Reflections on language documentation in the Chaco

Lucía Golluscio
Consejo Nacional de Investigaciones Científicas y Técnicas
Universidad de Buenos Aires

Alejandra Vidal
Consejo Nacional de Investigaciones Científicas y Técnicas
Universidad Nacional de Formosa

This chapter focuses on field research aimed at documenting Chaco languages with varying degrees of vitality, specifically those spoken in Argentina and in the vicinity of the Argentinian/Paraguayan, Argentinian/Bolivian, and Paraguayan/Bolivian borders. The case studies here selected provide an overview of recent experiences conducted in Chaco within the framework of Himmelmann 1998’s foundational program on documentary linguistics and subsequent publications along these lines. We emphasize the results of collaborative research on equal grounds and a discourse-oriented approach to language documentation. Our reflections also highlight the current threatening situation of indigenous peoples and their languages and discuss the function of language documentation, preservation, and archiving in this fragile scenario, with a view to supporting community language use and transmission as well as ongoing and future research in South America.

1. Introduction

In the early 1990s, a relevant publication in Language by Hale et al. (1992), placed the topic of endangered languages around the world on the international academic agenda. Explicitly linked to this topic, the seminal paper by Himmelmann (1998) brings the issue of linguistic documentation to the forefront while advocating for documentary linguistics with an autonomous status similar to descriptive linguistics.

*We are very grateful to the editors of LD&C for their invitation to collaborate in this volume and their recommendations to our first manuscript as well as to two anonymous reviewers for their comments and suggestions. Our special thanks go to the following colleagues for their generous contribution to the article: Elizabeth Birks, Florencia Ciccone, Luca Ciucci, Santiago Durante, Hebe González, Analía Gutiérrez, and Verónica Nercesian.*
In this article, we focus our reflection on contemporary collaborative research projects on Chacoan languages developed within the documentary linguistics framework. After this introduction (§1), in §2 we briefly consider a set of case studies; then, based on the results of those initiatives, we reflect on language documentation in the region and in South America (§3). Finally, in §4 we present the proposal for the creation of the South American Network of Regional Linguistic and Sociocultural Archives, with a view to the present and future work on language documentation and preservation, including the issue of data accessibility and exchange.

Chaco, in the heart of South America, comprises a vast lowland territory ranging from southeast Bolivia and the southwestern area of the Mato Grosso in Brazil northwards, to the westernmost area of Paraguay and northeast of Argentina southwards. It is a plurilingual region where twenty languages belonging to seven linguistic families are spoken with differing degrees of vitality. Moreover, extended multilingualism phenomena have been registered on the Bolivia/Argentina/Paraguay border (Campbell & Grondona 2010; Ciccone 2015).

The panorama among Chaco peoples is particularly complex. Many of them inhabit lands currently belonging to different countries and, therefore, ruled by different socio-educational policies. In Argentina, Intercultural Bilingual Education was until recently a program under the National Education Law enacted in 2006. However, resistance from many sectors, including some of the teachers themselves (Vidal & Kuchenbrandt 2015: 91), and the limited availability of materials for bilingual education mean these peoples are not offered equal opportunities within the educational system and literacy in their languages is not always valued by national or regional governments.

The situation regarding linguistic vitality is also heterogeneous. On the one hand, there are communities where use and transmission of the heritage language is supported by collaborative experiences in documentation and ongoing linguistic description, with community-led linguistic activism (2.2, 2.3, 2.4). On the other, important language attrition and shift processes have been documented (2.1). In particular, our field research has shown that transmission to younger generations does not always occur (2.1, 2.4). Language transmission and use is further affected by migration to the cities, where native language teaching is seldom on the curriculum. Living in rural settings helps strengthen shared socio-cultural ties. However, after moving to the cities, speakers usually become a minority within a Spanish-speaking majority. See the following striking comparative data (INDEC-ECPI, 2004-2005). Whereas less than half of Pilagá (48%) and Wichí (35%) speakers live in urban areas in Argentina, the Toba/Qom proportion of urban population is much higher (69%). This Census reports 99% and 91% of Pilagá and Wichí native speakers, respectively, while the Toba/Qom speaker proportion is 65%. In spite of these figures, a tendency towards incomplete language acquisition has been observed among Pilagá (2.4); see also Tapiete (2.1). Hence, speakers of these languages in peri-urban settlements do not achieve skills in some registers of their heritage language (principally, oratory and narrative).

In Argentina, these processes have been triggered within a broader context defined by a hegemonic Hispanizing language ideology fostered by the national state since the second half of the 19th century. Other factors include speakers’ marginal status; changes in organization, from hunter-gatherer to (semi)urban sedentary lifestyle; migration and

---

2For a complete view of of the distribution and situation of peoples and languages in the area, see Unruh & Kalisch (2003), Censabella (2009), Lewis (2009). Detaileded references on Chaco languages studies can be found in Fabre (1998; 2017a [2005]), Golluscio & Vidal (2009-2010), and Campbell & Grondona (2012).
integration with other indigenous or non-indigenous communities; everyday contact with the Spanish-speaking population and media, and a diminished sense of ethnic pride caused by racism and discrimination. In some cases, governments’ denial of these peoples’ Constitutional rights contributes to the situation. Even for those currently settled in rural enclaves in Argentina, pressure by the spreading agricultural frontier into the forested territories they inhabit is a crucial problem that puts not only their culture and language transmission, but also their very survival at risk, not to mention serious local and global damage caused by deforestation and soy plantations.³

Fortunately, the communities here considered have shown a great deal of interest in language revitalization. The still everyday use of their languages ensures secure format documentation to produce linguistic materials for educational purposes. These are the reasons that have led to undertaking linguistic documentation of most Chaco languages over the last twenty years (see §2).

2. Documentation of Chaco Languages focusing on Argentina and Paraguay

Fieldwork-based linguistic research by professional linguists focusing on Chaco languages began in the 1960s and 1970s, and included grammars, vocabularies, and phonological studies. From a descriptive perspective, they provided an analysis and written documentation of these languages. Subsequently, work on Chaco languages increased, especially through doctoral dissertations. Though relevant to our past and current knowledge of the Chaco, these studies were not necessarily intended for language documentation, neither did they adopt a discourse-oriented approach.

It was not until the early 21st century that practices and research projects within the framework of documentary linguistics on Chaco languages began. The first was the Chaco languages Project (2002-2005), “Endangered Languages, Endangered Peoples in Argentina. Documentation of four Chaco languages in their ethnographic context: Mocoví, Tapiete, Vilela, and Wichí”, carried out as part of the Dokumentation Bedrohter Sprachen (DoBeS) Program, under the auspices of the Volkswagen Foundation (see Golluscio & Hirsch 2006).⁴ Audio and video resources from this project, mostly having open access, have been deposited both in the DoBeS Archive and the Archive of the Laboratorio de Documentación e Investigación en Lingüística y Antropología (DILA) (http://www.caicyt-conicet.gov.ar/dila), created at the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina, by agreement with the Max Planck Institute (MPI) for Psycholinguistics, Nijmegen, The Netherlands, in 2007.

In the last two decades, there has been a significant increase in available audio and video records of Pilagá, Wichí, Chorote, Nivaclé, Mocoví, and Ayoreo language incorporated to the Endangered Languages Archive (ELAR) within the framework of the Endangered Languages Documentation Programme (ELDP) and the Hans Rausing Endangered Languages Project (HRELP). Likewise, Toba/Qom, Maká, and Mocoví materials can be found in the Archive of the Indigenous Languages of Latin America (AILLA), which, in 2009, donated a copy of these resources to DILA-CONICET.

³For more information on the socio-political and sociolinguistic situation in Argentina and Chile, see Zúñiga & Malvestitti, this volume.
⁴This collaborative documentation project settled at the University of Buenos Aires (UBA) was conducted by an interdisciplinary team under Lucia Golluscio’s supervision in academic collaboration with Bernard Comrie (Department of Linguistics, MPI for Evolutionary Anthropology) (http://www.mpi.nl/DOBES/projects/chaco).
Recent experiences, fieldwork techniques and results from some of the Chaco language documentation projects will be reviewed in the following subsections. We selected projects focused on languages belonging to different families with varying degrees of vitality and developed under international programs such as DoBeS, HRELP, the Foundation for Endangered Languages (FEL), and the Documenting Endangered Languages Program (DEL) from the National Science Foundation (NSF). These projects taken together represent a paradigmatic change in the field of Chaco linguistics, given the collaborative perspective and the technological possibilities for disseminating results. For many researchers in Argentinean academia, the focus on language work has shifted to include recording discourse exchanges and verbal art – see Messineo (2008) in this journal, among others – and set the records in formats useful to a number of parties, from communities and linguists to social organizations and national and provincial education ministries.

2.1 Tapiete and Vilela, two endangered languages in the Argentine Chaco  Tapiete (Tupí-Guarani) and Vilela (Lule-Vilela, affiliation under discussion) are the two most endangered languages in the Argentine Chaco. Both show evidence of linguistic attrition without obsolescence. However, their current differing sociolinguistic and sociopolitical situations raise very different opportunities for their future (Gollusco & González 2008). While multigenerational, closely-knit Tapiete communities do exist and have a relatively small number of speakers (about 2,200, Ciccone 2015) distributed around Argentina, Paraguay and Bolivia, the extreme paucity of Vilela speakers (only two speakers, actually remembers, have been located) and the lack of a speech community have proven to be critical threats for this language. Moreover, the interruption of Tapiete language transmission is quite recent (González 2005). In contrast, the current Vilela situation is the result of a lengthy cultural and political de-structuring process intensified in the second half of the twentieth century. The elderly speakers who are working collaboratively on the documentation of their language (ML and GC, siblings, now 85 and 83), though exposed to Vilela on a daily basis during their childhood and youth, do not currently use it in their everyday life (Domínguez et al. 2006).

The existence of a Tapiete speech community enabled researchers to adopt traditional fieldwork techniques with older speakers, but linguistic tasks involved in documenting contemporary ways of speaking among younger speakers proved to be methodologically challenging (Ciccone 2015). Younger bilingual speakers incorporate loanwords, grammatical innovations and extensive code-switching, without participating in the performance of traditional narratives. It was necessary to foster particular situations to elicit spontaneous performances in Tapiete (see Figure 1). Traditional linguistic techniques would not have enabled the recording of verbal exchanges among these younger speakers, given their high regard for expert Tapiete speakers and their recognized abilities as language consultants, which the younger ones feel they cannot match.

As part of documenting traditional knowledge, an ethnobotanic dictionary is underway, in collaboration between Hebe González and members of the Tapiete communities in Salta, Argentina (González 2017 and in press) (see Figure 2). Apart from its contribution to lexical and cultural studies, this work includes a collection of analyzed narratives on Tapiete life in the Chaco.
Figure 1: Tapiete people. Awara Montes, Florencia Ciccone and members of the community. Tartagal, Salta, Argentina.
Figure 2: Helena Cabeza (Tapiete community) holding a specimen of ḡambi 'spicy herb' (Acmella oppositifolia). Photo by Hebe González.
Vilela documentation has, instead, been based on a single-speaker-centered approach, including both the documentation of linguistic attrition in our main consultant’s speech and the systematization of language-remembering processes triggered during his participation in documenting his language. The remembering strategies achieving best results (see Figures 3 and 4) were his return home and joint work with his sister. Narratives about topics relating to their own childhood, when the mother tongue becomes engrained, were key to motivating remembering of the language. This situation brings up an interesting question on the existence of latent cognitive strata that may resurface when stimulated by emotionally-charged experiences. Finally, collaborative linguistic research highlights the essential contribution of these last speakers’ generation to the knowledge of: their heritage language, evidences of contact with other languages in and beyond the area (Golluscio 2015), and the persistence of Vilela structural characteristics and lexicon since the 18th century (Zamponi & Golluscio 2018).

Figure 3: Vilela language documentation. Returning to his place of origin. Mario López with Analía Gutiérrez, team member. Photo by Marcelo Domínguez.
Figure 4: Recovering Vilela basic vocabulary. Art session at Mario López’ house with his grandchildren, great-nephews and María Hellemeyer, team member. Photos by L. Golluscio.
2.2 Ayoreo discourse-oriented documentation The second project involves collaborative documentation of Ayoreo discourse conducted in Campo Loro, Paraguay. Ayoreo (Zamucoan) is still the language of communication in Paraguayan and Bolivian Ayoreo communities, although some signs of linguistic attrition have been documented (Durante p.c.). The 2012 census data for the above-mentioned countries estimate 1,862 Ayoreo people in Bolivia (CEDIB 2012) and 2,481 in Paraguay (DGEEC 2012), with some communities as yet uncontacted in Paraguay. The speech community exhibits speakers of all ages with the elderly and youngsters being mostly monolingual. The current vitality of the language is clear in the narratives collected in close collaboration with the community by Santiago Durante, a former PhD student at the University of Buenos Aires, under the auspices of ELDP (https://elar.soas.ac.uk/Collection/MPI192274). However, the presence of Spanish at school, in the media and in the social networks is jeopardizing the future of the language.

This collaborative research has centered on the documentation of texts of various genres in high-quality audio and video recordings. The collected stories, about life before contact, evangelization and sedentarization, are of great interest, since Campo Loro inhabitants have only recently come into contact with non-indigenous groups. The outcomes of the analyzed and annotated text-corpus include a significant volume of new information on Ayoreo grammatical structures based on naturalistic data and the publication of a collaborative anthology of narratives (Etacore & Durante 2016) already in use in the community. The book was welcomed by community members who interpret it not as a finished product but as a first step in the process of documenting their cultural and linguistic heritage (see Figure 5).

![Documenting Ayoreo](image)

**Figure 5:** Documenting Ayoreo. Benito Etacore and Santiago Durante editing the book *Campo Loro gosode oe ojñane udojo*, Boquerón Department, Paraguay.

---

[4]Our knowledge of the Zamucoan languages (Ayoreo and Chamacoco) owes much to Pier Marco Bertinetto and Lucca Ciucci (Figure 6). See complete references on these authors in Fabre (2017b).
Figure 6: Documenting Chamacoco. Luca Ciucci with Francisco García and Domingo Calonga, Paraguay.
2.3 Nivaclé: a single language and territory, two countries There are an estimated 1,000 Nivaclé speakers in the provinces of Salta and Formosa, Argentina, and 12,000 in Paraguay. Today, there are open access audio resources for the community and more broadly for linguists (https://elar.soas.ac.uk). With support from ELDP and CONICET, Analía Gutiérrez has been investigating dialectal differences in Paraguay and contributing to capacity-building for language transcription and decision-making with regard to competing alphabets (Gutiérrez 2015) (see Figure 7). In Paraguay, this language is the primary means of communication among family members within indigenous communities, but Spanish and to some extent Guaraní are used with outsiders (Fabre 2017b). There are incipient bilingual programs in Paraguayan Nivaclé community schools. In Salta, the community is multilingual, living in peri-urban settings. Nivaclé communities in Formosa have received no attention from the local government and there are no bilingual education programs to serve around 93 school and pre-school-aged children from 180 families. Communities are denied most civil rights and recognition as an indigenous group. There is extended Nivaclé-Spanish bilingualism in Formosa but no extended multilingualism as documented in the Bolivia/Argentina/Paraguay border (see §1). In recent years, a project awarded by NSF to Alejandra Vidal as co-principal researcher (see 2.4) has enabled the production of audio and video resources in Nivaclé spoken in Formosa, Argentina. Transcription, analysis and translation of 5 hours of texts (conversations, narratives, songs) for archiving is currently underway. A selection of Nivaclé narratives was published for community use (Vidal 2015).

Figure 7: Work session at the First Meeting of Nivaclé Teachers, Uj’e Lhavos, Paraguay, Photo by Analía Gutiérrez.
2.4 Collaborative documentation, description and revitalization activities of Wichí and Pilagá  Wichí language, spoken across the borders of Argentina and Bolivia, is transmitted intergenerationally. Characterized by their visibility, with their own radio programs and political organizations, representatives at government levels, some primary and secondary school teachers, the Wichí are one of the most numerous groups in the region (around 40,000 people).

Human resources training was a central issue in the DoBeS Chaco project (see page 295) and continues to be so. Linguistic training of younger Wichí speakers from settlements located by the Bermejo River, collecting oral discourse, developing reading materials, vocabularies and grammar courses for the study of their language, culture and history were activities pursued by Verónica Nercesian (Nercesian 2014a) (see Figure 8). Currently, the study of Wichi/Weenhayek dialectal varieties and sociohistorical processes in Northern Argentina and Southern Bolivia is underway. A published Wichí grammar (Nercesian 2014b) shows significant progress in new lines of research, such as the interaction between phonology, morphology, syntax and semantics. This model has opened up a new perspective on the study of similar phenomena in other Chaco languages.

Pilagá, with its 5,000 speakers, does not enjoy the vitality it did 20 years ago when Vidal’s research on the language began (Vidal 2001). Although intergenerational transmission of the mother tongue still occurs in some rural communities, and language teaching materials such as a talking dictionary and a learners’ grammar have been developed for community schools (funded by ELDP and FEL; see Vidal & Miranda 2010;
Vidal, Almeida & Miranda 2014a–d), there are symptoms of linguistic attrition especially in semi-urban and urban settlements.

A recent documentation project (NSF-DEL 263817) was established between the University of Oregon (with Doris Payne as PI) and the Universidad Nacional de Formosa, Argentina, with the purpose of obtaining high-quality audio and video recordings in Pilagá and Nivaclé spoken in Formosa, and text-collection for archiving (see also 2.3). These include narratives on past conflicts and contact between Pilagá, Nivaclé and Wichí groups (see Figure 9). The texts support very fragmentary data provided by earlier European travelers and ethnographers about encounters with indigenous peoples in the Chaco region and their distribution in the territory where fieldwork is conducted.

![Figure 9: Alejandra Vidal recording Pilagá stories, Km 30, Formosa, Argentina.](image)

3. Directions in Chaco language documentation: Thoughts from the field

The initiatives on language documentation presented in §2 raise a number of issues. Following Himmelmann’s design for a language documentation project, our considerations evolve in three directions: corpus collection, corpus theorization and the role of the participants.

Regarding corpus collection, the fragile situation and degree of endangerment of Chaco languages described above provide a strong warning that, although linguistic analyses may be conducted at a later date, documentation of indigenous languages in Chaco and South America is extremely urgent with much work still pending. On the other hand, documentation of these languages should allow for developing versatile fieldwork methodologies to deal with a range of speakers in different situations and scenarios.

As to corpus theorization, we believe that the sharp distinction between documentation and description proposed in Himmelmann (1998) had a foundational epistemological function and relevance: it was necessary to constitute documentary linguistics as a field of study having the same status as descriptive linguistics. Now, twenty years later, our experience and that of other researchers in South America confirm the intimate feedback relationship between documentation and description. In line with Evans (2008) and
Woodbury (2011), the collaborative preparation of a dictionary and or a grammar can at the same time raise the possibility of eliciting new texts, as shown by the Tapiete (2.1) or Wichí (2.4) experiences. Conversely, the collaborative collection and edition of Ayoreo (2.2) and Nivaĉle (2.3) texts constitute a primary source of information to ongoing studies on the phonology and the grammar of those languages. Likewise, community activities carried out during and after the language documentation projects here considered confirm Himmelmann’s claim about the creation of available multipurpose corpora which can be used for and beyond linguistic research.

Concerning participants, any scientific praxis involving fieldwork with communities cannot and should not avoid reflecting on the impact of the intrusion of field researchers in the lives of community members. This issue becomes relevant in the case of South American indigenous peoples, socioeconomically very vulnerable and historically threatened by non-indigenous society. Faced with this ethical question, the perspective of collaborative fieldwork and the model of empowerment (Cameron et al. 1997) appear as viable alternatives when proposing work agendas agreed on between the researcher and the community, based on common interests.

Furthermore, research on Vilela, the severely endangered Chaco language mentioned above (2.1), highlights the unique role of the last generation of speakers of a language in the documentation, description, and history of their language and people. In addition to the claim by Harrison & Anderson (2008) about the importance of including the speech of semi-speakers and passive speakers in the documentation of endangered languages, we affirm the relevance of applying language-remembering strategies in these situations. As said elsewhere, “the attested processes of linguistic remembering and recovery defy biological metaphors about the vital cycle of languages and their fate” (Golluscio and González 2008: 238).

Finally, the urgent need to document lesser-known still living South American languages, some of them critically endangered, and fill in some remaining gaps in genealogical, typological and areal contact knowledge of the languages makes the establishment of strong collaborative links a central goal. This issue is addressed in the next and final section.

4. Looking ahead: Towards a South American Network of Regional Linguistic and Sociocultural Archives Different academic institutions in South America involved in linguistic and cultural documentation with indigenous peoples in the region have set up a South American Network of Regional Linguistic and Sociocultural Archives, with the foundational aim of strengthening interaction and exchanging information and documentary resources between archivists, researchers and members of indigenous communities. Some objectives of the agreement include: contributing to the knowledge, preservation, valuation, transmission, and diffusion in national societies of South American languages and non-standard varieties of Spanish and Portuguese, as well as migration languages in contact with them; promoting the development of collaborative language documentation as well as typological and areal research projects; maintaining technological compatibility of files, and, ideally, developing shared criteria for the classification of the contents; implementing a unified code of ethics for researchers, donors, file users, community members and responsible archivists; ensuring long-term preservation of databases and increasing security for stored data; working on creating automatic copies distributed among the archives involved, respecting authorship rights,
ethical principles, regulating access restrictions; and establishing validation processes that can be used for modeling other databases.

This network has its origins in the more than decade-long cooperation between members of South American institutions. The exchange has long been active at informal and personal levels, in particular thanks to researchers participating in DoBeS and ELDP projects, as well as the use of similar technology in individual centers (Seifart et al. 2008). The project currently encompasses CONICET, Universidad Nacional de Formosa and Universidad Nacional de San Juan, Argentina; Instituto de Investigaciones para la Amazonía Peruana; Universidad de Chile; Centro de Estudios Antropológicos of the Universidad Católica de Asunción (CEADUC), Paraguay; Pontificia Universidad Católica del Perú, and Universidad del Azuay, Ecuador. The Museu do Índio and the Universidade Federal do Rio de Janeiro through Bruna Franchetto’s active participation, as well as the Museu Paraense Emílio Goeldi, Brazil have shared the Network’s objectives from inception (Drude et al. 2009; Golluscio et al. 2013). As of this year, this initiative has been assigned full legal status by an Agreement for Scientific and Technical Collaboration signed by all Network members. The proposal is open to other South American archives or institutions that wish to join this initiative on the understanding that the Network’s ethical guidelines are respected.
References


Lucía Golluscio
lgollusc@conicet.gov.ar

Alejandra Vidal
vidal.alejandra@conicet.gov.ar