Symbiosis between documentary linguistics and linguistic pragmatics

Edward Garrett & Leah Bateman

Proceedings of Conference on

Language Documentation & Linguistic Theory

Edited by Peter K. Austin, Oliver Bond & David Nathan

This paper can be cited as:

Edward Garrett & Leah Bateman. 2007. Symbiosis between documentary linguistics and linguistic pragmatics. In Peter K. Austin, Oliver Bond & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory*. London: SOAS. pp. 83-93.

or:

Edward Garrett & Leah Bateman. 2007. Symbiosis between documentary linguistics and linguistic pragmatics. In Peter K. Austin, Oliver Bond & David Nathan (eds) *Proceedings of Conference on Language Documentation and Linguistic Theory*. London: SOAS. www.hrelp.org/eprints/ldlt_11.pdf

Symbiosis between documentary linguistics and linguistic pragmatics

EDWARD GARRETT[1] & LEAH BATEMAN[2]

Eastern Michigan University/SOAS[1] & University of Massachusetts, Amherst[2]

1. BACKGROUND[1]

As the field of documentary linguistics has grown, so too has the software infrastructure to support it. Some of the most commonly used linguistic software tools, such as Shoebox/Toolbox, Transcriber, and ELAN, were either developed with the objectives of field linguists in mind, or co-opted and given new life by field linguists. Field linguists typically use such tools to carry out core tasks of language documentation and description: orthographic and phonetic transcription of audiovisual recordings, morpheme by morpheme glossing of source language text, and time-alignment of annotations to media files, for example.

To overemphasize these typical uses, however, would give the impression that such tools are only capable of meeting the needs of field linguists. To the contrary, such tools are often architected for general use, by linguists and non-linguists alike. In this paper we examine a particular problem in Tibetan linguistic pragmatics, showing on the one hand that software originally intended for documentary linguistics can be used productively with other, more theoretical linguistic objectives in mind; on the other hand, we demonstrate that tools designed for non-linguistic purposes can and should be adapted as aids for linguistic research.

2. A PROBLEM IN TIBETAN PRAGMATICS

It is normally claimed that the Spoken Standard Tibetan direct evidential auxiliary *'dug* is used to report situations known through direct sensory perception (vision, touch, etc.) and that its indirect evidential counterpart *yod.red* is used to report situations presented as non-evidentially grounded facts. In spite of this, we discovered in the Tibetan & Himalayan Digital Library's corpus of transcribed and time-coded videos a striking example of *yod.red* being used in a question (1a) and answer (1b) even though the target activity was in full view of both the questioner and the respondent. The example is transcribed in (1), and contextualized by Figure 1.

(1)   (a)   *khong*   *ga.re*   *gnang*   *gi*   *yod.red*
            (s)he   what   do   IMPFTV   INDIRECT
            'What is she doing?'

      (b)   *khong*   *gsol.ja*   *bzo*   *gi*   *yod.red*
            (s)he   tea   make   IMPFTV   INDIRECT
            'She's making tea.'

Figure 1
Video context for (1)

The questioner in (1a) is the foreigner in Figure 1. (Pretending to be[2]) unfamiliar with Tibetan customs, he asks the Tibetan man in the middle what the woman is doing. The man responds that she is making tea (1b).

It is evident from the picture that both men can see the activity in question, so the use of the indirect auxiliary *yod.red* by both is prima facie unexpected. So, two mysteries need solving: why doesn't the foreigner use *'dug* in his question? and why doesn't the Tibetan use *'dug* in his answer?

After elicitation sessions with several native speakers, some clarity emerged. First, we learned that knowing the intent and knowledge of the questioner was critical. In (1a), the questioner perceived but was unable to interpret the scene, and his intent was to acquire knowledge. In contrast, the use of direct *'dug* in (1a) would imply that the questioner already knew the answer, and had the intent, for

---

[2] In fact, this video was scripted – it is the dialogue from Chapter 17 of Tournadre & Dorje (2003). The foreigner, moreover, is Tournadre himself, who in real life knows Tibetan language and customs extremely well. Video, audio, and transcript files for this and other titles are available for free download from the Tibetan & Himalayan Digital Library's AV Archive, http://www.thdl.org/, Title #01030.

example, of checking the hearer's understanding. In more familiar terms, *yod.red* would be used in a request for information, and *'dug* in an exam question.[3]

In the case at hand, substituting *'dug* into (1a) would have no effect on the content of the answer. Whether presented with a request for information or an exam question, a Tibetan would answer 'She is making tea' regardless, since within Tibetan culture, the activity in Figure 1 is unambiguously an instantiation of tea-making. However, we were told that if the situation were ambiguous or difficult to interpret, the form of the question (*yod.red* vs. *'dug*) would affect the content of the answer. For example, given a scene of a man tapping a table, a question with *'dug* would provoke a simple descriptive answer ('He is tapping the table'), while a question with *yod.red* would provoke a creative interpretation ('He is thinking', 'He is summoning the gods', etc.)

We could have left it at that, but we wanted systematic confirmation of these intuitions. In part, we wondered whether the behavior of metalinguistically naïve native speakers would confirm the predictions of our consultants. More importantly, however, we felt that although we had a basic understanding of the data, we really had no theoretical framework with which to explain it. There was still, for example, no explanation for why both the question and answer should be affected.

So, we decided to do a controlled experiment.

3. THE EXPERIMENT

3.1. *Experimental stimuli and design*
To further explore the differences between direct *'dug* and indirect *yod.red*, we conducted and videotaped an experiment in which 18 subjects were presented with and queried about 4 visually perceivable situations by Gen Dekyi, our Tibetan colleague. Half of the subjects were asked only *'dug* questions, and half only *yod.red* questions.

Subjects were first presented with a video of Figure 1, with the sound removed. Their attention was directed to the woman's activity. Next, they were presented with the video in Figure 2, a woman crossing herself. After that, subjects were presented with Figure 3, a bizarre ritual in which a man pushes a wall with one hand, and then another, and then the first, and then the other... The fourth and final video was Figure 4, the BSL sign for 'please'.

---

[3]It is essential to point out that this dichotomy, as well as all other generalizations made in this paper about the use of direct *'dug* and indirect *yod.red*, only apply to the context where the situation in question can be perceived by both the speaker and hearer. Other contexts may allow other possibilities.

Figure 2
Video stimulus: Christian crossing

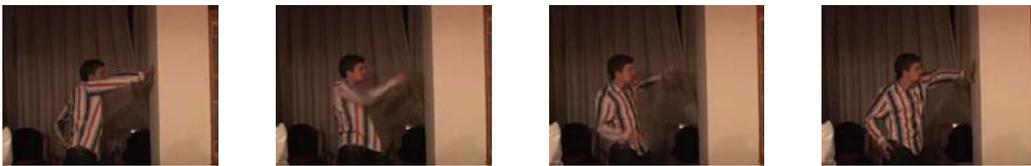**Figure 3**
Video stimulus: bizarre wall pushing

Figure 4
Video stimulus: BSL sign for 'please'

The videos were presented one after another, each set on loop, and as each video was playing the subject was asked either (2a) or (2b) by Gen Dekyi. Half of the subjects were asked (2a) throughout, and half (2b).

(2) (a) *khong*   *ga.re*   *byed*   *kyi*   *yod.red*
          (s)he   what   do   IMPFTV   INDIRECT
          'What is (s)he doing?'

     (b) *khong*   *ga.re*   *byed*   *kyi*   *'dug*
          (s)he   what   do   IMPFTV   DIRECT
          'What is (s)he doing?'

The experiment went on from there, with subjects seeing and answering several questions about images unrelated to the current investigation. In the midst of these questions, however, Gen Dekyi played a series of short BSL videos with the same signer. She watched them and then said (3):

(3) *a.las*   *khong*   *lag.brda*   *byed*   *kyi*   *yod.pa.red*
     I see   (s)he   hand signal do   IMPFTV MIRATIVE
     'Oh, I see! He's signing.'

    Finally, at the conclusion of the experiment, the subject was again shown the video in Figure 4, and asked the same question as earlier. The point of this was to see how the subject's newly acquired knowledge would affect their answer.

    The experimental scene generally looked something like Figure 5, with Gen Dekyi on the left, the subject on the right, and the computer with video stimuli in front of them both.

Figure 5
Experimental scene (subject on right)

3.2. Expectations

The videos in Figures 1-4 were selected for their diversity, with the assumption that some activities would be easier than others for Tibetans to understand. We knew that no Tibetan would fail to understand the meaning of tea-making (Figure 1), but the other situations were less clear. The bizarre wall pushing of Figure 3 was expected to be difficult to interpret (for Tibetans and Westerners alike, for that matter). The Christian crossing of Figure 2 and the signing of Figure 4 were assumed to be interpretable by those familiar with such codes, but not by others, and our experimental team felt that not all Tibetans would be familiar with these codes.

    Our hypothesis was that, as per our consultants' comments during elicitation, the form of the question (indirect *yod.red* vs. direct *'dug*) would affect the content of the answer. In other words, if a question were asked with *yod.red*, then the answer would be an interpretation of the activity, whereas if a question were asked with *'dug*, then the answer would be a visual description of the activity. We thought that the difficult to interpret cases (e.g. wall pushing), would be answered quite differently depending on the question: that, for example, a *yod.red* question

would likely elicit 'I don't know' as an answer, while a *'dug* question would elicit a simple description: 'he is pushing his hands against the wall.'

Our expectation seemed to fall in line with the evidential interpretations of the two auxiliaries. It made sense to us that *'dug* would involve description based on the senses, while *yod.red* would involve a layer of inference, given the status of the two auxiliaries as direct and indirect evidentials, respectively.

3.3. Results

Our results addressed two questions. First, as above, did the form of the question affect the content of the answer? Unexpectedly, the answer to this question was 'No' (more on this below). Second, did the form of the answer follow the form of the question? That is, we expected that *yod.red* questions would be answered with *yod.red*, and *'dug* questions with *'dug*.[4] Again, however, this turned out not to be the case.

3.3.1. Did the form of the question affect the content of the answer?

Recall our hypothesis that *yod.red* questions would be answered with interpretations of a scene, while *'dug* questions would be answered with descriptions of a scene. In fact, we found no such correlation in our results. Regardless of the form of the question, subjects always tried to characterize the scene in terms of its conventional interpretive meaning.

For example, in the case of the tea-making, 18/18 subjects answered that the woman was making tea. Everybody latched on to the broader activity of which the churning was a part, rather than the sub-activity (churning tea) before them. In the case of the Christian crossing (Figure 2), 16/17 subjects stated that the woman was praying (or some variant thereof), although there were some 'looks like' evidentials, meaning that subjects' overall degree of certainty was lower.[5]

The wall pushing scene was more difficult to interpret. The greater number of 'looks like' evidentials was again an indication of subjects' uncertainty about what was going on in the video. Still, however, whatever the question, subjects by and large attempted to explain the significance of the activity. Many said that the character was exercising; others said that he was channeling his anger.

Our subjects' responses to the BSL signing video was especially interesting. As noted earlier, the video was shown twice. The first time they saw it, subjects were not sure what to make of it. 4/15 subjects just said 'I don't know', and 6/15 used 'looks like' evidentials or modalized forms in their answer. Three of the 'I don't know' answers came from *'dug* questions, showing that the use of *'dug* in a question does not force the hearer to provide a simple visual description of the scene.

On the second go around (after subjects had been told that the man was signing), subjects responded quite differently. Only 1/16 replied 'I don't know', and 'looks like' and modalized propositions were down to 4/16. The remaining

---

[4] In obedience of the 'rule of anticipation' whereby an evidential form is selected to appear in a question in anticipation of the form that should be used in the answer (Tournadre & Dorje 2003).

[5] 'Looks like' evidentials include forms like *yod.pa.dra* and *yod.sa.red*.

11/16 speakers replied that the man was signing, regardless of the form of the question.

It seems, then, that our expectations were disconfirmed. Subjects did not take *yod.red* in the question as an invitation to speculate any more than they did *'dug*. The question 'What is (s)he doing?' was almost universally answered with the culturally constructed interpreted significance of an action, rather than a brute description of it.

3.3.2. *Did the form of the question affect the form of the answer?*
According to the rule of anticipation, the questioner uses in a question that evidential which she expects the hearer to use in his response. This is not an absolute requirement, in the sense that the hearer may always choose to phrase his answer differently, but there is certainly a tendency for question and answer pairs to be evidentially matched. In our results, however, we found patterned deviance from this tendency. Of the 37 questions asked with direct *'dug*, 24 were answered with *'dug*, and only one with *yod.red*.[6] However, of the 44 questions asked with *yod.red*, only 9 were answered with *yod.red*, while 19 were answered with *'dug*.[7] The clear pattern is that *yod.red* questions are often answered with *'dug*, but *'dug* questions are never answered with *yod.red*.

3.3.3. *Summary*
We have made the following observations (again, only considering those contexts in which the activity in question is in full view of both speaker and hearer): First, questions with indirect *yod.red* are taken as genuine requests for information, whereas questions with direct *'dug* are taken as exam questions. Second, according to our consultants, *yod.red* questions call for interpretive answers, while *'dug* questions call for descriptive answers. Third, by contrast, in our experimental setting, the evidential in the question does not determine the type of answer given (e.g. interpretive vs. descriptive). And fourth, *yod.red* questions are often answered with *'dug*, but *'dug* questions are never answered with *yod.red*.

4. DOCUMENTARY LINGUISTICS TO THE RESCUE

4.1. *Unfinished business*
In the previous section, we concluded by summarizing the four main findings of our research. Unfortunately, these findings are in conflict with each other, and despite our efforts, we have yet to come up with a satisfactory analysis of our results. One suspects that either further experimentation is required, or an

---

[6] It is probably safe to throw out the lone case of *yod.red* answering *'dug*, because it involved some distance between question and answer, and so may not constitute question-answer adjacency in the strict sense.

[7] Uncounted cases represent either 'I don't know' type answers, modal constructions such as *yod.na*, and 'looks like' evidentials.

appropriate theoretical apparatus in which to frame the results, or perhaps both. The path forward, alas, is not entirely clear.

One thing, however, is quite clear: those who found Sections 2-3 interesting and want to learn more will not be satisfied with this paper alone. Such curious souls will also want to get their hands on the data on which these sections were based. Perhaps we miscoded the data. Perhaps our experimental design was flawed. Perhaps subtle dialect issues interfered with our results. Perhaps the whole thing was made-up. Without being able to access the data, no one would ever know. Our data should therefore be stored in a digital language archive.

Moreover, those who wish to pursue aspects of this work further, including ourselves, will want to be able to explore and annotate the data, suggesting reinterpretations, future directions, and so on. It is unlikely that this work will simply and suddenly grind to a halt; therefore we need a framework by which to facilitate further research.

4.2. Documentary linguistics and data access

It is at this point that we turn to the field of documentary linguistics for assistance. The business of annotating and archiving language data has been a core concern of this field from the outset. The software tools favored by the field can help us to time-align and code our data, while other, more general tools, can be adapted for our specific purposes.

As a starting point, we note that our data is different from the kind of field linguistic data that is fetishised by some linguists: the traditional three (or more) tier IGT model including (at least) a phonetic transcription, an interlinear gloss, and a translation, which then needs aligning to a source audio or video file. What we have instead is a video for each subject, consisting of a series of experimental units (question and answers, essentially), and then assorted codes for each unit. In addition, there is a small amount of metadata for the experiment as a whole.

Creating the annotation files for each subject involved time-aligning the experimental units to the video file, transcribing the subjects' answers, and inserting coding fields such as the auxiliary used in the question—*yod.red* or *'dug*; the auxiliary used in the answer; the label for the experimental unit—'TEA', 'EXERCISE', 'PRAYING', and so on; and the subject identifier, e.g. 'Subject_E'.

For time-alignment, we employed *QuillDriver*, a tool used by the Tibetan & Himalayan Digital Library (THDL) to transcribe, translate, and time-code conversational Tibetan videos. (This was purely out of familiarity and convenience—in fact, ELAN or any other time-alignment tool could have done the job.) The result was an XML file with time-coded elements corresponding to each experimental unit.

Meanwhile, subjects' answers were transcribed into Unicode Tibetan using *Wylie Word*, an add-in for Microsoft Word, also distributed by THDL. We did this because Gen Dekyi was accustomed to using this tool, and since we had relatively little data to transcribe, we felt that merging it into our XML document later by hand would not be too onerous a task. Had the data volume been higher, however, then we would have had her transcribe directly into *QuillDriver*.

Once the transcriptions were incorporated into our time-aligned XML file, we then finished coding the data within *jEdit* (although any other XML-aware editor would have done equally well). The resulting files look something like Figure 6.[8]

Figure 6
XML ready for posting to Lucene/SOLR

```
<add>
 - <doc>
    <field name="MediaRef">Subject_F.mov</field>
    <field name="StartTime">25.631</field>
    <field name="EndTime">32.951</field>
    <field name="Subject">Subject_F</field>
    <field name="QLabel">TEA</field>
    <field name="QShortForm">ཡོད་རེད་</field>
    <field name="ALongForm">ཁོང་ལ་ཇ་བཞེ་ག་ཡོད་རེད།</field>
    <field name="AShortForm">ཡོད་རེད་</field>
    <field name="Flags"/>
  </doc>
 - <doc>
    <field name="MediaRef">Subject_F.mov</field>
    <field name="StartTime">34.295</field>
    <field name="EndTime">42.836</field>
    <field name="Subject">Subject_F</field>
    <field name="QLabel">PRAYING</field>
    <field name="QShortForm">ཡོད་རེད་</field>
    <field name="ALongForm">ཁོང་སྨོན་ལམ་འདེབས་ཀྱི་ཡ།</field>
    <field name="AShortForm">འདུག</field>
    <field name="Flags"/>
  </doc>
```

We can see from Figure 6 that our experimental units correspond to <doc> elements in our XML; and that each <doc> element contains one <field> element per coding field. The second <doc>, for example, shows a question that was asked with *yod.red* (QShortForm) but answered with *'dug* (AShortForm).

The XML data structure in Figure 6 follows the 'document-field' model of *Lucene*, an open-source Java search library. Lucene is a powerful toolkit for efficiently indexing and searching free text. To use Lucene, one first constructs an

---

[8] Note that Tibetan's non-spacing vowels and vertical stacks are being improperly spaced. in Figure 6. Unfortunately, support for Unicode Tibetan display and input still varies widely depending on the platform and application.

index, then adds 'documents' to it. Once an index is populated with documents, then users can perform complex queries against it, referencing data within specific fields or across multiple fields.

Rather than interacting directly with the Lucene application programming interface, we post our documents to *SOLR*, a enterprise search server built on top of Lucene, which facilitates and parameterizes interaction with Lucene. While direct use of Lucene would require knowledge of the Java programming language, users with a basic knowledge of XML and related technologies can get up and running with SOLR relatively easily. (This means technically savvy linguists.)

The Lucene/SOLR tandem has certain advantages over traditional database technology. First, Lucene is optimized for free text search. It efficiently indexes and analyzes globs of text, and is widely reputed for its search speed. Second, and of particular relevance to linguists, it supports language-specific text analysis, including user customized extensions. Built-in analyzers can perform advanced text analysis for major European languages, including lowercasing, stemming, stop-word removal, and synonym injection. In our work, we have developed a Tibetan text analyzer that tokenizes Tibetan orthography into Tibetan syllables and punctuation units, and is sensitive to various orthographic equivalencies and peculiarities. In future, we plan to enhance this analyzer with additional features, including word segmentation and part-of-speech tagging.

Lucene indices can be modified and updated as well, so the ability to tag and annotate existing data comes for free. New fields can be added on the fly, and a single index can contain disparate document types. Overall, Lucene/SOLR offer considerable flexibility for indexing and searching language data.

5. CONCLUSION

Conceiving documentary linguistics broadly, there is no reason to limit the focus of the field to the 'grammar, dictionary, and texts' model, as Woodbury (2003) points out. A key aspect of documentary linguistics is its focus on holding linguists accountable. The field's emphasis on digitally archiving its results, whether in the form of thinly or thickly analyzed data, implores linguists to publish not only derivative theories and works, but also the data they started with—recordings, transcriptions, field notes, and so on. All of this may be useful for current and future researchers.

Indeed, the tools and ethic of documentary linguistics have been as useful to us, in our investigation of Tibetan linguistic pragmatics, as they would have been to a more traditional field linguistic pursuit. Our investigation yielded some results, but these results remain contradictory and incomplete, and so it is incumbent upon us to make the data available to others, for further evaluation and

development. In the process, we have suggested adding the Lucene/SOLR search tandem to the documentary linguist's toolkit.[9]

REFERENCES

Tournadre, Nicolas & Sangda Dorje. 2003. *Manual of Standard Tibetan*. Ithaca, NY: Snow Lion Publications.

Woodbury, Anthony C. 2003. Defining documentary linguistics. In Peter Austin (ed.) *Language Documentation and Description, Volume 1*. London: SOAS.

---

[9] We expect to find a permanent and stable site for our data within the next several months. We are currently discussing the possibility of creating an archive of experiments on Tibetan language, including both adult and child language experiments, based at Smith College in the U.S.A.