigers]

b. *temente*                      *rana*      *mama*      *aytsemeka*  
       no.exist                      3PL.M    mother    truth  
       ‘Their mother is not there, for real’

c. *ooohhh,*                      *mamaaa*  
       ‘mamaaa’ [the puppies call]

d. *toooj,*                      *ajan=ka*      *etse*                      *ipia-ra-ri,*  
       [ideophone]    this.F=LOC    1SG.L.F                      firewood-VZR-PROG  
       ‘Here I am collecting firewood’

e. na ra kumitsa  
 QT 3SG.M say  
 ‘That’s what she says’

Because we documented language in interaction, we learned that switching back and forth between male and female forms is the norm among fluent speakers, in both narratives and spontaneous conversations. Heritage speakers, on the other hand, not only struggle to master their respective paradigm of forms, but also avoid the switching strategy altogether. They try to stick to either female or male forms throughout (as seen in (4)). Occasionally they would use the wrong form, although this has been observed mostly among women. For instance, there is a tendency to use the male demonstrative *ikian* ‘this’ as the default form. It is clear, then, that the speech of neo-speakers displays unique patterns that differ from those found among fluent speakers.

To conclude this section, I would like to highlight three things. First, my team partners played a key role in designing strategies to identify the different types of speakers. Recall that Victor Yuyarima and Rosa Amías implemented intuitive, but consistent, criteria to categorize speakers into fluent speakers, semi-speakers, neo-speakers and ghost speakers. It is very likely that a linguist working alone, or one that did not have the opportunity to build long-term relationships with the community, could have overlooked or failed to notice such variation.

Second, recording such variation for both revitalization activities and linguistic research is essential. For revitalization purposes, it is important to know the capabilities of the human resources available. In the Kokama context, it would be necessary to consider not only whether the teachers have the desire and training to teach the language, but also what their actual communicative competencies are for implementing specific instructional programs. The current concern is that the neo-speakers have already taken the role of teachers in elementary schools. If teachers are not proficient enough to instruct beyond simple constructions, children will be exposed to ‘simplified’ language. In the event an intervention to improve their language skills in Kokama is contemplated, having available information about their speech patterns could help in the design of a planned response. Beyond an understanding of the neo-speakers’ speech, the significance of the Kokama documentation project lies in having available naturalistic data that learners can access for exposure to the language. Given that it is difficult to find fluent speakers in many areas within the Kokama territory, the audio and video recordings generated within this project can help to recreate the sounds and patterns of the language as they are used in natural interactions.

As for linguistic research, data from nonfluent speakers can make crucial contributions to the discussion about language change and language attrition in endangered language situations. Themes of current debate include whether language decay and contact-induced change show similarities (see, for instance, Sasse 1992), whether dying languages show unequivocal reduction of one kind or another (Dorian 1978; Rice, Libben & Derwing 2002), and whether these processes show parallels with creole genesis (Trudgill 1978), among other topics. Data from Kokama can contribute to the examination of potential relationships between language decay and incomplete second language acquisition. In sum, by focusing our documentary efforts on fluent speakers alone, we would be missing opportunities to contribute not only to revitalization efforts, but also to linguistic theory.

**4. DISCUSSION AND CONCLUSIONS.** This paper highlights the significance of community involvement in documentation endeavors. Across the social sciences, responsibility to the communities with which the researcher is engaged has become a central theme. It is important to keep in mind, though, that given the pressures from academia, “community-based research as a core piece of documentary linguistics, will be valued not only for how it serves society in general, but also for how it serves the profession” (Rice 2011:203). There are already a number of projects that have managed to integrate documentary and revitalization work with descriptive and theoretical linguistic research. Czaykowska-Higgins (2009), Penfield et al. (2008), and Yamada (2011), among many others, all describe community-oriented projects that include revitalization components in cooperation with speech community members. The project described here adds to those types of initiatives. I have shown here that collaborating with community members produces more broadly useful outcomes for linguistic studies and for community initiatives. The activities presented here entail building long-term relationships with the speech community and recognizing that each participant contributes unique knowledge and expertise.

A second point highlighted here is the relevance of documenting language interaction for language revitalization and linguistic research. The involvement of community members throughout the project has been crucial not only for the naturalness and variety within the collected data, but also for documenting interactional patterns essential for revitalization initiatives. Documenting social practices is crucial because language maintenance is about documenting and promoting not only the use of a linguistic code, but also interactional patterns, communicative practices, and sociocultural perspectives. Before the community members took a hand in selecting the speakers and carrying out the actual collection of the data, the collected material was quite different in nature. When I conducted the interviews during the first trips, the language use was not only prosodically more plain, but the use of bare morphological units was the dominating pattern. The data collected by the team members included more vivid language, rich in metaphorical figures and prosodic contours. In terms of morphology, words attached with multiple suffixes and clitics formed the dominant pattern. In sum, when native speakers conducted the interviews and ran elicitation sessions themselves, we collected the most ‘natural’ data.

Quality data makes quality research possible. Within this project, my documentary work has fed my linguistic research. Documenting interactional patterns is essential for examining structures beyond the sentence level. As Rice (2011) notes, collaborative research can allow insights into a language that do not always emerge otherwise, no matter how experienced the fieldworker is. Within the Kokama documentation project, community-based research has ultimately led to better linguistic work, and to new linguistic questions. Concrete outcomes of the project consist of a collection of digital materials, a lexical database and a trilingual dictionary (Vallejos & Amfás in progress). Dictionaries and text corpora are essential for scholarship; they constitute a rich encyclopedia of Kokama knowledge and cultural history. Other products of the project include a comprehensive grammar and a number of articles and conference presentations. These products will contribute to advancing our understanding of the linguistic and historical complexity in this area of the Amazon.

As noted before, after some initial resistance, we decided to go beyond the documentation of fluent speakers to record the speech of semi-speakers and neo-speakers. This



data has allowed us to make preliminary observations that could have an impact on revitalization initiatives. Data from neo-speakers show that although they exhibit a range of proficiency levels, the majority have not yet acquired core areas of the Kokama grammar, such as bound morphology, the gender system, and complex constructions, among others. However, given the social pressures, the fact that a number of young Kokamas have gained some command of the language cannot be disregarded. The current concern is that those second language speakers have already taken the role of teachers in elementary schools. The relevant questions are, what parts of language should take priority in language preservation programs? What kind of variation can be tolerated in small, endangered language communities? Is (the lack of) tolerance related to specific types of speakers? Within the Kokama context, variation introduced by neo-speakers seems generally tolerated. This is not the case with semi-speakers. Of course, these issues would need further research in light of theories of language attrition, language obsolescence and language change in communities in which the endangered language is no longer used on a daily basis. All in all, it has become clear that any effort to revitalize the language must seek to increase the community engagement—e.g., include elders more effectively in the process—in order to shift the focus from formal settings to real language use outside the classroom. And those tasks lie well beyond the control of committed linguists and motivated leaders.

We cannot ignore, however, that documentary linguistics has its caveats. Engaging oneself in this type of research for graduate work presupposes the assumption that a graduate program may take longer to complete than if one carries out purely descriptive, theoretical work. For instance, processing video to create archival resources is extremely time consuming; working in teams usually entails allocating time to train the team members, committing to the production of language resources and archiving materials. This means that our work in the community does not finish after the data has been collected. This is especially true when working with highly endangered languages such as Kokama. Also, working in teams does not always go smoothly. This can be an issue when the members of the team have different approaches or priorities regarding basic issues, such as time management and following an itinerary. During the visits to the villages, community members had the tendency to engage in activities that were beyond the ones that we had planned. Following an itinerary could not be as crucial as helping relatives cultivate farms or build houses, or as participating in community celebrations. Thus, we had to constantly negotiate with these issues, which forced us to re-schedule our trips multiple times.

To conclude, it might fair to say that one of the major achievements of the Kokama revitalization initiatives implemented by several institutions so far has been at the attitudinal level. Community members are not only playing an active role in those activities, but most importantly they are engaging other people and generating excitement about their language. These small, enduring efforts are now evolving into a social movement to connect people to their roots and to stimulate self-recognition of their identity. During the last two decades, the Kokama language has become a key element in a movement among young adults to reclaim their Kokama indigenous identity.

## REFERENCES

- Austin, Peter K. 2010. Current issues in language documentation. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 7, 12–33. London: SOAS.
- Cameron, Deborah. 1998. Problems of empowerment in linguistic research. *Cahiers de l'ILSL* 10. 23–38.
- Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M. B. H. Rampton & Kay Richardson. 1992. *Researching language: Issues of power and method*. New York, NY: Routledge.
- Campbell, Lyle & Martha C. Muntzel. 1989. The structural consequences of language death. In Nancy C. Dorian (ed.), *Investigating obsolescence: Studies in language contractions and death*, 181–196. Cambridge: Cambridge University Press.
- Czaykowska-Higgins, Ewa. 2009. Research models, community engagement, and linguistic fieldwork: Reflections on working within Canadian indigenous communities. *Language Documentation & Conservation* 3(1). 15–50. <http://hdl.handle.net/10125/4423>.
- Dobrin, Lise. 2008. From linguistic elicitation to eliciting the linguist: Lessons in community empowerment from Melanesia. *Language* 84(2). 300–324.
- Dorian, Nancy C. 1973. Grammatical change in a dying dialect. *Language* 49(2). 413–438.
- Dorian, Nancy C. 1978. The fate of morphological complexity in language death: Evidence from East Sutherland Gaelic. *Language* 54(3). 590–609.
- Dorian, Nancy C. 1980. Language shift in community and individual: The phenomenon of the laggard semi-speaker. *International Journal of the Sociology of Language* 25. 85–94.
- Dorian, Nancy C. 1981. *Language Death: The life cycles of a Scottish Gaelic dialect*. Philadelphia, PA: University of Pennsylvania Press.
- Dorian, Nancy C. 1987. The value of language-maintenance efforts which are unlikely to succeed. *International Journal of the Sociology of Language* 68. 57–67.
- Dwyer, Arianne M. 2006. Ethics and practicalities of cooperative fieldwork. In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of Language Documentation*, 31–66. New York, NY: Mouton de Gruyter.
- FORMABIAP (Programa de Formación de Maestros Bilingües de la Amazonía Peruana). 2003. *Uwaritata: Los Kukama-Kukamiria y su bosque*. Iquitos, Peru: Instituto Superior Pedagógico Loreto & Asociación Interétnica de la Selva Peruana.
- Frawley, William, Kenneth C. Hill & Pamela Munro (eds.). 2002. *Making dictionaries: Preserving indigenous languages of the Americas*. Berkeley, CA: University of California Press.
- Grinevald, Colette. 2003. Speakers and documentation of endangered languages. In Peter K. Austin (ed.), *Language Documentation and Description*, vol.1, 52–72. London: SOAS.
- Grinevald, Colette & Michel Bert. 2011. Speakers and communities. In Peter Austin & Julia Sallabank (eds.), *The Cambridge handbook of endangered languages*, 45–65. Cambridge: Cambridge University Press.
- Himmelmann, Nikolaus P. 2006. Language documentation: What is it and what is it good for? In Jost Gippert, Nikolaus P. Himmelmann & Ulrike Mosel (eds.), *Essentials of language documentation*, 1–30. New York, NY: Mouton de Gruyter.

- Instituto Peruano del Bien Común. 2010. Sistema de Información sobre Comunidades Nativas de la Amazonía peruana (SICNA). [http://www.ibcperu.org/?page\\_id=158](http://www.ibcperu.org/?page_id=158). (12 June, 2010.)
- Leonard, Wesley & Erin Haynes. 2010. Making “collaboration” collaborative: An examination of perspectives that frame field research. *Language Documentation & Conservation* 4. 268–293. <http://hdl.handle.net/10125/5077>.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60. 847–894.
- Mithun, Marianne. 1989. The incipient obsolescence of polysynthesis: Cayuga in Ontario and Oklahoma. In Nancy C. Dorian (ed.), *Investigating obsolescence: Studies in language contractions and death*, 243–258. Cambridge: Cambridge University Press.
- Mosel, Ulrike. 2004. Dictionary making in endangered speech communities. In Peter K. Austin (ed.), *Language Documentation and Description*, vol. 2, 39–54. London: SOAS.
- Moseley, Christopher (ed.). 2010. Atlas of the World’s Languages in Danger, 3rd edn. Paris: UNESCO Publishing. <http://www.unesco.org/culture/en/endangeredlanguages/atlas>. (13 April, 2013.)
- Penfield, Susan D., Angelina Serratos, Benjamin V. Tucker, Amelia Flores, Gilford Harper, Johnny Hill Jr. & Nora Vasquez. 2008. Community collaborations: Best practices for North American indigenous language documentation. *International Journal of the Sociology of Language* 191. 187–202.
- Rice, Keren. 2009. Must there be two solitudes? Language activists and linguists working together. In Jon Reyhner & Louise Lockhard (eds.), *Indigenous language revitalization: Encouragement, guidance, and lessons learned*, 30–52. Flagstaff, AZ: Northern Arizona University.
- Rice, Keren. 2011. Documentary linguistics and community relations. *Language Documentation & Conservation* 5. 187–207. <http://hdl.handle.net/10125/4498>.
- Rice, Sally, Gary Libben & Bruce Derwing. 2002. Morphological representation in an endangered, polysynthetic language. *Brain and Language* 81. 473–486. doi:10.1006/brln.2001.2540.
- Sasse, Hans-Jürgen. 1992. Language decay and contact-induced change: Similarities and differences. In Matthias Brenzinger (ed.), *Language Death*, 59–80. New York: Mouton de Gruyter.
- Trapnell, Lucy & Eloy Neira. 2009. *De la cutipa a la icara: la experiencia de Formabiap 1998–2006*. Tarea: Lima, Peru.
- Trudgill, Peter (ed.). 1978. *Sociolinguistic patterns in British English*. London: Arnold.
- Vallejos, Rosa. 2009. The focus function (s) of= pura in Kokama-Kokamilla discourse. *International Journal of American Linguistics* 75(3). 399–432.
- Vallejos, Rosa. 2010a. Is there a ditransitive construction in Kokama-Kokamilla? *Studies in Language* 34(1). 75–107.
- Vallejos, Rosa. 2010b. *A Grammar of Kokama-Kokamilla*. Eugene, OR: University of Oregon dissertation.
- Vallejos, Rosa. 2014. Reference constraints and information-structure management in Kokama purpose clauses: A typological novelty? *International Journal of American Linguistics*. 80(1), 39–67.
- Vallejos, Rosa & Rosa Amías. In progress. *A trilingual dictionary for the Kukama-Kukamiria*. (Working title).

- Yamada, Racquel-María. 2007. Collaborative linguistic fieldwork: Practical application of the empowerment model. *Language Documentation & Conservation* 1(2), 257–282. <http://hdl.handle.net/10125/1717>.
- Yamada, Racquel-María. 2011. Integrating documentation and formal teaching of Kari'nja: Documentary materials as pedagogical materials. *Language Documentation & Conservation* 5. 1–30. <http://hdl.handle.net/10125/4486>.

Rosa Vallejos  
rvallejos@unm.edu