DESIGNING "COMMUNITY-TECH" WORKFLOWS: A FIELD LINGUISTS' GUIDE TO PUTTING GOOD PRACTICE LANGUAGE TECHNOLOGY INTO THE HANDS OF SPEAKERS

By

Andrea Berez & Gary Holton

Paper presented at

Designing "community-tech" workflows:
A field linguists' guide to putting good practice language technology into the hands of speakers[1]

Andrea Berez
Wayne State University
andrea@linguistlist.org
andrea.berez@gmail.com

Gary Holton
Alaska Native Language Center
University of Alaska Fairbanks
gary.holton@uaf.edu

1   Introduction

As we write this, we are coming to the end of a three-year language technology project focused on the Dena'ina Athabascan language of Southcentral Alaska. Known as Dena'ina Archiving, Training, and Access (DATA), this project was developed in order to address three primary concerns: (1) requests by the Dena'ina community for greater access to existing language documentation materials; (2) the need to implement enduring standards of digital language archiving; and (3) the need for intensive training of community members and graduate students in language technology and fieldwork techniques.

Dena'ina is a severely endangered language, spoken by less than 75 people in the Cook Inlet region of Alaska (Figure 1). While several nascent maintenance efforts are underway, English remains the dominant language of the Dena'ina community, and children are no longer acquiring Dena'ina as a first language. However, the current vitality of the language is in sharp contrast with the depth and breadth of existing documentation. While there is no comprehensive dictionary or grammatical description for the language, Dena'ina is among the better-documented of the Northern Athabascan languages. The ANLC Archive contains documentation going back to a wordlist collected in 1778 by William Anderson during Captain Cook's voyage to Alaska, and Dena'ina has received considerable attention since the early 1970s, particularly through the work of linguists James Kari and Joan Tenenbaum and Native speakers Peter Kalifornsky and Albert Wassillie (cf. Berez & Holton 2005). Among the materials at the ANLC Archive are extensive fieldnotes, grammatical descriptions, narratives, ethnographic information, pedagogical materials, and hundreds of hours of audio recordings. Unfortunately, the ANLC Archive is located in Fairbanks, far from the Dena'ina region (see Figure 1 inset), so access to these items remains difficult for Dena'ina heritage speakers and linguists worldwide. Furthermore, prior to the beginning of the DATA project, Dena'ina resources in the ANLC Archive had not been extensively cataloged, and lack of attention to preservation rendered paper and audio recordings at risk to long-term degradation.

Figure 1: Dena'ina Language Area (Fairbanks is indicated with a red star in the inset). Courtesy of Lake Clark National Park and Preserve and Barbara Bundy; text by James Kari.

To address these issues, the DATA project created an Online Archive of the Dena'ina language materials at the ANLC, using the technological expertise of the LINGUIST List and the recommendations for archiving standards of groups like E-MELD and OLAC. At the moment, the Archive contains metadata for more than 500 entries, all of which are searchable by a range of fields, including title, date, dialect, content type, and participants. The Archive is built on an Oracle database with a ColdFusion interface (see Holton et al, forthcoming, for a thorough description). Most of the resources are available for download in digital format. The Online Archive is also designed to protect intellectual property rights by permitting restrictions to be placed on access to materials as deemed sensitive to the Dena'ina community. The access system is flexible can easily be adapted to changing community consensus.

The Online Archive represents the first *A* in the acronym DATA: archiving. The second *A* stands for access, which is covered by the Qenaga.org website (the word *Qenaga* means 'language' in Dena'ina). The website acts as an online portal to Dena'ina language and community information. Here, users can learn of local news and goings-on in the Dena'ina community, as well as access resources for learning Dena'ina language. The site contains information on grammar, phonology, verb structure, music, literature and history, as well as instructional modules and suggestions for self-study. The web portal provides a kind of "value-

added" product, synthesizing existing archival documentation in order to make it more accessible to users. The web portal also supplements existing documentation by providing access to "known" but undocumented information. For example, the portal contains short biographical descriptions of people who have worked with the Dena'ina language, including linguists, educators, and speakers. This information is "known" by linguists and others who have worked with the language, but it has for the most part gone unrecorded.

We put the *T* in DATA by including a significant training component in the project. In the summer of 2005 we developed and taught a course in language technology at the Dena'ina Language Institute, a three-week intensive summer language program in Soldotna, Alaska.[2] The course covered basic HTML programming and audio recording and digitization. Students were all Dena'ina heritage speakers and friends of the community, and their ages and levels of experience varied widely. By the end of the three-week class, each student produced two web pages about Dena'ina language or culture, one of which incorporated audio clips of Dena'ina language that they had collected from elders during the Institute.

But our training goals for DATA as a whole reached beyond the limits of the summer class, and we worked to provide more in-depth instruction to interested students. One of the language products resulting from DATA is a CD-ROM containing nineteen Dena'ina narratives with time-aligned audio and text. The audio comes mostly from legacy audio recordings, some nearly forty years old. These recordings had been digitized as part of the archiving component of the DATA project, and those recordings are available in the Online Archive. However, the linking of recordings and their translations had never before been available to the Dena'ina community at large.

The nineteen stories included on the CD-ROM do not begin to scratch the surface of the supply of legacy narrative recordings available in Dena'ina. Given the positive community response to the collection, their nearly constant quest for new language materials, the eagerness of some students to broaden their technology skills, and the authors' commitment to good practice technology infrastructure, we decided to train community members to create similar products in the future. In the remainder of this paper we describe this training process and reflect on our experience with community technology training.

2   What is a "community-tech" workflow, and what good is it?

By "community-tech workflow" we mean a procedure or set of procedures which can be taught to community members, who can then use it to create new language technology products on their own. The workflow is meant to supplement, and maybe eventually replace, the situation in which the technologically proficient linguist supplies digital language materials to the community without direct community involvement in the technological aspects of product development. Hence in the case of the DATA project, our community-tech workflow is the set of steps a Dena'ina community member would undertake to create a CD-ROM similar to the collection of narratives we produced ourselves.

2.1   Motivation

But why bother training community members, when we were perfectly capable of producing the story CD ourselves and handing the finished product over to the community? In truth we did

---

[2] ANL 295: Technology for Alaska Native Languages was offered as a University of Alaska Fairbanks course at Kenai Peninsula College by instructors Andrea Berez and Sadie Williams.

make the CD, but we first needed to train *ourselves* to do it. Before we started, we had very little knowledge of the software and programming languages we eventually used, and selecting among the many available options required careful consideration of each in learnability, ease of use, long-term viability, and the like. But as we worked out the bugs—a process that took several months, during which we relied heavily on our professional contacts in the language technology world for advice—we realized that it might be a good idea to put as much attention into the process itself as we were putting into the finished product. We knew that we were going to be available to do the work in the short term, as the grant had a time frame of three years, but it seemed a shame to waste all that investigation and self-training. By focusing on streamlining the *process* of making the CD, according to principles we outline here, we were able to turn our work into a community-tech workflow.

Certainly, developing such a workflow is a larger undertaking than simply making the product and handing it over to the community, so the potential benefits of doing so need to be made clear. We have identified several potential benefits, including the following.

- Often, the person who makes language materials gains as much benefit, in terms of language learning, as the consumers who use the materials later. We learned a great deal of Dena'ina language simply from manipulating texts and recordings.
- Teaching and learning materials of all kinds are very desired. A class in curriculum and materials development is offered at the summer Dena'ina Language Institute, and at the weekend-long Dena'ina Language Workshop in Nondalton in November 2005, an entire evening was devoted to sharing new pedagogical materials.
- For severely endangered languages, many aspects of language work are in a race against the clock. Many hands make light work, and an army of tech experts can do more than just a few.
- Community members have access to documenting authentic, spontaneous language-use contexts and events that outsiders do not have (cf. Villa 2002).

Most importantly, perhaps, community-tech training keeps financial and intellectual resources inside the community. Often, expensive technology work is contracted to non-community members great expense. As Hinton (2001) cautions:

> "(a)n easy pitfall is for a language program to get a grant to buy computers and hire a consultant to make some CD-ROMs or programs for the language—and then be left without the money or expertise to upgrade or make repairs or transfer the program or CD-ROM to the next stage of technology." (268)

We were nearly witness to such an experience. During the course of the DATA project a local tribal organization received a grant that included $85,000 budgeted as a subcontract for digitizing audio recordings. Through the DATA project we were able to provide training in digitization to community members who were then able to do the digitizing themselves, thereby allowing that budgeted funds to be spent on community members while at the same time developing technological expertise within the tribal staff.

To invoke an adage, a good workflow is like teaching a person to fish, to become self-sufficient in the use of technology as a tool for language revitalization. Villa (2002) also advocates technology training as a response to the ills of cultural mining:

"Certain researchers have entered indigenous groups for study, taken the data they sought, and then left to publish their research without returning any information to the people with whom they worked. As (a) result … many indigenous groups in the US closely control who is allowed to enter the community to carry out work on language…. One solution to this problem is to train members of the language group in the use of advanced technology in order that materials for language preservation and teaching can be collected, archived, and prepared by in-group members for other in-group members." (93-94)

Thus, the benefits of technology training are not limited to improvements in the quality of products produced during or shortly after the training period. In a larger sense technology training can bring ownership to language documentation and maintenance efforts within the community.

2.2 Goals

In designing the technology workflow we established a number of goals. These can be broken down into general "macrogoals" and more specific "microgoals". The general goals would likely apply to any language technology project. They are the methodological goals that remain constant from application to application. On the other hand, the microgoals are specific to this particular project and help define the function of the final product. We conceived of our macrogoals as an overarching list, to be adapted to most community-tech situations, independent of the nature of the final product (that is, these goals would still apply if the finished product were something other than an interactive CD of narratives, such as a dictionary or language-related game).

Table 1: Community-tech macrogoals

1. adhere best practice procedures for data preservation
2. be easy enough to be learned in a short time by a motivated student
3. significant training component
4. include instructions for getting assistance if a problem occurs
5. require only free or inexpensive software
6. bug-free fully functioning software
7. easy installation of final product

These goals are motivated by several factors. The first macrogoal reflects our desire for materials to be of lasting value. Even if the materials are not intended to be of documentary value, we cannot ignore best practice. We see at least two reasons for this. First, even when documentation is not the intent, language maintenance materials may end up being the best record of a language. For example, some the best documentation for Han, one of the most endangered languages of Alaska, is arguably in the form of a series of pedagogical recordings which were never intended to be of documentary value. Second, without attention to best practice, the resulting products may not even serve their immediate pedagogical purpose. This tends to occur especially when products have long development times, so that the underlying technology is nearly obsolete when development is complete. Examples from Alaska include many products produces using Hypercard and Macromedia Director.

The remaining macrogoals are motivated primarily by our experiences with teaching language technology. The technology must be intuitive and easy to learn, and users must have

access to help resources. Training must be incorporated into the workflow so that communities members acquire sufficient expertise to continue development and maintenance of the product. The final two macrogoals on the list are apparently in conflict and must be balanced against each other: while it is important that inexpensive tools be made available for community work, such tools frequently contain more bugs and are prone to crashing.

As a final macrogoal, we wanted any resulting product to require very little configuration of home or office computers to use the CD. Through experience with previous HTML-based projects intended for community use we have found the tolerance for installation difficulties to be extremely low. Many of the installation difficulties have centered around the use of audio plug-ins within web browsers. Indeed, after a preliminary beta distribution of the Dena'ina CD in 2005 we received numerous requests to hand-configure users' machines to run the required browser and audio plug-in combination. Similarly, a Dena'ina multimedia phrasebook produced recently at ANLC (Balluta & Evanoff 2004) has not been able to be accessed by all potential users because it requires the Quicktime audio plug-in with a specific web browser. Even though detailed and accurate instructions for installing and configuring the plugin with the requisite browser are included on the CD, many users have proven reluctant or unable to complete the installation and configuration progress. A similar experience has been witnessed with trial versions of a Tanacross multimedia CD which includes Quicktime video clips (Holton & Thoman 2006). The CD automatically attempts to install and configure the Quicktime plugin if it is not already present. However, many are reluctant to click "OK" on the installation confirmation dialog.

While these macrogoals were developed with our specific project in mind, the underlying principles will also be relevant in other kinds of language technology training, where the outcome is something other than multimedia product. For instance, if speakers are trained to do, say, morphemic analysis on video segments of conversations using some kind of software tool, these macrogoals will still apply. We also developed a number of project-specific microgoals which reflect our implementation of the macrogoals. These may or may not apply more generally to other language projects. For example, in order to implement the best practice macrogoal we chose to employ an XML structure and to include metadata for each narrative. However, in order to meet the goal of easy installation we made a goal of including support for non-unicode enabled machines—a choice which is apparently in conflict with best practice. This choice reflects our experience with browsers and machines that still fail to display some Unicode characters correctly. For example, the Polish hook or nasal hook, glyph 808, is properly represented by few browser and machine combinations other than Safari on Mac OS X.

3   Developing the workflow

In this section, we describe the process of developing the workflow. This was largely a process of self-training, in which we investigated possible options for converting legacy materials into the finished presentation product.

To select the stories for inclusion in the collection, we searched through the existing archival materials to find matching pairs of transcriptions and audio recordings. Some transcriptions had been previously published without accompanying audio, while others were part of the archive as manuscripts. However, existing metadata were generally not rich enough to aid significantly in the matching of text and audio. Audio recordings and transcriptions were archived as separate resources. In addition, the audio recordings for the most part did not exist as separate resources but as part of a larger field recording which might contain several stories in

addition to elicitation and other field work data. We began with legacy transcriptions of several different types. Still other audio-transcription pairs were in the collection of linguist James Kari. In one case matching text and audio had been previously published and was available to us (Johnson 2005).

Kari provided transcriptions as either handwritten documents or as WordPerfect files. In the case of the previously published stories and the handwritten transcriptions, we typed them into a text editor and passed them back to Dr. Kari for proofreading. In the case of the WordPerfect files, we simply opened them in the text editor and checked for character conversion problems. Dena'ina has only two non-Latin characters in its orthography, Ł and ł, which in pre-Unicode days were typed with the backslash symbol (\). This is fortunate for us, because we discovered no real character conversion problems that can plague older files in other languages. We quickly cleaned up some small formatting issues, and the new text file became our master copy of the transcription.

Most of the audio recordings, with the exception of a few recent digital recordings, were digitized from original cassette or reel-to-reel. The process of selecting audio recordings thus involved editing large audio files. Generally, we tried to select audio that was free of background noise, but there is a significant hiss in at least one very old recording. We did not attempt to remove this noise.

At this point, we were ready to begin creating a time-aligned, interlinearized version of the story. We already knew that the ELAN software could be used to create time-aligned XML files in accordance with best practice archiving recommendations.[3] After some experimentation, we found the easiest approach to alignment turned out to be from the back-end: we first created the time divisions in the Elan GUI, then we opened the ELAN file in a text editor and cut-and-pasted the Dena'ina and English transcriptions directly into the ELAN XML. Our alignments contained only two tiers, one for Dena'ina and one for an English free translation, but the back-end method would work for any level of interlinearization, provided the user takes care in navigating through the XML structure. This method is generally easier than pasting text directly into the ELAN interface.

Once we had our annotation file in hand, it was time to turn our attention to presentation. We decided to create pages in HTML in order to be consistent with our macrogoal of making the product readily available on home computers. All modern computers have a web browser installed, and most people, even those living in the more remote villages, know how to use them. In order to transform the ELAN XML into HTML we attempted to use an XSLT transformation. As neither of us was familiar with XSLT, we relied on books and our colleagues at the LINGUIST List, and managed to teach ourselves enough XSLT to create a passable, though not elegant, stylesheet. This stylesheet essentially iterates through ELAN tiers and transforms the result into paired Dena'ina and English text with accompanying Flash file (see Figure 2). Notably, this stylesheet does not manipulate the ELAN timecodes. Note that we also needed to add a little more markup to the ELAN XML by hand, in particular, the <DENTITLE>, <ENGTITLE>, and <SPEAKER> tags. We then used the LINGUIST List's ColdFusion server to perform the transformation, which met one of our macrogoals by necessitating the upload of the .eaf to the Qenaga.org Online Archive.

---

[3] http://www.mpi.nl/tools/elan.html

Figure 2: Excerpt of the stylesheet

```
<tr align="left">
   <td width="50">
                <object height="40" width="40">
                <param value="audio/audiofile-1.swf" name="movie" />
                <embed height="40" width="40" src="audio/audiofile-1.swf" />
                </object>
             </td>
<td><span class="den">
   <xsl:value-of select="//TIER[@TIER_ID='Denaina_text']/
ANNOTATION/ALIGNABLE_ANNOTATION[@ANNOTATION_ID='a1']/ANNOTATION_VALUE"/>
</span>
<br/><span class="eng">
    <xsl:value-of select="//TIER[@TIER_ID='English_Free_Translation']/
ANNOTATION/REF_ANNOTATION[@ANNOTATION_REF='a1']/ANNOTATION_VALUE"/>
</span>
 </td>
        </tr>
        <tr align="left">
            <td><br/>
            </td>
        </tr>
```

Note that this stylesheet does not output immediately useable HTML. That is, the resulting output must be edited further before it can be used. In addition, this stylesheet deals only with the visual portion: it does not deal with the audio playback.

In order to decide upon the best method for presenting the audio, we investigated several different audio presentation technologies, including SMIL, QuickTime, and Flash. Our microgoals for the project included a choice between uninterrupted audio and line-by-line clips. Offering clips proved to be the most challenging aspect of the entire project, one that took us several months, and two versions of the CD-ROM, to solve. Early in the project, we had no easy way of breaking the full audio file into smaller chunks, and cutting audio by hand would have proved too costly in terms of time and the potential for mistakes. Our first solution came in the form of QuickTime embedded media players. QuickTime recognizes <STARTTIME> and <ENDTIME> tags, and so can be programmed to play only a section of the entire file. This method has two advantages: there is no need to cut the audio into smaller files, and memory space on the CD is saved by including only one audio file for each story.

However, our QuickTime solution also had a number of significant disadvantages. First, we had to copy the timecodes from the ELAN file header into the HTML by hand for each line of text. This probably could have been done using XSLT but was beyond our limited expertise. The second disadvantage is that while Apple uses QuickTime as the default media player plug-in for their machines, PCs do not. Because most of the PC default players do not recognize the <STARTTIME> and <ENDTIME> tags, PC users would be forced to download QuickTime and then configure their computers to use it. As described above, community testing showed this to simply be too much of a hassle for most users to bother with, even when we provided hands-on instruction. The final disadvantage of QuickTime is that numerous players in one page will often crash the computer, or may simply fail to load. In short, QuickTime, as used here, is neither very programmer-friendly nor very user-friendly.

A better solution came in the fall of 2005, when Aric Bills, a graduate student at the University of Alaska Fairbanks, made a small piece of software with big potential for Elan users. Using Tcl, he created an "audio chopper" that reads the timecodes from the ELAN file header and creates audio clips with unique file names with the push of a button. Now we were free to investigate Flash as a method for playing audio. We created Flash buttons for each line of audio and embedded those into the HTML.

This solved several problems at once. First, it was fairly easy for us to modify the XSLT to include the unique audio filenames, which meant much less massaging of the HTML output. Second, Flash has none of the user-friendliness problems that QuickTime has. According to Jeffries (2002), 73-98% of computers can view Flash with no additional configuration, and recent figures from Macromedia (2006) claim 97% penetration of Flash Player as of December 2005. Pages with multiple Flash buttons are not nearly as buggy or susceptible to browser crashes as those containing multiple QuickTime players.

Another microgoal was to include support for non-Unicode-enabled machines, which we accomplished by creating a parallel set of story pages wherein all the *l*'s and *Ł*'s are replaced with a backslash. The navigation within the parallel pages is designed to be as unobtrusive as possible. Once the user has navigated to a page containing backslashes, all the linked pages also contain backslashes, eliminating the need to switch back and forth for each story. Of course, the user can chose to return to the pages containing the Unicode characters at any time. Finally, we added metadata for each story to a "credits" page.

In the end we were able to meet most of our macrogoals and many of our project-specific microgoals. Our level of success in reaching our stated technology goals for the workflow is summarized in Table 2 below.

Table 2: Achievement of workflow goals

Goal	Achieved?	Comments
Easy to learn	maybe	With sufficient student motivation. Some parts easier/more appealing than others.
Training	maybe	Attempted but perhaps not successful.
Bug-free	yes	As of now, yes. Future bugs difficult to predict.
Free or inexpensive software	no>yes	Flash is expensive, but may be available through a university license; Quicktime is free.
Able to contact linguist later	yes	Linguists are in close contact with the community and readily available by email.
Audio available whole or clips	yes	
Built upon Elan XML	yes	
Metadata included	yes	However, metadata is not integrated into archival version.
Non-Unicode support	yes	
Little configuration of computer to use	yes	Except computers without Flash Player (a small number).

As this table reveals, training was perhaps the most significant area of underachievement. Undeniably, community training has been one of the most challenging aspects of the DATA

project.[4] As noted previously, the three-week technology class at the 2005 Dena'ina Language Institute successfully met the goals stated in the syllabus, and each student who participated learned something new. However, as technology training was only one of three components of the DATA project, there was simply not enough scheduled contact time to address our advanced training concerns. Most of our time during the three years of the project was spent developing the Online Archive and the Qenaga.org website from our offices in Fairbanks and Michigan.

Providing advanced training was further complicated by difficulty in locating and recruiting interested community members. Only a few students in the 2005 summer class undertook projects with the intent to build on them later on. One student created an interactive map of the Lake Clark area, with clickable hotspots linked to Dena'ina audio of the placenames. The student has continued to develop this project in her spare time during the last year. In addition, her experience with this project has helped her to contribute in a significant way to development of GIS placenames database for Lake Clark National Park by the National Park Service.

The two most successful students from training class (who, interestingly, are two non-Native men with close ties to the Dena'ina community) have secured more than $20,000 from the University of Alaska and Kenaitze Indian Tribe to upgrade their equipment and produce digital products on the Kenai dialect of Dena'ina. One of these products is an interactive CD-ROM suitable for internet publication. Linked pages cover Kenai Dena'ina phonology, vocabulary, grammar, verb structure, narratives, and ethnogeography of Dena'ina territory.

Our continued assistance to these students has taken the form of emails and a few extra hours in front of a computer, but at the time of writing, no additional formal training has taken place.[5]

These students also have plans to produce a collection of narratives similar to the one produced during the DATA project, but to our knowledge they have not yet started work on it. Whether they will adopt the workflow discussed in this paper, in whole or in part, is questionable. It will be interesting to see which aspects of the workflow are attractive to them.

4 Balancing goals: the BP cline

For the Dena'ina stories project, we were largely able to meet the technology workflow goals we established. However, in the end we realized that some of our goals were actually in conflict, causing a kind of cline of best practice to emerge within our workflow. We now believe that such a "BP cline" may be inherent in the community tech workflow more generally. Best practice goals are easily violated in favor of learnability and user-friendliness goals, and vice versa. For example, we chopped the audio and used Flash because we wanted an end product that could be easily installed and was bug-free. However, the pathway from the archival time-aligned XML version and the Flash-enabled end product is rather non-intuitive, violating our workflow goal of easy learnability. Additionally, Flash software is expensive, violating our goal to use only cheap software.

Another example of the BP cline is the issue of transforming the XML into HTML. Originally we used the LINGUIST List ColdFusion server to perform the transformation, which

---

[4] In contrast, the graduate assistant training was highly successful, providing two years of training to two students and limited training to two other students. None of these students were from the Dena'ina community, in spite of extensive efforts to recruit Dena'ina students.

[5] We have plans to conduct additional advanced training at the 2006 Dena'ina Language Institute in early June. You'll have to wait until the E-MELD conference in late June to hear about it!

had the advantage of forcing the storage of the alignment file in a secure location. However, it is also possible to perform the transformation in XML editing software on a local machine. While doing the latter is certainly quicker and less of a hassle than uploading the file to the server, the archival-quality file is not automatically stored in a secure location.

While all community trainees believe intrinsically in the value of best practice as it applies to language technology, in reality best practice goals are often traded for other more immediate or more practical goals. The hardest thing for trainees to accept is the value of ensuring the longevity of their work, preferring instead to focus on only those stages of the workflow that directly result in an attractive final product. People are very interested in learning HTML, but they would prefer to skip the Elan alignment altogether in favor of just transcribing directly in the source code. While many of our students have now become proficient in editing HTML pages, none has yet mastered the art of massaging the XML representation into HTML. Thus it seems that at several points in the workflow, two paths are available: an easy way, and a best practice way. Given the choice, people will very often choose the easy way.

This failure is at least in part owing to a shortcoming of our workflow design. We failed to develop (or locate) an appropriate tool for managing the conversion to presentation format, and we failed to develop (or locate) an audio playback tool which didn't rely on expensive proprietary software. While better solutions could surely have been found for this project, similar problems are bound to arise in future projects. In any case, the stated goal of the DATA project was not to develop new tools but instead to investigate a subset of the existing tools to see how they can be applied to Dena'ina. As a case study, our experiences with the Dena'ina narrative CD may indicate that the field needs better tools for making presentation products.

In fact, in order to better understand the shortcomings of our training program, we need to revisit not only our stated technology workflow goals, but also our unstated ideologies regarding those goals. From the outset, we assumed best practice to be an integral part of the community technology workflow, incorporating this concept as our first workflow goal. As linguists who have been involved with digital language documentation and archiving, the need for best practice seemed not only obvious, but universal. For us—as indeed for many E-MELD participants—best practice had become a creed. In incorporating BP into the community workflow we felt we were participating in a form of technology repatriation, responding to Villa's (2002) admonition to provide training in advanced technology. And yet in doing so we may have inadvertently engaged in another form of cultural hegemony, imposing the culture of best practice upon a community that may not want it. Indeed, the Dena'ina community may prefer to develop its own indigenous sense of best practice.

Whether or not this is the case, it is worth re-examining the locus of best practice within the technology workflow. This is not an issue of whether the community cares about best practice. Of course they care. We all want language materials to be preserved and to be of lasting value. But to care about best practice is not necessarily to care about implementing it. A more effective community-tech workflow might strive to assign tasks in a manner which takes best advantage of relevant skills. In small communities with limited resources it may not be unreasonable to have outside linguists assist with best practice preservation while community members focus efforts on designing and implementing presentation formats.

5 Conclusion

It's always wonderful when by working together, two parties can help one another to meet their respective goals. Whether it is necessary—or even possible—for each party to actually *adopt* the

goals of the other as its own is unclear. As field linguists, we cannot expect communities to commit to implementing best practice recommendations for long-term data preservation as eagerly as we do (as we know, it is hard to convince all linguists to commit to such a task).

It's probably fine, even a good idea, to help communities become self-sufficient in language technology when they request it. But they may not approach language technology with the same goals we have; indeed, they may not use their new knowledge in ways we expect. If a community requests training, it would of course be unfair of us to not share our skills, or to not take the time to help as much as possible. But we need to be sure to serve the needs of the community. Asking them to serve our BP needs will likely lead to frustration for everyone.

Any best practice component of a community-tech workflow should be as unobtrusive as possible. There is a tenuous line between ensuring long-term data preservation and creating a workflow that is too cumbersome to be attractive. The last thing we want to do is create a roadblock between the community and their language goals (cf. Dauenhauer & Dauenhauer 1998). In the case of the DATA project, the responsibility for BP fell primarily on the linguists rather than on the community. This may or may not prove true for other projects. As linguists, we cannot be disappointed or frustrated if community members are not interested in best practice. They simply may not be.

Then again, they may be. In the end the goals of linguists and the community are not really all that different. Both are interested in creating materials which both document the language and aid in language revitalization. And both linguists and community members are interested in ensuring the long-term preservation and accessibility of those materials. No one wants to devote hundreds of hours to developing CD-ROMs which will be inaccessible in just a few years. To the extent that workflows designed with best practice in mind will help ensure the long-term viability of language materials, both linguists and community members will surely benefit from employing such workflows. Thus, while the results of the workflow developed for the DATA project are not yet encouraging, we remain optimistic that future initiatives to develop community tech workflows will be successful. And we hope that others will benefit from knowledge of our experiences.

References

Berez, Andrea and Gary Holton. 2005. *Dena'ina qenaga:* Previous linguistic work. Online: retrieved February 25, 2006 from http://qenaga.org/lingwork.html.

Balluta, Alec & Gladys Evanoff. 2004. Dena'ina Qenaga Du'idnaghelnik, ed. by Olga Müller. Fairbanks: Alaska Native Language Center.

Dauenhauer, Nora Marks and Richard Dauenhauer. 1998. Technical, emotional, and ideological issues in reversing language shift: Examples from southeast Alaska. Endangered Languages: Language Loss and Community Response, ed. by Lenore A. Grenoble and Lindsay J. Whaley, 57-98. Cambridge: Cambridge University Press.

E-MELD. 2005. E-MELD School of Best Practice: Dena'ina. Online: retrieved April 18, 2006 from http://emeld.org/school/case/denaina/index.html.

Hinton, Leanne. 2001. Audio-video documentation. The Green Book of Language Revitalization in Practice, ed. by Leanne Hinton and Ken Hale, 265-271. San Diego, London: Academic Press.

Holton, Gary & Rick Thoman. 2006. The Sound of Tanacross. Fairbanks: Alaska Native Language Center.

Holton, Gary, Andrea Berez and Sadie Williams. Forthcoming. Building the Dena'ina language archive. Information Technology and Indigenous People, ed. by Laurel Dyson, Max Hendricks, and Stephen Grant. Hershey, Pennsylvania: Idea Group.

Jeffries, Andy. 2002. Percentage of Flash users. Online: retrieved April 29, 2006, from http://www.andyjeffries.co.uk/documents/flash_penetration.php.

Johnson, Walter. 2005. *Sukdu Neł Nuhtghelnek/*I'll tell you a story: Stories I recall from growing up on Iliamna Lake. Fairbanks: Alaska Native Language Center.

Macromedia. 2006. Flash Player penetration survey. Online: retrieved April 29, 2006 from http://www.macromedia.com/software/flash/survey/.

Villa, Daniel J. 2002. Integrating technology into minority language preservation and teaching efforts: An inside job. Language Learning and Technology 6(2): 92-101.