Text translation in the context of
endangered language documentation: The case of Iquito*

Christine Beier
cmbeier@gmail.com
April 23, 2015
Fieldwork Forum, UC Berkeley

1 Introduction

○ This talk\(^1\) explores the realm of translating ‘texts’ of various types from Iquito, a highly endangered language spoken in Peruvian Amazonia, into Spanish and English, in the context of a long-term, community-oriented language documentation project, the Iquito Language Documentation Project (ILDP).

○ I will discuss some key assumptions, objectives, challenges, and strategies relevant to text translation as part of endangered language documentation; and will examine some of their impacts on project planning, work flow, and eventual tangible outcomes, illustrated with specific examples from a text produced by ILDP team members.

○ It is my goal to make a strong case for why the quality of our free translations matters; and to offer some practical strategies for improving our work in this domain – both the ease of doing the work, and the quality of the eventual output.

2 Background: The Iquito Language

○ Iquito [iqu] – or, in the new official orthography, Ikíitu – is a member of the Zaparoan family. It is one of the two Zaparoan languages still spoken; the other is Arabela, with \(~50\) speakers. The other two members of the family are Sápara, a.k.a. Záparo [zro] (nearly extinct); and Katsakáti, a.k.a. Andoa [anb](extinct).

○ Iquito is highly endangered; at present, \(~18\) fluent speakers remain, all over 65 years of age.

○ See the maps below for the approximate location of Iquito, which is primarily spoken in San Antonio de Pintuyacu, Loreto, Peru (about 40 miles straight-line distance from Iquitos); as well as the approximate locations where rememberers of Sápara and Katsakáti now live.

\(^*\) Many thanks to all of the ILDP team members with whom I have worked, and from whom I have learned, in the course of our collaborations over the years. Special thanks to Lev Michael, Nora England, and Tony Woodbury for their input over the years on matters related to the substance of this talk; and to Ema Llona Yareja, Ligia Inuma Inuma, Hermenegildo Díaz Cuyasa, and Jaime Pacaya Inuma for all that they have shared with me in, through, and about the Iquito language. Last but not least, special thanks to Line Mikkelsen for a series of stimulating conversations on this topic during the spring of 2015.

\(^1\) This manuscript is available at: www.cabeceras.org/beier_translation_fforum_20150423.pdf
The segmental inventory of Iquito is relatively typical for Western Amazonia. The ten consonants (j, k, m, n, p, r, s, t, w, y) include no voiced obstruents; and the four vowel system (a, i, i, u) also demonstrates contrastive vowel length (aa, ii, i, uu).

Its prosodic system is a mixed stress-tone system with H and Ø tones (represented thusly: a, á, aa, áa, aá, etc.).

The orthography for Iquito/Ikíitu was changed by the Ministry of Education and made official in September 2014. The official alphabet is: a, i, i, j, k, m, n, p, r, s, t, u, w, y.

Syllable structure in underived contexts is (C)V(V).

Iquito exhibits nominative-accusative alignment and SVO basic constituent order. This basic order is changed through a number of processes, including topicalization and focus fronting.

Iquito verbs are exclusively suffixing; tense and aspect are obligatory categories.

Much Iquito morphology is non-agglutinative, and involved changes in quantity or tone on vowels adjacent to the segments associated with a given morpheme, or consonant mutation in adjacent consonants.

3 Background: The Iquito Language Documentation Project

The ILDP was launched in 2002 as a team-based, community-directed documentation project.

While Lev Michael and I were grad students at the University of Texas at Austin, he and I designed the ILDP, and collaborated with Dr. Nora England on its administration. The project was funded by the Endangered Languages Documentation Programme (ELDP, www.eldp.net) from 2003 through 2006 (Grant #MDP0042).

Since its inception, the ILDP has included a total of 6 fluent speaker consultants; 4 community linguists; and 15 visiting graduate student researchers, including 11 linguists and 1 ethnomusicologist from UT-Austin, and 3 Peruvian licenciatura-level linguists from the Universidad Nacional Mayor de San Marcos in Lima.
○ Fieldwork on the ILDP has been done in 9 years so far: 2002 through 2006; 2008, 2009, 2014, and 2015, primarily during the US academic summers; fieldwork has mostly been carried out by a team, but a handful of solo trips have been made by licenciatura and doctoral students.

○ As an endangered language documentation project, the ILDP is producing an encyclopedic dictionary in both Iquito-to-Spanish and Iquito-to-English forms; a descriptive grammar; a set of Iquito-Spanish educational materials for community use; and a wide-ranging text corpus, to be published in forms useful and accessible to multiple audiences – primary among them, of course, being linguists and community members, but also members of the broader public, especially in Peru where attitudes (and prejudices) regarding the value of indigenous languages has tremendous consequences for the longevity of those languages.

○ Information on and products of the ILDP are available at: www.cabeceras.org/ildp_home.htm

4 Background: The ILDP text corpus

○ The ILDP’s text corpus currently includes ~170 audio-recorded ‘texts’, amounting to ~35 hours of spoken Iquito.

○ When I say ‘texts’ here, I mean audio (and video) recorded spontaneous streams of connected speech (even if the content is memorized); which are typically monologic; and which narrate an event or events, or otherwise relate or exchange ideas and information.

○ So far about 100 of the 170 texts have been partially or fully transcribed.

○ The recordings range from ~1 minute to 75 minutes; and many, but not all, are accompanied by video recordings.

○ At present, the longest fully transcribed and translated text is Kaaya asáana, at 68 minutes and 900 lines.

○ Types of texts in the collection include:

  ◦ traditional/shared narratives
  ◦ historical narratives
  ◦ personal narratives, reminiscences, and dream narratives
  ◦ procedural descriptions
  ◦ instructive descriptions and narratives
  ◦ hortative discourse
  ◦ ayahuasca chants
  ◦ public political speeches
  ◦ multi-party conversations
  ◦ and even some gossip (ergo, generally not for public consumption)

○ We produced a printed collection of 21 texts (2,372 lines) for the community in 2006, which is available online at: www.cabeceras.org/ildp06_textos.html
We plan to publish a substantial portion of the text corpus in two versions at the end of 2016: one for a lay audience with Iquito, Spanish, (and sometimes English) free translations; and another for linguists with Iquito, word- and morpheme-level breakdowns and glosses, and English free translations.

At that time, we also plan to publish Iquito-Spanish and Iquito-English versions of our dictionary.

Taking a brief step out into the bigger picture, let me mention that I consider richly detailed, engaging, and aesthetically pleasing texts and text artifacts to be indispensable tools for both learning a language and learning about a language and culture, as well as for revalorizing them. So a strong value of both content and form in text work runs through this entire discussion for me.

5 The need for a strong foundation in the theory and methods of translation

Within the next decade or two, the Iquito language will have no remaining native speakers. And language revitalization efforts have born little fruit thus far – although the revalorization of the language has been much more successful, both in San Antonio and regionally.

Therefore, for our text corpus to be useful and accessible to the heritage language community, it will have to be accessed through other languages, primarily Spanish. Whether the ILDP’s texts are used as a cultural resource, as a political resource, and/or as a language-learning resource, Spanish will soon become the gateway to Iquito.

Therefore, I feel that the free translations and annotations provided in our published text collection must have ‘good’, if not ‘excellent’, free translations into local Spanish; as well into English when possible (both to broaden the general audience and to increase the prestige of the text artifacts); and must include ample annotation to connect the local Spanish to more standard varieties.

Knowing that translation is a complex and context-sensitive enterprise, I have sought guidance and insight from others who have undertaken careful translation work in the context of language documentation.

However, when I started digging around in the language documentation literature for discussions of the theory and methods of translation relevant to our work, I found astonishingly little.

It is my goal, therefore, to examine the disciplinary context surrounding this notable absence; and then to provide some experience-based commentary and preliminary resources in order to begin to fill that lacuna, both for my own self and for you.

6 Disciplinary context

6.1 Rallying calls for comprehensive documentation

In the programmatic literature on language documentation and documentary linguistics that has emerged since the early 1990s, calls are repeatedly made for “the making and keeping
of records of the world’s languages and their patterns of use” (Woodbury 2003) that are “lasting,” “multipurpose” and “cover all registers” (Himmelmann 2006: 1-2); that “provide a comprehensive record of the linguistic practices characteristic of a given speech community” (Himmelmann 1998); and which are produced through documentation projects in which “native speakers are among the main players in determining the overall targets and outcomes of a documentation project.” (Himmelmann 2006: 16).

- In my personal experience in this field, it is now fairly widely accepted, at least at the ideological level, that aspects of the use and circulation of language and discourse, as well as its inseparability from culture, are part of what good language documentation work will engage with.

- It is also fairly widely accepted that language documentation projects will engage with the community of speakers and inheritors of the language, and take their desires, priorities, and ethics very seriously.

- All of that discourse and ideology adds up to a clear mandate that we produce text artifacts that are useful to and valued by the heritage language community. In my mind, this mandate in turn demands that we produce and publish “good translations” of endangered language material, and most especially those materials that convey endangered cultural knowledge and endangered forms of verbal art.

- Another entire area of our practice that I must mention but won’t explore in great detail here are the theory and methods necessary to represent aspects of the poetics and the performance of verbal art in the context of language documentation; just a few relevant comments on this are included below.

- As an aside, this discussion about translation in the context of text work intersects in important ways with the work of ‘glossing’, ‘translation’, and ‘exemplifying’ in dictionary writing and grammar writing, but I won’t venture into those areas in this talk.

- To situate this constellation of issues in the intellectual history of language documentation for a moment: because I was well-trained in the ‘discourse-centered approach’ to language and culture at UT-Austin, which itself is deeply rooted in the Boasian tradition, I seem to be far more sensitive to this set of issues than the average bear.

### 6.2 Insufficient disciplinary attention to the theory and methods necessary for ‘good translation’

- Creating a ‘good translation’ is variously considered by translators and translation theorists to be a ‘science,’ a ‘craft’, a ‘skill’, and/or an ‘art’. I myself consider ‘good translation work’ to be all of these things. The perspective that I have developed over time is that we can deliberately learn the science and skills of translation in order to hone our craft and eventually our artistry as translators of endangered languages and forms of expression.

- So, how can I learn to do translation well? By seeking out and working with time-tested and field-tested theories and methods that are appropriate to the specific enterprise of translating endangered language material for diverse audiences.

- Unfortunately, as I intimated a moment ago, most of the discussions of translation that I found in the language documentation literature either:
... said nothing at all about the substance of the issue, merely mentioning ‘translation’ as something to be done;
... provided brief, largely un-theorized lists of ‘do’s and dont’s’;
... problematized the issue, sometimes quite thoroughly, but without offering strong ‘solutions’;
... or even outright disparaged the work of translation in language documentation.

- Lest you think I am being hyperbolic, let me share with you a quote that explicitly articulates an attitude which underlay a number of resources that I consulted. In her chapter on ‘Linguistic Annotation’ in Gippert et. al.’s (2006) *Essentials of Language Documentation*, Eva Schultze-Berndt wrote the following about free translation:

Translation is a skill (many will say, an art) which, if undertaken to professional standard, usually requires a lot of training, and is fraught with methodological problems. It seems highly unrealistic to burden documenters or annotators with the expectation that they ought to provide translations that meet the standards of professional literary or scientific translation. This is all the more so as the translation is often undertaken by someone who is not a member of the speech community and, moreover, is only just beginning to learn the language to be documented and to understand its structure as well as its cultural background. In addition, often a documenting linguist will translate into a language which is not his or her native language... Therefore, all users and potential users of language documentations should be discouraged in the strongest possible terms from using the free translations which are provided as part of the annotation as more than a clue to the meaning and analysis of the documented utterances. (p. 234)

- My immediate and strong reaction to this quote was, “OK, I’ll wait here while you go ‘discourage’ ‘the users and potential users’ in the heritage language community ‘in the strongest possible terms’.”

- The facts that translation requires “a lot of training” and “is fraught with methodological problems” are not justification, in my mind, for not providing/obtaining some good basic training in relevant aspects of translation theory and in field-tested methodologies for producing good free translations of the rich material that it is our charge to document.

- And yet, in both the 545-page *Oxford Handbook of Linguistic Fieldwork*, (Thieberger 2012) and the 340-page book *Language Documentation: Practice and Values* (Genoble and Furbee 2010), the word ‘translation’ doesn’t even have an entry in the index!

- Many of the resources on language documentation and fieldwork that I have surveyed so far (and there are a number more that I have yet to scour) that do mention or discuss translation work take it as a matter of fact – that is, work that must and will be done; but they provide no, or minimal, theory and methodological guidance for producing skillful translations (for example, Bowern 2008).

- Also in Gippert et. al.’s *Essentials of Language Documentation*, Bruna Franchetto provides a thoughtful and helpful, if brief, discussion of the role and importance of translation in language documentation in her chapter on ‘Ethnography in language documentation’. She

---

2 Thereafter, she provides four pages of commentary on a few of the issues one may encounter in the work of translation.
states that translation “must be understood in the widest possible sense, ranging from kinds of transcription and annotation... to translation properly speaking, working from a source-language to a target language.” (Franchetto 2006: 185). She points to what she considers a “vast literature” on translation found primarily in “the areas of literary criticism and poetic theory” and refers us to Swann (1992), Rubel and Rosman (2003) and Bringhurst (1999) as “useful starting points” for language documentation purposes. In addition, for foundational approaches to the representation of verbal art forms, she refers us to Hymes (1977, 1992), Sherzer (1990), and Tedlock (1983).

7 What is a “good translation”?

If we are going to attend to creating ‘good translations’ of endangered language materials, what might that mean, in concrete, practical terms? In brief, these are my basic criteria for a “good free translation” of a text, ordered roughly by increasing complexity:

◦ A good translation attains a high level of ‘faithfulness’ between the source language and the target language(s); everything that is present in the source language is present or conveyed in the target language(s), and nothing more.

... is designed for and appropriate to a specific target audience, based on their particular interests and priorities. The audience may be broad or it may be specialized, but it must be conceptualized and targeted.

... handles disfluencies in the original recording in a skillful, and perhaps even illuminating way.

... is richly and imaginatively annotated and footnoted, with a strong Translators’ Note or Translator’s Introduction that provides ample contextualizing and clarifying information.

... conveys the elements of fluency, fluidity, style, coherence, and artistry that are present in the original ‘performance’ of the text, to the greatest degree possible.

◦ In the same chapter mentioned above, Bruna Franchetto characterizes a “good translation” as “trustworthy” (herself echoing Malinowski) and “competent, something only made possible by allying linguistic and ethnographic knowledge”; one which respects “the meanings carried by the “source” language... thereby respecting its speakers.” (Franchetto 2006: 185)

8 Common attitudes and (false) assumptions

Recognizing the minimal disciplinary discussion of translation theories and methods among documentary linguists and language documentarians, I would like to articulate (perhaps sometimes provocatively) a number of attitudes toward, and (false) assumptions about, translation that I have heard or seen articulated in sometimes explicit, sometimes subtle, ways over the last 20 years:

◦ “We just need to produce one version of a given text, as quickly as possible, so that we can get on to the ‘real’ work of linguistic analysis.”

... and that one version will, and must, be lean, monologic, dense, error-free, and a definitive representation of the source recording, so that we don’t have to think about it anymore.

◦ “As long as I gloss the words right, I’ve done my job.”

◦ “Free translation doesn’t matter anyway, as long as you have interlinear glosses.”
→ Put another way, sometimes the realm of ‘language’ that we work to document is unwittingly limited by a basically “structuralist notion of a language as a system of rules and oppositions” (Himmelman 2006:20) and as a result our default mode is to think and work as though the sentence is the largest unit of meaning that we need to explore when we translate texts.

○ “It is easy to translate between languages as long as you and/or the author is/are bilingual; it’s just like talking combined with taking dictation.”

○ “I am a transparent, passive conduit between the source language and the target language; the author of the source language text is the best and only acceptable author for the target language translation, and must have the final, immutable word on all free translations.”

○ “The cultural, historical, political, etc. contexts of texts are just too complicated and confusing to document well. (So I’m going to ignore them.)”

9 Real and substantial challenges

Having spent some time arguing for greater sensitivity to and commitment among us to take on the work of ‘good translation’, (and especially good free translation in the context of verbally artistic texts), it is appropriate now to face head on some of the real and substantial challenges that we must engage with, both in our theories and in our practice of endangered language translation work:

○ Challenge: I may lack fluency in the source language and culture in which the text is embedded.
  → Strategy: This ignorance is not irremediable. If I can be intensely attentive to detail, self-reflexive, humble, and porous to learning, I can reduce my ignorance steadily over time.

○ Challenge: We may lack fluency in the target language (and culture) as well.
  → Strategy: Again, with time and effort, our fluency will improve.

○ Challenge: The author may lack fluency in the target language.
  → Strategy: In such circumstances, we can seek out and pull together additional resources and collaborators to create a ‘good translation’ by local standards.

○ Challenge: While working with difficult source material, the author may wish to retell or rewrite the text, rather than work patiently through a close transcription and translation, making ‘faithfulness’ among the representation and the translation of the text, and the source recording, extremely difficult to achieve.
  → Strategies: In such cases, we can provide alternate written versions – one more faithful to the recording, the other honoring the author’s revisions. We can also work with difficult material on multiple occasions, negotiate with patience, and consult multiple fluent speakers, until a satisfactory outcome is obtained.

○ Challenge: My time is always limited and full; how much time can I really afford to allot to the subtleties of translation? How many versions of a single text can I really commit to producing?
→ **Strategy:** Large-scale project planning and smaller-scale time allocation boil down to our individual strengths, interests, and commitments, so only you can come up with appropriate answers to these questions – recognizing that your answers may change as your work evolves.

→ **My argument here is that our plans and commitments can be better made if they rest upon better disciplinarily-grounded theories and methods for our inevitable translation work.**

- **Challenge:** This aspect of language documentation work is under-valued, even disparaged, in our field, so it is counter-cultural and professionally risky to prioritize ‘good translations’ and concomitantly dedicate the time and resources necessary to produce artfully transcribed and translated text artifacts (and/or to engage with the theories that produce ‘good’ translations).

→ **Strategy:** Language documentation itself has been disparaged by linguists (and anthropologists), especially of prior generations. The path forward has been to make a case for why this work is important to our larger shared goals; and to improve its standing, along with its own theories and methods, *through doing it*. The same holds for translation work within language documentation.

### 10 A Concrete Example: *Kériiku*

Having talked at length about various issues surrounding the work of translation in endangered language text work, I would like to look at a concrete example of a text transcribed and translated by members of the ILDP team. The text, *Kériiku*, is a short dream narrative told by Ligia Inuma Inuma and recorded in 2008. The recording and the relevant transcription/translation are available at: www.cabeceras.org/ildp_products.htm

It is crucial to underscore the fact that the original transcription and translation that I worked with were produced in the context of a dissertation, with specific goals in mind. I have approached it anew with different goals, and a different audience in mind. No disrespect is intended toward the linguist who produced the original version.

I chose to use this particular text for this talk because it demonstrates a number of important issues, including the following, which are footnoted in the second and third versions in the accompanying text artifact:

- At various points, it lacks full correspondence between the ‘source language’ (SL) and the ‘target language (TL)’; and there are no notes explaining these additions and subtractions of material.
- There is ambiguity between possible senses of single words.
- There is insufficient contextualizing information for a non-local audience (for example, no *Translators Note* and essentially no footnotes).
- Both the style of speaking and the articulation of the author are unusual in many spots, which means that a non-expert would have a very hard time recovering both form and content at some future time.
- It shows the importance of providing a clear and detailed free translation in order to substantiate the inclusion of fast speech material that is impossible to recover from the recording alone.
- It includes several nice examples of the San Antonio variety of Loreto Spanish.
And, regrettably, the author has already died, which means certain information is now unrecoverable, so we have to do the best we can by other means.

11 Concrete strategies for improving the quality of our translations

- Immediately after recording a text in the source language, I can ask for and record a retelling in a target language(s).
- Later, I can obtain a free translation of the text from the author (or another fluent speaker) while playing back the original recording.\(^3\)
- Meanwhile, I can learn as much as possible about the text, its context, and its most salient features (key concepts, characters, places, events, times, etc). I can discuss the text and these other aspects not only with the author but also with other fluent and knowledgeable members of the heritage community.
- Likewise, I can familiarize myself with as many existing resources as possible that provide necessary context, insights, and models for creating my translation (for example, other text recordings and translations from the same or a closely related language, if these exist; ethnographies, collections from other languages that have the qualities you desire, etc.)
- I can record (and archive) my discussion sessions with the author (and others) to clarify key points of the text and its context.
- I can work through the original, line by line, or even word-by-word, with the author – and with other consultants as well in tricky spots – as many times as necessary to provide the best possible translation under current conditions.
- I can work through my transcriptions and translations on multiple distinct occasions, knowing that my understanding will blossom as I deepen my familiarity with the bigger picture of a text.
- After I think I have a good result, I can set the work aside for a week or two, so that I am able to return to it with fresh eyes and a clear mind.
- After I think I have a good result, I can work through my translation with other linguists and collaborators, in order to benefit from their distinct perspectives and areas of expertise.
- I can ask clear and detailed questions about any aspect of the text that is unfamiliar to me, without being embarrassed to acknowledge my ignorance. (Over and over again, I have seen my fellow humans feign knowledge to save face with their interlocutors, and of course I do it sometimes too. Alas, especially in the context of our research, this does us and everyone else a great disservice in the bigger picture.)
- I can resolve my confusions regarding inconsistent or unstable translations of certain words or concepts by probing them on other types of occasions and in different contexts. This kind of ‘sneak attack’ can be an excellent strategy for pinpointing the salient distinctions between similar concepts.
- When I realize that I thought something in the source language ‘meant’ something other than the gloss or translation given by a native speaker, I can ask how exactly my previous interpretation is wrong. (For example, “Hey, I thought that was a gavagai, so what is the difference between that and a gavagai?”) This too can draw out salient contrasts and differences between

\(^3\)Woodbury (2007) calls this a ‘U.N.-style translation’.
similar concepts. Similarly, by asking subtle and careful questions I can actually improve my collaborators’ estimation of my competence and willingness to learn well.

- In general, I can be systematic and patient.
- I can prepared, in terms of both attitude and resource allotment, to produce more than one version of a (at least a sub-set of) the text corpus, because I recognize that the interests and desires of linguists are not the same as those of the heritage language community.
- Knowing that an untranslated text is essentially useless, I won’t let my perfectionism get in the way of my progress toward producing a good translation.

12 Thick translation

- Conceptually, as well as procedurally, we can also aim to produce a ‘thick translation’ in the sense described by Tony Woodbury.
- In his 2007 paper On thick translation in linguistic documentation, Woodbury uses Evans and Sasse’s (2006) paper as a point of departure for recommending “thick translation” of indigenous language texts, which he characterizes as “documenting textual meaning in such a way as to offer maximum transparency to those who may interpret the records we make in a context highly different from that in which we ourselves work: the ‘philologist 500 years from now’. . . . My basic [principle] is that to do this best, the documentary record must show not one ‘finished’ translation, but as many tacks into the translation of an original language entity as can be documented, each representing a certain theory of inter-language correspondence.” (Woodbury 2007: 120)
- He states that thick description “involves at least the following curatable artefacts:”
  - Audio recordings of real-time oral free translations
  - Word-by-word and sentence by sentence translations by an original-language speaker (written or represented in field notes or on tape)
  - Linguists’ morphosyntactic parses, with invariant glosses for minimal elements
  - Linguists’ compositional renditions of parses
  - Drafts of (ever-more) refined literary translations by source-language speaker, target-language speaker, or a collaboration of both
  - Formal poetic analyses of the original that were factored into translations
  - Alternative versions of the same text
  - Literary exegeses, discussions, footnotes, hypertext (written, or in notes or tapes of interactions leading to any of the above)

13 Concluding remarks

- At the most basic level, (and as is sometimes said with despair inside old archives) endangered language materials without some form of translation are (perhaps beautiful but) meaningless shapes and sounds. Ergo, we are all operating from some default “theory” of translation, whether we are consciously aware of it or not.
- Doesn’t it make sense, then, to include an examination of our default theories of translation in our profession – both in our training, and in our practice – and to engage those default theories via more explicit and fine-tuned ones? From there, we can much more effectively
design and implement our strategies and methods for meeting our own professed goals as language documentarians.

- To do good translations that do justice to the unique character, style, and artistry of the original text takes time and focused effort. If this is important, then we must make the time, and make the effort, to produce good, even excellent, translations.

- Excellent translation, whether it is an ‘art’ or not, rests upon some solid, learnable principles and skills, which can be ever improved through practice. So again, if this is important to us, we must acquire those skills and practice, practice, practice.

- The bibliography below offers various useful starting points if you wish to engage more with the science, craft, and art of translation in the context of your own language documentation work.

14 Bibliography


Bringhurst, Robert. 1999. *A Story as Sharp as a Knife.* University of Nebraska Press.


