POSITION PAPER

# Training graduate students and community members for native language documentation

Judith M. Maxwell

In a world where "remote" communities have cell phones, some Internet access, and are tied to national and international webs of commerce and socialization, linguistic study must be attuned to the desires and expectations of the host community and graduate students must, in addition to training in phonetics, transcription, use of technology for recording and analyzing the language, learn to work with language brokers and develop goals for their research conjointly with the speakers who facilitate their data, while writing for the (academic) community. Student projects may range from salvage to revitalization, from documentation to theoretic exploration. Projects co-constructed with the host community promise to give both local and academic communities the greatest satisfaction in goals, products, applications and dissemination.

## 1. Preparing to go to the "field"

While it is still possible to pick an un(der)documented language from the large set of endangered languages (which includes languages with unique case systems, evidentials, or other linguistic features of interest) and set off to "the field," this is a vanishing scenario. Some nations, such as Canada and Brazil, strictly regulate research contact with indigenous peoples. Even without explicit government policies overseeing interventions in indigenous communities, many autochthonous groups have grown weary of being "objects" of investigations, investigations whose final reports, published findings, books, journal articles, and so on seldom make it back to those whose knowledge is being (re)presented. Many community members suspect that past researchers have made fortunes off their publications and so now demand a share in supposed future profits. In other cases, members of studied groups have been able to pursue education (formal or informal), gain fluency in a hegemonic language, and have found in print exegeses of their home languages and cultures. Still others have trained as linguists and anthropologists,

obviating a need to appeal to "outside" experts for documentation. In a world where the myth of the "naive" informant has been exploded by the mass media, educational institutions, cell phones, and the Internet, it is the alternative of ethical field practice, which involves the native speakers and home communities in the co-construction of knowledge, that becomes not only "right" but increasingly the most feasible path to data collection.

Granting agencies now routinely include in their checklist of concerns to be addressed (a) collaboration with local institutions,[1] and (b) plans for dissemination of the findings in the host country (if not community). In 1985 at the American Anthropological Association meetings, a group of linguists working on Mayan languages of Guatemala and Mexico, produced a manifesto of best practices for field research. A key point held that native speakers be involved at all stages of the research, from hypothesizing a research question through data collection and analysis to publication and dissemination of findings. Some researchers present at the meeting refused to subscribe to the document. They felt that a requirement to consult with anyone as to the topic of investigation was an infringement upon academic freedom. But as access to speaking communities comes more and more to depend upon the goodwill and cooperation of these host communities, agreement between researchers and language users on the ends, goals, products, and methods of investigation becomes *sine qua non*.

Admittedly, it is not always possible for a nonautochthonous researcher to have extensive contact with the host community before first writing for funding for feasibility studies, but these pilot studies should bring the researcher into close contact with the language users. Their desires and expectations can then be incorporated into the plan of investigation eventually proposed. When the researcher is a member of the language group studied, the central concerns of the group may be more patent. Luckily, with a language death rate of about 34 languages a year, some foundations[2] have made funding available for documentation for documentation's sake, saving corpora for posterity.

Expected outputs, results, and products of field research may structure the project itself. Let us examine a few typical goals of field research, with their concomitant skill sets and resource requirements.

---

1. Collaboration or affiliation with a "local" institution is most commonly required by agencies funding international research. In the United States, grantors show less concern for explicit institutional ties with groups internal to the U.S. borders, even if recognized legally as separate nations, cf. with the Navajo, Ojibwe, Mohawk, etc.

2. These funding agencies include the Foundation for Endangered Languages, the Endangered Language Fund, the Documenting Endangered Languages Program of the National Science Foundation and the Hans Rausing Endangered Languages Project.

### 1.1 Salvage linguistics

In "salvage" linguistic fieldwork, the scenario is typically one of assumed imminent extinction of a language. The primary goal of the research is to record as much of the language as possible in a short period of time. Some projects, such as Robert Howren's 1965–69 survey of the languages of the northern tier of Canada and Alaska, were envisioned primarily as salvage operations. Each researcher was to document as much language as possible on audiotape, write a sketch grammar, and file these materials in an archive to be established at the University of Iowa. The charge to fieldworkers did not include any deposit of data or analyses in the host communities. As time ran out on the grant, languages were reprioritized by danger of extinction. I was shifted from my slated research language, Ingalik (which already had a sketch grammar written), to Han.

Han is an Athabaskan language spoken in Eagle, an indigenous Alaskan community on the Yukon River, five miles from the Canadian border. Upon my arrival in June 1969, via mail plane, I explained that I wanted to document the indigenous language before it disappeared. At that time, only adults over 50 spoke the language fluently; adults between 20 and 50 showed comprehension but responded to their parents in English; children of these younger adults knew and used only isolated words of the language, such as *guu'guu'* for 'owl'. The college-educated chief, then in his midtwenties, was happy to have some scholarly attention paid to his language and culture. He and his wife introduced me to the heads of the 50 households in the town. The most fluent elders adopted my project as theirs. One man, Mr. Paul, appointed himself as my spiritual as well as linguistic guide. He carefully chose the stories he would tell me and my tape recorder to highlight the principles of living a indigenous (moral) life. Our sessions were usually attended by a bevy of children, who repeated along with me when I struggled with new words, tried to understand verb conjugations, or ran through syntactic paradigms. Mr. Paul saw his mission of working with me as leaving important lessons recorded for these children and theirs. However, my brief did not include archiving these tapes locally, nor did I have the equipment to make copies from my trove of reel-to-reel tapes made on my Uher recorder. I made sure both Mr. Paul and the chief had the address of the University of Iowa. Though it was not part of the project plan, the enthusiasm of the children and the concern of the elders prompted me, during the last several weeks of my field session, to spend the time after daily recording sessions transcribing some materials into reading lessons and exercises to be used in school language classes. The previous year, the Bureau of Indian Affairs had closed the all-Indian school. Alaska was integrating indigenous students in the "general" population. One school would serve Indian Eagle and its white namesake three miles away; one teacher would teach all six grades. When I

consulted him about teaching in or about Han, a language that he, as an Anglo from another community, did not speak and had not heard, he had expressed enthusiasm about enlivening the classroom for both Anglophone and indigenous students. I presented him with three workbooks when I left in August. Two months later, I called to see how the "enlivening" was proceeding. He announced that as there was one more white child than Indian in the school, bilingual education was not required by state law, and so he had discarded the materials.

The goals of this project, as set by the principal investigator, Robert Howren, were met. I cached reel-to-reel tapes of Han tales, example paradigms, and carefully enunciated conversations in the University of Iowa archive. I wrote a simple descriptive grammar. My supervising professor, John Ritter, and I added a study of Koyukon, a neighboring language that Ritter had investigated, as a comparative codicil to this project. The goals of the indigenous collaborators can only vaguely be said to have been met. The language has been preserved should future Han children seek out the archive in Iowa City. The moral lessons that Mr. Paul carefully chose are waiting to be rediscovered. But bilingual education never brought texts in Han to the Eagle schoolchildren, thanks to the 50 percent rule enacted by the State. We must note that such education was not a goal of the research plan, though it resonated with the hopes of heads of families who did not understand why simple language competency was no longer assimilated along with fry bread. The five remaining fluent speakers of Han today are living large parts of the year in Fairbanks or Whitehorse to be near medical facilities and their grandchildren, who have left Eagle to search for work and education. John Ritter remains in the Northwest Territories as founding director of the Yukon Native Language Centre. He has added to our knowledge of Han. Publications of the Centre are available in the Northwest Territories, and, to a degree, in Alaska.

## 2. Salvage linguistics, education, and revitalization

At times, local school districts have worked together with indigenous peoples off-reservation to combine the ends of salvage linguistics (data collection) with education and revitalization projects. The Lansing Michigan School District serves as another short case study. In 1976, the School District received federal funding for bilingual education. Ricardo Briones, then head of the bilingual program, defined his mandate to include all the major heritage languages of his district: Spanish, Vietnamese, and Ojibwe/Anishinaabe. Having decided to offer Anishinaabe, Briones was faced with the problem of finding pedagogical materials. Canadian schools had texts for K-12 in Chippewa/Anishinaabe. The dialect differences were

minor, but the local Ojibwe community rejected the Canadian materials. They did not find the orthography transparent or accessible, nor did they find the cultural content commensurate with the lived experience of their children. I was called in to assist an elder, Elmer Miner, who was selected by the Lansing Band leadership, in developing basic reading materials and a glossary to be used in the municipal school system's bilingual education program. Mr. Miner and I wrote short readings and grammar lessons and compiled a glossary. At the Band's Thanksgiving celebration, we were able to present the tribal council with the first draft of the materials. We sat in council with the elders for four hours as they read through each lesson, evaluating the cultural relevance and accuracy. The approved materials were then passed on to the Lansing school system. Ojibwe/Anishinaabe language lessons were offered to indigenous and nonindigenous students alike.

This program did not survive Briones's term as director of bilingual education. However, the Band retains the materials, and indigenous schools do offer language classes to children, youths, and adults, so they can proudly say, *Awi-gikinoo'amaagozig*! 'Get to class!'.

For a short time, the program brought an indigenous language to the curriculum and consciousness of nonindigenous as well as indigenous students, and thus met the goals set by the Lansing School District. In terms of the goals of the Lansing Band of the Ojibwe, the project produced materials that are still in their hands and are in use. While the language is not replacing English as the language of daily communication, even among Band members, it is actively deployed as a marker of identity. Its study binds the group together, providing shared symbols for lived and historic experience.

Colette Grinevald (Craig)'s work with the Rama of Nicaragua displays the triad of linguistic salvage, education, and revitalization in a slightly different light.[3] The Rama are divided into two groups: one, largely urban, dwelling on Rama Cay; the other, largely sylvan, living in jungle areas along the river. Grinevald worked with the Rama revitalization project from 1985 to 1993. In her initial work, she found that the urban dwellers had essentially switched from Rama to Spanish. The sylvan Rama, whom the town dwellers referred to as "Tiger" Rama, were still fluent, even Rama-dominant. However, Grinevald's attempts to simply teach "Tiger" Rama language to Rama Cay dwellers were resisted. The urbanites rejected the encapsulated world view they felt and heard in

3. An official view of this project can be found in Grinevald (2003a). Much of the affective information I adduce here comes from informal conversations with Dr. Grinevald. Her early work on Rama was published under the name Craig, while more recently she publishes under the name Grinevald. Citations here and in the References use the name of the publication.

this speech. Still, the Rama Cay dwellers were representing the indigenous nation in the Congress of Indigenous Peoples established by the Sandinistas. Addresses in this Congress were given, or, at least begun, in the language of the autochthon representative. The Rama did not wish to be forced by monolingualism to speak Spanish in this counterhegemonic arena. They accepted a bilingual education plan that brought back a select and reduced vocabulary for the emblematic deployment of language as a key emblem of identity. This limited language reintroduction has resulted in resurgence of language use within restricted contexts. A Rama teacher, "Miss Nora," is credited with bringing the emblem into active spoken discourse (Craig 1992b).

While not the revitalization of the language envisioned by Grinevald when she began her work, the current use of Rama by speakers on Rama Cay, as well as on the "Tiger" mainland, has led speakers to proclaim that their language has been "rescued" (Craig 1992b: 80, 88).

## 3. Documentation and ethnic identity

Another set of documentation studies has been commissioned or added to the process of ethnic recognition. In the United States, recognition as an indigenous group by the federal government allows the group a degree of autonomy in self-governance and political, social, and economic organization. Recognition is thus a boon much sought after. Many indigenous groups have retained lawyers specialized in land-tenure issues; others have contracted with linguists to document their linguistic heritage, possession of a distinct native language being a crucial criterion for proof of legitimacy. The case of the Louisiana Houma, an historic amalgam of Muskogean groups recently disbanded in an attempt to more clearly document their cultural continuity as required for recognition, is illustrative of the dilemma of indigenous groups who quickly assimilated linguistically. Today's disamalgamated Houma speak French. They trust that it is a French not only significantly different from the surrounding Cajun and Louisiana Regional French, but also that each band has enough unique language forms to show indigenous substrate influence and an unbroken chain of descent.

Similarly, in 2003 the Congress of Guatemala recognized a 2nd Mayan language, Chalchiteko. Speakers of this language had petitioned for recognition on the basis of their dissatisfaction at being classified as speakers of Awakateko. While Awakateko and Chalchiteko are mutually intelligible and share core vocabulary and rules of morphology and syntax, Chalchitekos felt excluded from the process of standardization, materials development, and bilingual education,

processes funded through the Academia de las Lenguas Mayas de Guatemala and through the Dirección General de Educación Bilingüe Intercultural. The Chalchitekos highlighted unique lexical items in their neighborhoods and persuaded legislators that they spoke a previously undocumented language. Now, with federal recognition, they get separate budget lines from the two funding entities and can carry out their own linguistic surveys, develop reading materials, and produce academia-mandated documents, such as translations of laws, health warnings, and official notices. Without the help of trained linguists, the Chalchitekos have been able to defend and present their language, both as an emblem and as a tool for development.

## 4. Documentation and linguistic theory

Much linguistic research is not simply documentation for storage. Linguists also actively seek to understand the human language faculty. The breadth of variation among and between languages suggests the limits of the possible structures of linguistic cognition. For example, consider what has been said about the milestones of development in first-language acquisition. The concentration of such research in modern times on English-speaking children, or children learning Indo-European (I-E) languages, has led to some assumptions about meaning and acquisition. The tendency of children learning I-E languages to use uninflected roots in the early phase of holophrastic speech suggested that acquisition was meaning-driven. Children learn the roots because they are the chief content-bearing elements. However, cross-linguistic research, such as that of Clifton Pye (1980, 1981), showed that, at least for K'iche' children, the perceptually salient syllables were those reproduced. In K'iche' Mayan, these tend to be derivational and inflectional endings, often leaving the roots unexpressed. Reexamination of the I-E languages reveals that stress and intonation tend to fall on the root syllable(s). Thus, perceptual saliency has come to replace semantic weight as the key explanatory mechanism in the process of morpheme acquisition.

Similarly, linguists often choose to study a language or languages known to have certain grammatical properties in order to contribute to our understanding of syntax, morphology, phonology, or pragmatics. Dixon (1994) surveys languages to trace the split in ergative-absolutive systems. Languages from a number of unrelated families have been compared to understand the function of noun classifier systems (Craig 1986). Evidentiality provides a rich body for cross-linguistic study (Chafe and Nichols 1986; Aikhenvald 2004). Comparative cross-linguistic studies enrich our understanding of child language acquisition (Slobin 1985,

1992). Indeed, any facet of the linguistic system can serve as the focal point for a problem-oriented linguistic study.

## 5. Preparation for the field

These different types of documentation projects require different kinds and levels of linguistics skills and training. Before heading to the field, most graduate students take courses in phonetics, phonology, morphology, and syntax. The Linguistic Society of America's Linguistic Institute offers intensive summer courses that hone these "nuts and bolts" skills, while also deepening appreciation of current theories and controversies. Field-methods courses provide practical training in the application of these skills, while refining skills in elicitation, recording data, and restructuring interview schedules through feedback from elicitation settings, analysis, and presentation of findings. For example, Carol Genetti has designed a summer field-study institute meant to alternate with the current LSA Summer Institute, which would prepare students, professors, and other linguistic professionals for the exigencies of field data collection. This institute, the Institute on Field Linguistics and Language Documentation (InField), offers training in elicitation techniques; data collection from narratives, texts, and so on; practice with data organization and storage; tools to aid with analysis; and resources for archiving data and findings. Workshops are offered, covering available technologies; principles and tools of transcription; "life in the field"; and involvement with community efforts in language planning and revitalization. (See http://www.linguistics.ucsb.edu/faculty/infield/ for a description of the 2008 program.)

University courses in pragmatics, discourse, and sociolinguistics add in the complexities of cultural presuppositions, shared understandings, schemata, and communities of practice. Style, register, and code shifting introduce students to variable rule analyses. Practice with statistics and SPSS (Statistical Package for the Social Sciences) helps students perform multivariate analysis of sets of factors. Thus prepared, a graduate student could take any language, documented or not, and proceed to record its forms, applying modern theories to the analysis and presentation. If the language has already been studied, the student can avail herself of the grammatical analyses presented, read through texts provided, and focus her study on underexplored areas of the grammar. An example of this can be seen with Colette Grinevald Craig's analysis of Jacaltec [sic] grammar (1977), built on the foundation of Christopher Day's phonological and morphological treatise (1973). In her doctoral dissertation on verbal categories in Kaqchikel (Maya), Roberta Hendrick Krueger (1986) built on William Cameron Townsend's work on Kaqchikel verbal morphology (1937).

## 6. Quick and dirty run-through of pre-field considerations, needs, intellectual tools, and physical tools

### 6.1 Field consultants

Different types of field research require different kinds of sampling and interviewing techniques. For any extensive interaction with a speech community, however, one or more principal native speaker consultants will be needed. For survey-style research projects that sample across a population, native speaker consultants can facilitate contacts with different demographic segments of the community; they may be able to recruit talented interviewers with fluency in the language and the requisite cultural skills to know "how to ask" (Briggs 1986). For in-depth study of a set of phonological, morphological, or grammatical categories, one or two fluent speakers may provide the deep-level competence necessary for a theoretically nuanced description.

The ideal native speaker language consultant is a trained linguist who can read and write his or her own language, as well as a contact language, if the field researcher is not fluent in the language under study. Competence with word-processing and text-analysis programs, such as the SIL Interlinear Text editor program (IT), Shoebox, or Toolbox (for interlinear text transcription and preliminary analysis) and Concordance, Micro-OCP, MonoConc, TA, or TACT (for building concordances from texts and producing preliminary glossaries and word lists) is a plus. Consultants who can help with or take primary responsibility for transcribing texts electronically and archiving data not only streamline the process of collecting, organizing, and storing data but may also serve as checkers on errors of perception, recording, or analysis. Often, this ideal native speaker does not exist, or is otherwise employed. However, engaged native speakers with a keen interest in the language can quickly add to their skill set in order to assist the researcher.

Finding fluent speakers of moribund languages may be a challenge. If the research is community sponsored, the community liaison may be able to steer the research to interested fluent elders. Elders often have time available to work with the researcher for extended periods. If, however, these elders have lost their habitual co-locutors, they may be more in the status of "rememberers" than speakers of the language. Getting extended "natural" language samples may present difficulties; elders may have developed a "performance" repertory for the language, playing guessing games as with children: "Do you know how to say 'dog'?... dog is/.../?" and so forth. Still, the senior speakers may be willing to dedicate much time and energy into recouping as much of their proficiency as possible, especially when the community is committed to preservation or revitalization of the language.

Work with such elderly speakers can be richly rewarding in cultural heritage terms as well. Texts may serve as sources for ethnographic as well as linguistic study. Some researchers have found their willing elderly consultants physically limited, however, by loss of hearing and/or teeth. The community leaders sponsoring the Xinca revitalization initiative in Guatemala have taken their elders to Antigua Guatemala to be fitted with hearing aids and dentures.

Research that focuses on a particular segment of the population may require concentration of consultants within that segment, rather than sampling across genders, age groups, and socioeconomic and educational levels. Sometimes the gender of the researcher impedes access to "natural" language sampling across gender lines. David Carey (2001, 2006) found that despite the level of trust he earned through four years of presence within the Kaqchikel community, research into women's narratives was best facilitated through the intercession of female assistants as interviewers. Access to children, in the study of language acquisition, socialization into gendered speech patterns, grammatical innovation, word play, and so forth, may also require careful attention. Generally, internal review boards are especially sensitive to issues of informed consent when dealing with research on children.

In those projects with strong community participation in planning and execution of the research project, some consultants may be designated by the community leaders. These participants may be chosen for their skills and availability or for political connections, as well as for their interest in the project or fluency in the language. The *principales* (council of elders) of the Chuj community of San Mateo Ixtatán selected the men who would receive training as linguists through the Proyecto Lingüístico Francisco Marroquín (PLFM) from 1973 to 1977. These men met the criteria set by the PLFM: they could read and write some Spanish, they were fluent speakers of Chuj, and they had no more than three years of formal schooling. One of the goals for the training period was to write a dictionary of the language, including "all" the commonly used vocabulary. Having no women chosen limited access to domains of weaving, embroidery, cooking, washing, childcare, birth, counting the days for 'beating' the fruit trees, and other "feminine" enterprises. Another requirement set by the PLFM, working outside of San Mateo for a week at a time, effectively cut off later recruitment of women, since married women had responsibilities in the home and unmarried women could not travel alone or in the company of unrelated men.

In community-sponsored projects, payment for the time and expertise of the consultants may be managed by the community partner. In those cases in which an outside researcher is organizing the work and payment schedule, care should be taken to adjust the pay scale to local standards. It may be difficult to determine what these standards might be for consultant work. The comments of Fiona

McLaughlin and Thierno Seydou Sall (2001) indicate how integrated into the local social network of status and responsibilities consultant pay can be. Some consultants may be willing to donate their time for the good of the project and the benefits to accrue to the community. However, donated time may cede precedence to financial exigencies, so it may be difficult for unpaid consultants to consistently set aside blocks of time. Other consultants may have expectations for payment set by experiences with nongovernmental organizations (NGOs) and so be unwilling to work for an outsider at the local scale.

Once a cadre of consultants has been selected, there may be some continued turnover as interest or opportunities for individuals vary or wane. Consultants may also show differential interest in acquiring technical understanding of their language and/or skill with recording and transcription.

How the consultants understand the research project will greatly influence their commitment to it and their approaches to working with the researcher. Some linguists present themselves as wanting to learn the language. Consultants who see themselves as language teachers may be frustrated if or when the linguist is more interested in full paradigms than speaking fluency. Consultants for studies which seek to explore "naive" patterns of language use in relations of power, gendered environs, or formal oratory may rapidly lose their naiveté. Not all consultants on a single project need have the same range of interests, abilities, or responsibilities.

Consultants may be especially adept at fashioning appropriate final products for the home community, understanding the relevant media, styles of presentation, and the like.

## 6.2 Equipment

Digital recording equipment, both audio and visual, is now readily available. Most computers come with playback facility, if not with high-quality recording capabilities themselves. The recording quality can be upgraded with good microphones and software, but a separate small digital recorder is a handy tool, as well as a welcome backup. Recording speech sounds requires a signal-to-noise ratio of 45 decibels with a frequency response range of 60 Hz to 12 kHz. A good microphone is essential. If one is recording a single speaker, one can ask that speaker to wear a lip mike. These are actually dual microphones, one aimed at the speaker's mouth, and one angled away; sound picked up by both mikes is factored out of the signal, effectively eliminating ambient noise. For conversation, multiplier jacks can give conjoint input from lip mikes. It is still difficult to transcribe through overlap with this system, unless each mike is programmed to lay down a separate track that can be isolated. For conversation, researchers may still have to rely on a single centrally located microphone, which can pick up multiple speakers, as well as the situational background noise.

Many good software programs are available for acoustic analysis of recordings. A few of these are Praat, Dr. Speech, Multi-Speech, Signalyze (for Macs only), and SoundScope. Such programs can process the input-recorded speech, yielding displays of spectrographs, pitch contours, and fast Fourier transforms.

The addition of video recording can help situate the language use in its cultural context. When used in close-ups, it may also aid with identifying articulatory gestures. Some field phoneticians also carry equipment for palatography (olive oil, powdered charcoal, dental mirrors, and a camera). Equipment for more complicated articulatory monitoring and tracking, such as magnetic resonance imaging and computed tomography, is not usually taken to the field. Rather, speakers are asked to come to laboratory facilities.

## 6.3 Keeping up

It is generally easier to elicit data, especially when recording it electronically, than it is to transcribe, input, sort, and analyze. In order to keep on track for each work session with a speaker, one should have analyzed the preceding elicitations, checked hypotheses, and prepared a general plan for the upcoming session before it begins. Transcription, even with the aid of new software, still takes more than twice as long as initial recording. Transcription must be followed by analysis to keep the research on track.

The level (fineness of detail) of transcription will vary by research project. Some kinds of discourse analysis can be adequately carried out with a "practical" orthography, which may be the standard orthography of the language or a field product of phonemic analysis. Other studies may require phonetic detail. For example, Giles et al. (1976) found that aspiration of voiceless stops among Francophones signaled political ideology, even though it was subphonemic. Studies of gender differences in the speech of American women and men have shown differences in the degree of raising of /æ/, the incidence of alveolar versus velar nasals in the gerundive ending /-ing/, and the neutralization of the /a/ ~ /ɒ/ contrast.

In discourse analysis, some care should be taken, at least in preliminary transcription, to include all the utterance details, including pauses (length can be sociolinguistically significant), hesitation syllables, false starts, and repetitions.

The abundance and richness of linguistic data recorded can be a challenge for "keeping up." When field time is limited, it is tempting to record every day, for as many hours as possible, from as many speakers as possible. Such intensive collection makes daily transcription and analysis of the whole corpus impossible. When the researcher is also a primary interviewer, presence during the recording session can help monitor the issues and forms that arise, clarify paradigms, and

suggest the structures for successive sessions, where these are not set by a predetermined interview matrix or schedule.

## 6.4 The ad rem experience, a.k.a. "it's all data"

While the field linguist is responsible in partnership with community counterparts for structuring the research project, framing the question, setting goals for specific parts of the investigation and daily work sessions, the consultants also contribute their knowledge, expertise, and sense of necessity. Consultants may become attuned to what the linguist wants (beware, lest this skew the data provided) and can rapidly complete paradigms or provide forms that have no ready translation equivalents (like sets of incorporating antipassive constructions). They may also express discomfort with an elicited phrase or criticize productions heard in recordings (their own productions or those of others). These observations, along with corrections, emendations, and other spontaneous productions, are important insights that can speed up, redirect, or complicate analysis.

Part of being in the field is co-living, participating in community life. It will inevitably happen that a key phrase, word, or form will be heard in the "wild," not in the "tame" confines of the consultant recording session. Some research projects in the U.S. have sought to capture spontaneous production, if not these random incidents, by leaving recording devices on (with voice activation) in public spaces of the home (Leto de Francisco 1998). But even this massive sampling will miss unique, often crucially relevant productions. When heard "on the fly," these forms should be preserved in the amber of the researcher's mind (or pocket notebook) until s/he can work again with the consultants and explore the well-formedness, ubiquity, and pragmatics of the production.

Just as the researcher is part of the community in terms of interacting verbally and thus hearing data at all times, the researcher is socially part of the community. It is important to negotiate an identity in which one can be relatively comfortable. In early research projects, such as the Harvard and University of Chicago Chiapas projects, many female fieldworkers assumed, or were cast in, male roles. Norman McQuown (personal communication) noted that a young woman known to the community was depicted in an ethnographically adjusted Thematic Apperception Test that was used as a prompt for text elicitation. Though female, this researcher wore pants, had short hair, and spoke in the street with men, daily. She did not work in the home. She was then functionally male and was referred to with male nouns in the elicited narratives. Rosemary Joyce (personal communication) says she prefers a masculine self-presentation for the field, as it gives her more freedom of action. Other researchers, such as the Rosaldos (M. Z. Rosaldo 1980; R. Rosaldo 1980), found that, as a married couple, they had access

to different community bases and interactional spheres. Carol Hendrickson (1986) found that learning to weave provided her with the perfect format for extended periods of interaction.

One's happiness in the host community and satisfaction with the field experience will depend in large part on the social networks established and one's relation to them, but small physical things may also enhance or detract from the overall fulfillment. A simple example is that those who are vegetarian might be happiest working in vegetarian societies or should be prepared to gracefully accept meat obtained at great expense and to be served to honored guests. Those with corn allergies would be well advised not to work in Mayan communities. The degree to which one can be independent of local norms of food, dress, and hygiene varies with the degree of integration: Does one live with a family as an adoptive member (one way of defining rights and obligations, as well as necessary terms of address)? Does one have a separate household? Does one live not in the local community but in a neighboring urban center, either commuting or asking consultants to do so? Those doing comparative research among communities, dialects, or languages may find that the advantage of central location outweighs the serendipitous possibilities of participating in everyday conversations and life in a given community. Even with speakers commuting to a central place to study with the researcher, or a host institution, or simply to work as consultants, data will abound.

## 6.5 Know as much as you can before you get to the field

For languages with long histories of linguistic documentation and research, such as English, French, Sanskrit, and Korean, it may not be possible to read everything that has been written on the structure of the language before beginning one's own research project. Nonetheless, to properly situate a particular study, one should have a solid background in that research that most closely pertains. For languages with less documentation, one can try to be more exhaustive in covering the published literature before heading out. Some languages may have a literate tradition that has not been subjected to close linguistic analysis; novels, poems, and religious treatises can provide rich sample texts. Other languages may have had early linguistic surveys that resulted in little published beyond word lists and sketch grammars. Still other languages may have been partially researched or briefly studied, with no publications. One may have to travel to archives to locate holdings. Many of the northern Athabaskan languages surveyed by Robert Howren's team in the late 1960s and early 1970s were never described in written grammars or glossaries, though the data are stored at the University of Iowa. Nearly all of the prodigious research conducted by John P. Harrington on California indigenous languages is archived in Suitland, Maryland, at the National Anthropological

Archives of the Smithsonian Institution (Harrington 1981–91). The increasing use of electronic databases and clearinghouses is ameliorating this problem.

If the language itself is unstudied or understudied, and if it is not an isolate, looking at works on closely related languages may help orient the researcher to forms, structures, and problems of interest for the field season.

In some cases, earlier work done on the language of study may not have been made available either in raw data form or in final published format to the original consultants and/or host community. Taking a copy or copies of earlier work to the field to "repatriate" can be much appreciated if a reasonable representative recipient (such as the original consultant, an heir, the town council, the school library, or a community library) can be found. Having one's own copy of earlier works, especially dictionaries or reference grammars, can also ease one's processing of the data on a day-to-day basis.

For those languages with a literate and/or research tradition, familiarity with this literature and prior research will also ease integration into the local scholarly community, establishing a common basis for discussion and respect.

## 6.6 When the researcher is "from" the field

Preparation for the field may be different when the "field" is one's home community. Just as speakers of English in the United States can and do study the speech around them (cf. Labov 1982; Tannen 1994; Bucholz 1999), it is also the case that speakers of "minority" languages may become actively involved in the preservation of their native tongues. In some cases, they move from naive to sophisticated language consultants. Outside researchers from Sapir on have noted the utility of teaching native speakers to read and write in their own languages. Sapir (1949/1985) based his argument on the psychological reality of the analytic unit – the phoneme – on native speaker transcription. But native speakers are increasingly seeking formal training as professional linguists and working to document their own languages.

Leanne Hinton has been instrumental in establishing mentor-apprentice learning modules for speakers of endangered California languages. The model developed there has been successfully transported elsewhere (for details, see Hinton 1994 and Hinton and Hale 2001).

In 1972, the Benedictine brothers in Antigua Guatemala turned over to Robert Gersony their brainchild, a program for speakers of Mayan languages that would train them to become linguists and to control their own language resources for materials development, education, documentation, and preservation. Gersony established a lay institution, the Proyecto Lingüístico Francisco Marroquín (PLFM). Financed by teaching Spanish to foreigners, the PLFM funded

Maya students in their study of linguistics. The first students were chosen from the three largest groups of Mayan language speakers – Mam, Kaqchikel and K'iche'. The criteria for selection were (a) fair spoken competence in Spanish, (b) literacy in Spanish, and (c) no more than a third-grade education.[4] In the second year of the program, Chuj, Q'anjob'al, Akateko, Q'eqchi', Tz'utujil, and Awakateko were added. In the third year, Ixil, Popti', Ch'orti', and new variants of Kaqchikel and Mam were added.

These students received formal, intensive courses in phonetics, phonology, morphology, and syntax. However, the bulk of their training came during the research phase. The students had two mandated projects: the construction of a bilingual dictionary, and the writing of a sketch grammar. In the lexicographic project, they learned the use of regional checklists of flora and fauna, grammatical worksheets, monosyllabic root tables, and text collection. In the process, they documented dialect variation, differences in male versus female terminology, specialized professional jargons, and register and style shifts. The Maya students were the primary researchers.

Early in the training process, the Ministry of Education supported this initiative, providing the physical space for the linguistics school, and providing typewriters and paper. They awarded the graduates of the three year course a *perito* degree, roughly equivalent to a high school diploma. However, with time, the Ministry withdrew support. First it reclaimed the school building, then the typewriters. Finally, it refused to recognize the training as degree worthy.

Nonetheless, students in the three entering classes finished their training. Publication of the dictionaries was delayed by the violence of the civil war. Tz'utujil students buried their work in sealed clay ewers. The Chuj separated their dictionary printouts and stored them along roof poles, both in their home communities and in supposed safe houses in urban centers. Much of this work was never found and reclaimed, as over half the members of the Chuj research team were killed or driven into exile. Today there are dictionaries of all the original languages.

Nora England, one of the original PLFM linguists, persuaded the board of directors of the PLFM, all alumni of the training program, to resume professional preparation of Maya linguists. This time the restriction on academic preparation was lifted. England recruited a corps of bright young Maya, excited to learn about their languages and the science of linguistics. Rather than resorting to the earlier model of preparing technicians capable of carrying out lexicographical

4. At the time, there was a strong tendency for students who completed elementary school to leave their natal communities and go to urban centers in pursuit of further education and/or employment. It was felt that students with less formal education would return to their homes, assuring that their training would be put to use within those communities.

tasks, writing and reading in their language, England trained a cadre of linguists, who could and did examine their language for insight into universals of grammar. The first cohort of graduates incorporated in 1990 as Oxlajuuj Keej Maya' Ajtz'iib' (OKMA), a research group that funded itself by winning competitive grants. OKMA has trained two more "classes" of graduates. Together they have conducted dialect surveys in the major languages and written and published pedagogical, descriptive, and prescriptive grammars. They are currently involved in developing Internet and CD-ROM language courses, both for native speakers and for nonnative learners. The group has elaborated on monolingual and bilingual dictionaries, and advised the Ministry of Education. In August 2009, OKMA formally disbanded, turning over their publications and research to the Ministry of Education. The Vice Minister of Education, Manuel Salazar, pledged to establish a research section of the Ministry to continue OKMA's linguistic investigations and efforts toward standardization.

In addition, the Universidad Mariano Gálvez de Guatemala established the Escuela Lingüística and provided indigenous students with 50-percent tuition waivers. Students in linguistics at the Mariano Gálvez learned basic descriptive linguistics (phonology, morphology, and syntax) but also took classes in applied linguistics, particularly bilingual education and language planning. Students from this program have gone on to form the core of the Ministry of Education's Bilingual division (Dirección General de Educación Bilingüe Intercultural, or DIGEBI), as well as serving their communities in the Academia de las Lenguas Mayas de Guatemala (ALMG).

Likewise, the Universidad Rafael Landívar established an Instituto Lingüístico and, with USAID (U.S. Agency for International Development) money, trained three cohorts of Maya as linguists. Most of these graduates either continue to work in materials development with the Institute or are serving their linguistic communities in the ALMG.

One of the most recent initiatives is the elaboration of pedagogical neologisms, begun by the Kaqchikel Cholchi', the Kaqchikel branch of the ALMG, in 1994. This was taken up by ALMG nationally in 2002 and cosponsored with DIGEBI in 2003–4. Kaqchikel and other Maya educators felt that, in adhering to the requisite national curriculum, they were straightjacketed into teaching in Spanish, because many of the concepts taught had no translations in a Mayan language. New teams of pedagogues and linguists were brought in to receive specialized training in the creation of neologisms. In 1995, Martín Chacach and I published a booklet outlining the criteria and strategies used in the 1994–1995 Kaqchikel Cholchi' project (Chacach and Maxwell 1995). In 2003, I revised these criteria for the project with DIGEBI and the ALMG (Maxwell 2003) and trained volunteers from 11 linguistic communities to work on vocabulary needed for the

new intercultural curriculum developed for grades K-12. The principal criteria were (a) respect for the rules of the language, (b) transparent derivations, and (c) that the newly created forms be "non-kilometric" in the sense that they be lexical items and not descriptions. The first criterion required knowledge of the rules of the language beyond simple native-speaker competency. Understanding of derivational and compounding processes was crucial. Not all the volunteers had been trained as linguists, and some were schoolteachers. But by working in teams that each had at least one linguist, over 2,000 new words in each language were created. These were later ratified by panels of community elders, DIGEBI and ALMG linguistic technicians, teachers, and students.

Linguists, both native speaker and foreign, have also been involved in recent DIGEBI initiatives to create "easy" language courses in the big four "majority" Mayan languages of Guatemala. These courses are recorded on CDs, which are distributed along with portable CD players to non-Maya schoolteachers teaching in the Maya area.

Another facet of training for fieldwork, for foreign and Mayan scholars, involves co-construction of knowledge about the research-host community. Since 1987 I have been running a field school in Guatemala. This school, Oxlajuj Aj, named for the day of its inception, is billed as an Intensive Kaqchikel Language and Culture course. A core cadre of Kaqchikel native-speaker scholars teaches their language to non-Maya, both foreign and national, as well as to other Maya and to Kaqchikel who do not control their heritage language. The non-Mayan participants are chosen for their knowledge of their professional discipline and Western scholarship on the Maya. Throughout the course, the Kaqchikel and their counterparts are engaged in co-learning. The Kaqchikel teach the spoken language. Some of the Kaqchikel are trained linguists and/or teachers and know how to read and write their language. For others, a challenge of the first days shared with the non-Maya is learning the official orthography. Compositions in Maya and daily grammar lessons for the non-Maya require understanding a new linguistic system; for the Kaqchikel, they require learning the jargon and tools of linguistic analysis of their language. In the cultural classes, Western conceptions of Mayan history and culture are tested against native-lived perceptions. Kaqchikel and their counterparts are paired for field research projects, topics of which are co-determined. The non-Kaqchikel learn what are considered proper and interesting questions and how to ask them. The Kaqchikel learn to structure inquiry for critical analysis and to present findings in timed oral paper format. The write-up is also a joint construction, with a director who helps participants pitch their exegeses as professional papers. Non-Kaqchikel participants come away from the experience with conversational fluency in the language, contacts in several Kaqchikel communities, and insights into how to ask properly within this realm.

Kaqchikel participants learn to describe their language and culture in technical jargon, as well as mastering techniques for teaching, writing, and dissemination.

### 6.7 Dissemination

In the past, linguistic fieldwork has often been extractive. Outside linguists would go into a community with a pre-set agenda, gather data, leave, and publish and archive the data outside the host community. This pattern is changing. Linguists now often consult with host communities and native-language scholars as to the needs and wants of the community. Projects are designed to serve both abstract intellectual ends and practical applications.

Whereas published works in the past were largely in hegemonic languages, now some reports are being provided in the local languages or, at least, a locally intelligible "world" language. In addition, native researchers are producing their own works. The ALMG has a full run of dictionaries and grammars of 21 official Mayan languages of the country, all researched and written by Mayan scholars. FLACSO (Facultad Latinoamericana de Ciencias Sociales) has published critical historical and sociological fascicles by Maya. The discourse of these publications is often counterhegemonic, as new perspectives restructure the observed data (cf. Cojti' Cuxil 1994, 1997; Racancoj 1994; Montejo 1998) New media make it possible to share information widely. Websites can be accessed by scholars and community members around the world. The University of Pennsylvania now hosts the Open Language Archives Community (OLAC). This "community" provides a base for scholars to post their field notebooks (raw data), analyses, and queries. Other databases – for example, the AusAnthrop Australian Aboriginal tribal database and the Sino-Tibetan Database and Retrieval System Project – are regionally specific. Others are language-group oriented; for example, the Comparative OnLine Bantu Dictionary and Jonathan Amith's Nahuatl Learning Environment, a text-rich site. Some universities provide language materials online, through publicly accessible portals: Tulane University's e-Kaqchikel (http://ekaq.stonecenter.tulane.edu/). Nahuatl materials prepared by Jonathan Amith and a team of Nahuatl speaker scholars can be accessed through the Nahuatl Learning Environment: www.balsas-nahuatl.org.

## 7. The bottom line

All field linguists, native-speakers or otherwise, need basic training in phonetics (transcribing below the level of the phoneme can present a challenge for native speakers, but also great insights), phonology (a prerequisite for devising practical

orthographies), morphology (especially useful in applied projects such as neologisms), and syntax (grammars quickly become emblematic of status as "real" languages). Outside linguists may need help in learning what is acceptable to ask; inside linguists may need preparation in theoretical literature to learn what academia deems are the burning issues to address. Co-constructed projects promise to give both local and academic communities the greatest satisfaction in goals, products, applications, and dissemination.