Language Documentation in the Americas

Bruna Franchetto
Federal University of Rio de Janeiro

Keren Rice
University of Toronto

In the last decades, the documentation of endangered languages has advanced greatly in the Americas. In this paper we survey the role that international funding programs have played in advancing documentation in this part of the world, with a particular focus on the growth of documentation in Brazil, and we examine some of the major opportunities and challenges involved in documentation in the Americas, focusing on participatory research models.

0. INTRODUCTION. In the last ten or so years the documentation of endangered languages has advanced greatly in the Americas. The papers in this special section of Language Documentation and Conservation grow out of a session on language documentation in the Americas held at the June 2013 DoBeS Conference "Language Documentation: Past– Present– Future." This was one of a series of organized sessions on language documentation in various regions of the world. The organizers of each session were asked to focus on what was special about documentation in that part of the world. As the organizers of the session on the Americas, we decided to focus on community engagement, with this being a close to essential part of language documentation in the Americas.

The large international documentation programs – DoBeS (Documenting Endangered Languages, funded by the Volkswagen Foundation), ELDP (Hans Rausing Endangered Languages Documentation Project, funded by ARCADIA), and DEL (Documenting Endangered Languages, funded by the National Science Foundation and the National Endowment for the Humanities in the United States) - have played a crucial role in the development and expansion of language documentation in the Americas. We begin the paper with an overview of the impact that these international funding programs have had in terms of number of projects undertaken. We then examine the impact that these programs have had on the development of national funding programs by providing a more in-depth picture of the growth of language documentation in Brazil. Finally, we ask what the major opportunities and challenges are in documentation in the Americas, focusing in particular on participatory or community-based models of research.

1. DOCUMENTATION PROJECTS AND FUNDING AGENCIES. In this section we provide an overview of the documentation projects funded by DoBeS, ELDP, DEL, and the Endangered Language Fund (ELF) as well as the development of language documentation in Brazil.

1.1 DOBES-FUNDED DOCUMENTATION. DoBeS has been a very important contributor to the funding of documentation projects in the Americas, with the first grants awarded in 2002. The last projects were funded in 2012, and work on projects funded in the last round is still in progress.

   The maps below, taken from the DoBeS website (http://dobes.mpi.nl/), show the locations of the DoBeS projects in the Americas.
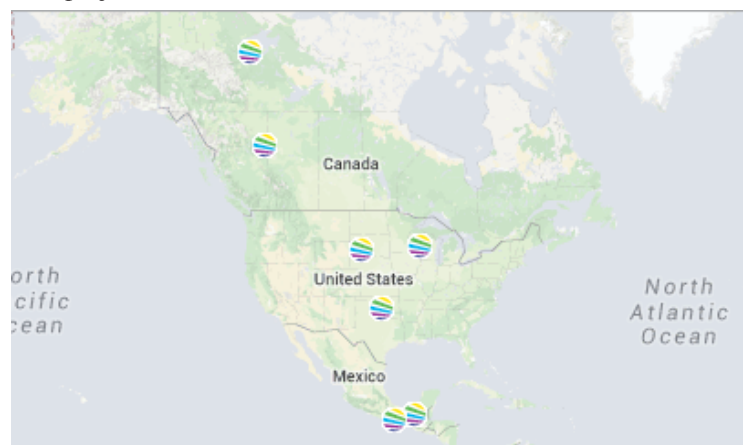
There have been five documentation projects on languages of North America (from north to south, and west to east, these are as follows: Déline (Athabaskan), Beaver (Athabaskan), Lakota (Siouan), Hoocąk (Siouan), and Wichita (Caddoan)). Two projects were funded on languages of Central America (Mexico in particular): Chontal (unclassified), and Lacandón (Mayan). In South America, fourteen projects were funded, with documentation of 24 languages resulting. The languages include Tsafiki (Barbacoan), People of the Centre (five languages of the Northwest Amazon in southeastern Colombia and the north of Peru:

Bora (Boran), Ocaina (Witotoan), Witoto (Witotoan), Resígaro (Arawakan), Nonuya (Witotoan)); Cashinahua (Panoan); Movima (unclassified); Baure (Arawakan); Aikanā/Kwazá (Tupian); Kuikuro (Carib); Yurakaré (isolate, Bolivia); Chipaya (Uru-Chipaya); Aweti and Aché (Tupi-Guarani); and Chaco languages in Argentina (Mocoví (Guaycuruan), Tapiete (Tupi-Guarani), Vilela (Lule-Vilela); Wichí (Mataco-Mataguayuan)) as well as Trumai (isolate; not on the map). The documentation not only involved a large number of languages, it also covered languages of three major linguistic areas of South America. The People of the Center languages project examined the languages of five of seven ethnolinguistic groups in Southeastern Colombia and Northern Peru. Three of the languages spoken in the Upper Xingu area of Brazil (Aweti, Kuikuro, Trumai) were documented in different projects, and four languages of the Chaco region were documented as well.

1.2 ENDANGERED LANGUAGES DOCUMENTATION PROJECT. The Hans Rausing Endangered Languages Documentation Project has also provided strong support for documentation of languages of the Americas. The number of grants awarded between 2002 and 2012 is summarized below.
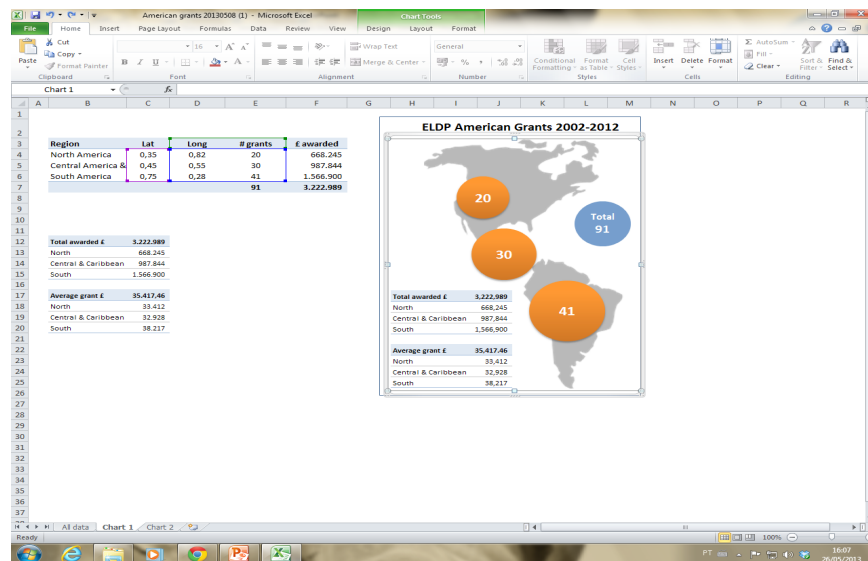
As this figure shows, ELDP funded documentation of a total of 91 languages of the Americas between 2002 and 2012.

1.3 DOCUMENTING ENDANGERED LANGUAGES, NATIONAL SCIENCE FOUNDATION/NATIONAL ENDOWMENT FOR THE HUMANITIES. The Documenting Endangered Languages (DEL) program also funded documentation projects on many languages of the Americas between its inception in 2005 and 2012. The numbers given here are approximate: North America, around 45 languages; Central America, around 23 languages; South America, around 21 languages. In addition, the DEL program provides funds for

BCS: DOCUMENTING ENDANGERED LANGUAGES
Projects Locations Worldwide

archives, conferences, and videos. This program is different from DoBeS and ELDP in that it only funds people in the United States and American citizens elsewhere.

1.4 ENDANGERED LANGUAGE FUND. The smaller Endangered Language Fund was begun in 1996 to provided small grants to support the preservation and documentation of endangered languages around the world. This fund has supported many languages of the Americas, funding approximately 60 projects in North America, around 9 in Central America, and approximately 21 in South America between 1996 and 2012.

1.5 SUMMARY. The major funding initiatives of the past decades have allowed for documentation of languages in the Americas that is probably unparalleled. While numbers only tell a piece of the story, it is worthwhile to consider the numbers of projects that have been funded through the major programs (DoBeS, ELDP, DEL): a total of around 70 in South America, 55 in Central America, and 78 in North America. (This does not count number of languages, but rather number of projects.) In the next section, we discuss the kind of impact these projects have had in one country, Brazil.

1.6 BRAZIL: A SKETCH OF DOCUMENTATION. The various programs discussed so far provide support for work on endangered languages around the world. Partly as a response to this international funding, some countries have introduced support for documentation of languages of that country. We focus here on support for the documentation of languages of Brazil, provided through the government and other institutions within the country.

We begin with a brief sketch of the linguistic situation in Brazil. There are approximately 150 distinct indigenous languages spoken in the country, each with dialects. The degree of language endangerment is high, with 21% of the languages considered in immediate danger because of small numbers of speakers (the median number of speakers is around 250) coupled with low transmission. At the same time, the documentation picture is not a positive one: 13% of the languages are considered to have complete descriptions,

with a grammar, texts, and a dictionary, 38% have what might be called advanced description (e.g., a doctoral thesis), 29% have 'incipient' description (e.g., a Masters thesis), and 19% have no scientific description (Moore and Galucio 2013).

The important work that began with the international programs, DoBeS, ELDP, DEL, and others, allowed for documentation of several languages of Brazil. In addition, it had a perhaps unexpected side effect: these projects served to raise the level of awareness in Brazil about the importance of language documentation and helped to make documentation of endangered languages more prestigious, with the development of structures to encourage descriptive, theoretical, and diachronic studies of endangered languages, based on large amounts of data.

One direct outcome of the South American DoBeS projects was the creation of a network of four archives in three countries, using the Language Archiving Technology (LAT) developed at the Max Planck Institute for Psycholinguistics in Nijmegen, Netherlands. The countries and institutions involved are summarized below.

> • Argentina: Buenos Aires, Laboratório de Documentación y Investigación em Linguistica y Antropologia (DILA)-CONICET, Ministério de Ciencia y Tecnología (www.caicyt.gov.ar/DILA).
> • Brazil: Rio de Janeiro, Museu do Indio-FUNAI, Program for the Documentation of Indigenous Languages (ProDoclin, http://prodoclin.museudoindio.gov.br/), and Belém, Museu Paraense Emílio Goeldi, Ministério de Ciência, Tecnología e Inovação (www.museu-goeldi.br/linguistica).
> • Peru: Iquitos, Instituto de Investigaciones para la Amazonia Peruana-IIAP (www.iiap.org.pe).

This regional network began through the cooperation of the South American DoBeS projects. It established a set of goals involving archiving, access, research focus, and maintenance/revitalization of languages, as follows.

> • to guarantee the availability of the resources from national and international researchers and institutions hosted in the archives to future generations
> • to bolster data exchange among the archives, creating mechanisms to facilitate mutual access
> • to do typological and comparative research in order to address contact and genealogical relations among South American languages
> • to support South American native people in their struggle for language and culture maintenance and revitalization

We return to the final point below.

The agreement for establishing this initiative was signed in 2007 in Rio de Janeiro. Present at the signing, in addition to representatives from the Museu do Índio-FUNAI, the National Museum (Federal University of Rio de Janeiro), the Museu Goeldi, and the Argentinian DILA, were international representatives, including representatives from the VolkswagenStiftung and, as a signatory, representatives from the Max Planck Institute for Psycholinguistics to support the LAT technology.

In 2009, the Brazilian Program for the Documentation of Indigenous Languages (ProDoclin) began at the Museu do Índio in Rio de Janeiro, with financial support from the Brazilian government through the administration of UNESCO (http://prodoclin.museudoindio.gov.br/). ProDoclin is one of the three components of a broadly encompassing program, the Program for the Documentation of Indigenous Languages and Cultures (PROGDOC). The other two programs are ProDocult (Documentation of Indigenous Cultures) and Prodocson (Documentation of Indigenous Sonorities/Music).

ProDoclin was created to promote the documentation of indigenous languages and cultures in Brazil, with attention to the sustainability of languages and cultures. In addition, the program was designed to foster the training of indigenous and non-indigenous researchers in Brazil. ProDoclin supports scholarship, fieldwork, equipment, publication, and training courses. Each project includes up to three indigenous researchers, not just consultants; this is one of ProDoclin's innovative aspects.

The ProDoclin projects have thus far led to documentation of 13 languages. Documentation has been carried out in 54 communities and 25 indigenous researchers have received training in language documentation.

While ProDocult does not have language documentation as its major goal, its 23 projects feed a digital archive with a huge amount of linguistic material, targeting 53 communities of 24 indigenous groups and languages, with 45 indigenous researchers being trained in video-documentation and anthropological research methodologies.

One task of PROGDOC is the creation of audio/video digital archives. This is an important aspect of building national capacity. Without the infrastructure for such archives, the materials might eventually be lost, be difficult to access, or be housed at an archive outside of the country. The materials in the archives come from the projects funded by ProDoclin, and also from materials given by individual researchers whose research has been funded in other ways.

The map below shows the communities where funding from ProDoclin and ProDocult has been important.
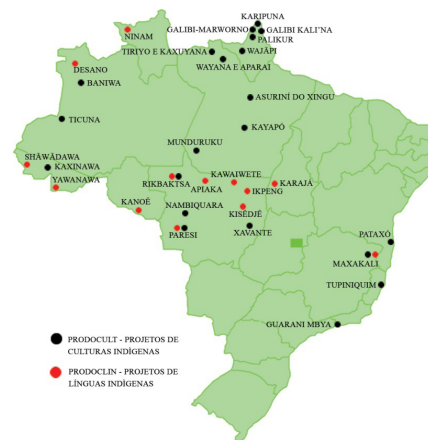
In 2013, five pilot ProDoclin projects began a new type of activity, the development of grammars to be used by indigenous researchers, speakers, and teachers in bi- and multi-

lingual contexts. This work is based on the outcomes of language documentation projects. The language areas involved include Karaja (Macro-Ge), Pareci (Arawakan), Ikpeng (Cariban), Kawaiwete (Tupi), and Wapichana (Arawakan). In the same year, a new branch of PROGDOC was created: Prodocson (Documentation of Indigenous Sonorities/Music), with five projects, focusing on Maxakali (Maxakalian), Guarani-Mbya (Tupi-Guarani), Enawene-Nawe (Arawakan), Krahô (Ge), and Baniwa (Arawakan).

The Museu Paraense Emílio Goeldi has also been actively involved in language documentation. This museum has a long tradition of maintaining scientific collections about Amazonia. The museum also supports extensive language documentation and hosts one of the Brazilian digital archives mentioned above. Through the museum, around twenty Amazonian languages have been documented in the last decade, thirteen with the support of international programs (www.museu-goeldi.br/linguistica). The museum archive contains around 80 individual language collections including recordings of whistled and instrumental forms of three Amazonian languages (Zoró, Gavião, Surui).

In addition to ProDoclin, the Brazilian federal government has undertaken a number of social projects in recent years, one of which is important for those interested in language documentation. *The national survey of the languages of Brazil / INDL* began in 2008 with five pilot projects, focusing on Nahukwá-Matipu (Carib), Ayuru (Tupi), Mbya-Guarani and Assurini (Tupi-Guarani), and Yudjá (Tupi). In December 2011 the president of Brazil signed a decree to formally establish the INDL, coordinated by the Non-material Patrimony sector of the Institute of National Institute of Historic and Artistic Patrimony (Instituto do Patrimônio Histórico e Artístico Nacional, IPHAN, Ministry of Culture) and charged it with developing a plan to survey and 'register' all of the languages of Brazil as cultural patrimony. A national survey of languages would address the existing lack of knowledge of the degree of linguistic vitality of various languages and provide the basis for directing preservation and revitalization. In addition, it would inform the development of other linguistic policies, such as education in the mother tongue in indigenous schools. The survey will also offer answers to many key questions about the factors affecting native languages, both positive or negative.

Recent years have seen an increased attention to indigenous languages and cultures of Brazil. An important development in Brazil is the inclusion of indigenous researchers in projects, and we turn to a discussion of this in the context of the Americas in the remainder of this paper.

2. WHAT'S SPECIAL ABOUT THE AMERICAS? As noted at the beginning, the papers that follow this one grow out of a session on documentation of languages of the Americas. As we considered how to organize a session on this topic, several things came to mind. While we use the term 'the Americas', it is important to recognize that in the Americas, as elsewhere in the world, the linguistic situation is far from homogeneous. Throughout the region, there are not only many different languages, but also many different histories, with very different contact patterns. There are communities at the extremes – with those where the ancestral language is no longer spoken at one end and those where the language remains strong, with transmission from one generation to another, at the other. Many communities fall somewhere between these extremes in terms of language use.

Just as language use varies between communities, so too does interest in language documentation and revitalization. In some communities there is little interest in language documentation and revitalization, while in other communities there is tremendous interest. In many ways, then, characterizing what is common about the Americas is very difficult.

However, there was one thing that we identified as likely to be shared across this vast geographic area, namely the need to develop models that fit the local situation, and we selected this topic as the theme of the session on documentation in the Americas. More specifically, we asked the participants in the session to consider community. There are numerous ways in which relationships between community and researchers might be realized, with these ways shifting over time, but the importance of understanding the critical nature of human relationships is at the heart of most documentation research in this part of the world today. Community control, or what is sometimes called OCAP – Ownership, Control, Access, and Possession (http://cahr.uvic.ca/nearbc/documents/2009/FNC-OCAP.pdf) – thus formed a common theme of the papers in the session, as was a focus on language revitalization and reawakening as critical parts of language documentation. Careful consideration of research ethics, the value and effects of research, and the importance of research that meets needs, embedded in the context of human rights, are at the heart of OCAP, and of the papers in this section.

2.1 COMMUNITY-BASED RESEARCH MODELS. In recent years, there has been much attention paid in the language documentation literature to participatory models of research. These go by various names – participatory research, action research, and community-based research, among others. This type of research model recognizes that different ways of doing research are appropriate in different places and at different times, and it does not espouse a single way of doing research. Rather it takes as critical careful thinking about issues such as control and access. Kristine Stenzel provides an overview of this research in her paper, and we refer the reader to that paper for a summary of that work. In the remainder of this section, we first look at Brazil and then summarize the papers included in this section.

2.2 BRAZIL. We begin with a brief discussion about Brazil. Moore and Galucio (2013) reflect on documentation in Brazil, writing that "Country-specific macro factors and trends at work strongly influence the prospects for language documentation and revitalization and are different from those found in other world regions."

As funding for documentation projects in Brazil became available, it became apparent to the researchers involved that such work needed to take place in conjunction with thinking about what language documentation meant for the people whose languages were the object of study. This involved thinking about issues such as human rights, intellectual property rights, and national and local cultures, policies, and laws. Discussion of these issues was often passionate, leading to a type of paradigm for documentation that is sometimes called decolonizing documentation. In this paradigm, a balance is sought between the efforts required to achieve good scientific results and those directed to the multiple and complex demands from the communities. Resulting from this, an emphasis on the principles of autonomy and self-determination is at the core of Brazilian documentation: these projects are collaborative in nature, with a focus on the genuine involvement of the indigenous people and the empowerment of indigenous researchers. Accompanying the

documentation, community control of access to and use of materials is also important. (See Rice 2010 for discussion of similar models in Canada.)

The careful thought that has gone into this bottom-up approach to language documentation has had many positive results. At the same time, the documentation projects that have utilized this paradigm have encountered challenges of various types. On the one hand, there are challenges that are external to the projects themselves. For instance, documentation projects often must deal with complex official bureaucracy. There is frequently suspicion of initiatives supported from abroad, often associated with old colonialist practices; there can likewise be suspicion of government-supported research. Many documentation projects in South America face challenges in working with communities in that communities are often primarily concerned that the researchers meet their needs – the establishment of a practical orthography, the production of teaching materials, the training of local consultants and researchers in documentation. Meeting these needs is often put as a *sine qua non* for the documentation work to take place. Indigenous communities see this collaboration as an implementation of the principle of reciprocity. Outsiders are often accepted if there is a sharing of knowledge and experience, but, more and more, to use the words of some indigenous people: We want to learn how to fish, not just receive the fish.

On the other hand, there are challenges that come from the nature of documentation today. The technical aspects that are a critical part of language documentation today require training, and there is often a lack of the specialized human resources needed for these technical tasks. Technology may be far from user-friendly, but many funders require that grant recipients use their tools, often leading to conflict. The logistics of fieldwork can also be very complex for both insiders and outsiders to a community. Conflicting demands coming from researchers/collectors, archivers and communities can paralyze the activities of a project. Documentation projects can thus easily become a seemingly endless struggle, frustrating expectations of researchers and community alike. At the same time, in many cases, the demands of the indigenous communities determine changes in the direction of documentation and, in addition, changes in the profile and expectations of funding institutions, the emergence of regional and local programs, and a healthy discussion of ethical standards for the relations between communities, collectors and archivists.

The papers outline many of these kinds of challenges.

2.3 THE PAPERS AT THE WORKSHOP. In sections 1.5 and 2.1, we surveyed some of the kinds of projects going on in Brazil and identified some of the opportunities and challenges that arise in participatory models of research. The papers at the workshop examined in some depth some of the opportunities and challenges that people involved in participatory projects in the Americas have encountered. We organized the workshop to include examples of language documentation in South, Central, and North America. Two of these papers, one about work in South America and the second about work in Central America, are included as part of this issue, namely those by Kristine Stenzel and by Emiliana Cruz and Tony Woodbury. The third paper, by Cecil Garvin and Iren Hartmann, is not included, but we summarize it here.

Kristine Stenzel discusses her experiences working with the Kotiria and the Wa'ikhana of the Vaupés basin of northwestern Amazonia. Stenzel examines what she terms "pleasures and pitfalls" of participatory documentation projects. She addresses logistical, tech-

nical, cultural, and philosophical issues, and examines what it means to work 'with', and what the balances mean.

Emiliana Cruz and Tony Woodbury describe their experience of linguist-community collaboration in the Chatino Language Documentation Project in Mexico, examining how the project developed over time, focusing on the changes that took place. As both a speaker and a linguist, Emiliana Cruz's position in the research is particularly interesting, with her being an insider and an outsider to her community.

The final presentation, by Cecil Garvin and Iren Hartmann, addressed an extremely interesting issue, namely how far it is possible to harmonize language documentation and language revitalization. In the initial development of the major language documentation projects described in section 1, documentation was considered to involve research on a language, with a focus on recording a variety of genres of the language. As noted above, in many ways this goal was antithetical to the types of research models that were developing in the Americas, with a focus on OCAP. Cecil Garvin and Iren Hartmann spoke about their work in the Hoocąk community in the state of Wisconsin in the United States. Harmonizing language documentation and revitalization faces challenges, as funding agencies often are interested in documentation but not in revitalization, yet communities tend to be more interested in revitalization. Garvin and Hartmann reviewed the Hoocąk program, and showed how documentation and revitalization can indeed work in harmony, with each feeding the other. More particularly, they outlined a three-phase project. The first phase was funded by DoBeS, with a focus on documentation. The second phase concentrated on revitalization, and the third phase involved documentation and language revitalization both. Garvin and Hartmann asked whether language documentation and language revitalization are two separate animals that are basically incompatible. They outlined some of the perceived challenges, the prime one being that materials that are useful to both researchers and the speech community are almost impossible to produce and neither audience is served properly. They concluded, however, that in fact documentation and revitalization go hand-in-hand. In the second phase of their project, there were outcomes that were important to both documentation (recording of texts; additions to lexical database; standardized orthography) and to revitalization (training of speakers in reading, writing, and teaching; completed textbook; better knowledge of language in use; lesson plans, games, activities; foundations for language teaching and materials development within the community). Garvin and Hartmann concluded that language documentation makes contributions to language revitalization – these include, among others, the development of resources, training and linguistic awareness, corpora of texts and so on, and an element of prestige. At the same time, language revitalization contributes to language documentation; Garvin and Hartmann identify the collection of text types that are hard to get (conversation); new lexical entries; better relationships with the community; more materials of all sorts; and more helpers. For a researcher, they concluded, there are many benefits, the largest being that the researcher gains a deeper understanding of the language.

Taken as a whole, the papers presented at the workshop and those included in this issue provide case studies showing how language documentation has come to be defined in many parts of the Americas. While the nature of documentation varies depending on many factors – the people involved, the attitudes towards the language, the use of the language, the particulars of the time, the nature of the relationships, the history - we feel confident in

saying that the basic notion of the communities playing a central role in defining the questions and in deciding how things progress is typical of the Americas now. We end with a brief summary of some of the things that people involved in language documentation in the Americas have learned. As the speakers at the workshop pointed out, things take time. This means that patience is required, as are champions who believe in the importance of the languages and are able to keep at the work, even when the going is tough. Community-researcher partnerships bring many challenges with them, but also are part of what is leading to new ways of thinking, combined with social activism.

REFERENCES

Garvin, Cecil and Iren Hartmann. 2013. Towards revitalization and language documentation in harmony: the case of Hoocąk. Talk presented at the DoBeS Conference Language Documentation: Past – Present – Future. June 2013.

Moore, Dennis A. and Ana Vilacy Galucio 2013. Perspectives for the documentation of indigenous languages in Brazil. Museu Paraense Emílio Goeldi - Ministry of Science, Technology, and Innovation. Unpublished.

Rice, Keren. 2010. The linguist's responsibility to the community of speakers: community-based research. In Louanna Furbee and Lenore Grenoble (eds.). *Language Documentation: Practices and Values*. Amsterdam/Philadelphia: John Benjamins Publishing Company. 25-36.

Bruna Franchetta
bfranchetto@yahoo.com.br

Keren Rice
rice@chass.utoronto.ca