International Journal of Lexicography
doi:10.1093/ijl/ecr018

AN ONLINE DICTIONARY WITH TEXTS AND PEDAGOGICAL TOOLS: THE YUROK LANGUAGE PROJECT AT BERKELEY

Andrew Garrett: *University of California, Berkeley (garrett@berkeley.edu)*

Abstract

In this paper, I report on an online dictionary project for a highly endangered Native language of North America. This project involves a dynamic lexicon, linked to a corpus of texts and enriched with several associated tools, which is designed to be useful for (and is regularly used by) scholars, language teachers, and language learners. The interest of this project for a broader audience emerges not from lexicographic innovations as such, but from how texts and lexicon are combined and how the interests of diverse user communities are addressed.

1. Background

1.1 *Language context*

Yurok is spoken in northwestern California, near the Oregon border; it is one of three distantly related branches of the Algic language family. One of the other two branches also consists of a single language, Wiyot, formerly spoken just to the south along the Pacific coast; Teeter (1964) presents an overview (the last first-language speaker died in the 1960s). The third Algic branch comprises the widely dispersed Algonquian language family of central and eastern North America; for a brief survey see Mithun (1999: 327-340).

Like most indigenous languages of California, Yurok was spoken in pre-contact times by relatively few people in a small territory; there were about 2,500 Yurok people occupying 80 miles along the lower Klamath River and the Pacific coast (Kroeber 1925: 17). In 2011 there are still a few elderly first-language speakers of Yurok, but it is no longer easy for them to play an active role in language teaching. (At the other end of the North American spectrum, compare Navajo, with over 150,000 speakers in Arizona and New Mexico, or the Algonquian language Ojibwe, with some 50,000

speakers in over three Canadian provinces and five US states.) For this reason, sound recordings and other linguistic documentation created over many decades are essential resources for language learners, teachers, curriculum designers, and academic researchers. At the University of California, Berkeley, the Yurok Language Project (YLP) has created a corpus of such documentary materials, and makes it available to users of an online lexicon at linguistics.berkeley.edu/~yurok.

To understand the structure and purposes of the YLP lexicon, it is helpful to know about the language's external circumstances—in particular, its history of documentation and the current situation of community members. Even among the hundreds or thousands of endangered languages in the world today, these circumstances are somewhat distinctive.

First, Yurok has a substantial record of previous documentation. Strictly speaking, this began in the nineteenth century with vocabulary and sentences taken down by explorers, settlers, and US government ethnographers, but serious linguistic research began in 1901. In that year the anthropologist A. L. Kroeber, a student of Franz Boas, came to the University of California and initiated a program of ethnographic and linguistic documentation of California's indigenous languages. California had some 80-100 indigenous languages belonging to over 20 distinct families; most were still spoken in thriving speech communities as late as 1850, and almost all were still spoken by communities or at least individuals when Kroeber began his research. As it happens, Yurok was the California language on which he worked most intensively. Thus, at a time when everyone in the community still spoke Yurok, and elderly monolingual speakers had grown up prior to White contact (which was around 1850 along the Klamath River), Kroeber was able to document traditional speech genres (ceremonies, narrative, song, usufruct) and record extensive information about grammar and language usage (including register and style). In the first half of the twentieth century J. P. Harrington and Kroeber's Berkeley colleague T. T. Waterman also recorded ethnogeographical information in Yurok, and in the mid-twentieth century the linguists Edward Sapir and R. H. Robins did significant grammatical documentation. Most of these scholars published important publications on Yurok (Waterman 1920, Robins 1958, Sapir 2001), but the bulk of the documentary material remains unpublished, in archives in Berkeley, Philadelphia, and Washington. In more recent decades, especially in the 1980s and 2000s, linguistic documentation with a more modern orientation has enriched the corpus further, as has community-based material prepared by native speakers working with the tribal language program (Exline n.d., Trull 2003).

Second, though like all Native communities in the US the Yurok community is economically disadvantaged, with very high unemployment and associated health and social problems, it has a good local technological infrastructure. Schools and tribal offices have computers and broadband internet connections,

and computer literacy is common; Humboldt State University, a part of the California State University system, is located in the area and is quite involved in Native education and cultural projects.

These facts, combined with the very small number of remaining first-language speakers of Yurok, make it both desirable and feasible to make use of earlier material in forms that can be distributed online in collaboration with language teaching programs. The Yurok Tribe has an active language revitalization program, situated administratively in its Education Department. Language classes are available in most area schools, from preschool through secondary school, and there are informal evening adult classes as well. Members of the Yurok community are very motivated to learn their heritage language, with hundreds of people having acquired good basic vocabulary skills and some rudimentary conversational ability. But as the fluent first-language speakers are elderly, those who teach the language in schools are themselves also learners; they need access to information about vocabulary, usage, and pronunciation. Obviously, there is no national or even state-wide curriculum support for teaching this one language from among the dozens of California indigenous languages whose communities would need such support.

1.2 *Project goals*

The Yurok Language Project has three main goals. One is language documentation, that is, linguistic fieldwork with the (few remaining) fluent speakers of Yurok—recording vocabulary, researching grammatical topics, and documenting language usage. A second goal is to develop a generically and chronologically diverse corpus of texts (described further below). A third goal is to produce scholarly publications and community-oriented language materials. Work associated with all three goals is organized around a lexicon project, and is driven by two assumptions.

Our first assumption is that lexicographic and other linguistic claims should be transparent, in the sense that the data justifying those claims should be directly accessible. Evidence supporting claims made in a traditional grammar or dictionary is usually not presented in full; space may make this impossible even if the researchers would prefer it. But Oxford English Dictionary users can now easily check lexicographic claims against online corpora such as Literature Online and the Corpus of Contemporary American English, and it would be equally desirable for users of small-corpus languages to be able to do the same. Insofar as possible, as a matter of scholarly transparency, the corpus underlying linguistic claims should be available to users.

Our second assumption is related: any stakeholders may engage with the lexicon or the text corpus as researchers. In many traditional projects, there is a sharp distinction between scholarly researchers, whose findings determine the content of a grammar or dictionary, and other users. In an

endangered-language project, with many engaged participants from the heritage community, maintaining this separation would inhibit progress. In this context, stakeholders include not only scholars interested in typological or theoretical claims about language, but teachers who wish to understand why a certain grammatical claim is made, why a word in the dictionary is defined or classified in a certain way, or why it is said to have the pronunciation it has; learners who wish to learn about usage patterns that are implied but not clearly exemplified in a dictionary entry; and community members who wish to take issue with definitions of plant or animal terms, based on the usage in their own families. The quality of documentary products is improved if such users are involved with analytic decisions, and if they have the ability to intepret the underlying data themselves. How this works in the YLP database will be seen below.

2. Lexicon and documentation database

The database associated with the Yurok lexicon and text corpus has three main elements—a corpus of audio recordings, a set of texts, and a lexicon—with some additional material as well. These are described separately in the following sections.[1]

2.1 *Audio recordings*

Audio recordings in the YLP database are of two sorts: primary and secondary. A *primary* recording is a recording of one or more speakers of Yurok using the language. Some recordings are 'natural' in the sense that they document an unplanned linguistic interaction, for example a conversation in which the speakers forget that they are being recorded. But as was typical of twentieth-century language documentation, most recordings are not natural in this sense; rather, they record either an elicitation session (in which a linguist asks a fluent speaker questions about the language, how to translate English sentences into Yurok, or how to pronounce Yurok words) or a planned linguistic activity such as a narrative, procedural description, or song ('performed' for the researcher, not in the setting it might otherwise have).

The *secondary* recordings in the YLP database are derived by editing primary recordings and making shorter audio clips. Three types of secondary recordings are produced: words, sentences, and texts. *Word* recordings are selected to illustrate pronunciation; the goal is to have examples of as much of the lexicon as possible, each spoken by as many speakers as possible. (In typical cases the recording is excerpted from contexts in which an elder was asked questions like 'How do you say 'raccoon'?', with answers spoken clearly and repeated; typical word recordings are thus quite helpful for learners.) *Sentence* recordings are excerpted from elicitation sessions with a grammatical purpose,

in which Yurok sentence patterns are illustrated. Finally, *text* recordings are excerpts with self-contained stories, anecdotes, songs with words, and the like; the current database contains about 150 of these.

The audio recordings in the YLP database are quite extensive. Primary recordings include many dozens made by A. L. Kroeber and T. T. Waterman on wax cylinders between 1902 and 1909, and a smaller set made in the 1950s and 1960s by Robins and the linguist William Bright. Several dozen hours of recordings were made in the 1980s by the linguists Paul Proulx and Jean Perry, and since 2001 over a hundred hours of recordings have been made by participants in the Yurok Language Project.

All YLP recordings are associated with metadata. Given in (1) are metadata fragments for a pair of secondary recordings of words. Each fragment shows the speaker (coded by initials), details about the primary audio recording, and information about the audio file location, followed by the Yurok text, translation, and references to the lexicon ID number for each vocabulary item.

(1) Audio metadata: Secondary word recordings (XML fragments)

a. A single word

```
<item speaker-id="JVP" audio-source="JE1b" audio-
  start-time="17:01"  url="/Words/JVP/
  wohkelo'_JVP.MP3">
<tx>wohkelo'</tx>
<tr>pepperwood</tr>
<lx id="3907"/>
</item>
```

b. A phrase

```
<item speaker-id="JJ" audio-source="AG-07-1"
  audio-start-time="13:57"
  url="/Words/JJ/ku-'ne-ch'wona'_JJ.MP3">
<tx>ku 'ne-ch'wona'</tx>
<tr>my coat</tr>
<lx id="1124"/>
<lx id="354"/>
</item>
```

As of mid-2011, the YLP database includes nearly 4,000 word recordings such as these (with thousands more to be excerpted from primary recordings in the future).

2.2 *Text transcriptions*

Text transcriptions comprise the second major element of the YLP database. For the purposes of this database, these may be transcriptions of texts like

those mentioned in section 2.1 above (narratives, conversations, etc.), or they may be sets of sentences recorded during an elicitation session. Some texts, including almost all that result from modern documentation, are transcribed from audio recordings, but numerous earlier texts were transcribed from dictation (elicitation, of course, but also narratives and the like). As of mid-2011, the YLP database includes 135 text transcriptions of these various types.

Texts are transcribed in XML documents, including metadata about the text as a whole and then a series of sentence transcriptions. A pair of XML fragments representing sentences from text transcriptions are given in (2), from a traditional narrative in (2a) and a 'text' comprising elicited sentences in (2b). In each fragment, 's' stands for 'sentence' and the 'tx' and 'tr' tags precede the Yurok text and translation. The fragments also refer to 'level' (discussed below) and have a set of tags beginning with 'parsetx'. The latter parse the full sentence into words, each with a lexicon ID number, and play a role in the online user interface. Finally, in (2b) but not (2a), some attributes identify the location of the audio file itself. This is because (2b) is taken from the corpus of secondary *sentence* recordings: a separate audio file exists. The associated audio file in (2a) is a recording of the whole narrative, and is identified in text-level metadata.

(2) Text transcriptions: Two sentences (XML fragments)

a. A sentence from a recorded narrative

```
<s level="1">
<parsetx>
    <word id="1267">Kwesee</word>
    <word id="4399">'okw'</word>
    <word id="2568">'ue-peechowos</word>.
</parsetx>
<tx>Kwesee 'okw' 'ue-peechowos.</tx>
<tr>He had a grandfather.</tr>
</s>
```

b. A sentence from a recorded elicitation session

```
<s audio-path="internal/MP3/AG/" audio-filename="AG-01-
  2_07.mp3" level="2">
<parsetx>
    <word id="3895">Weet</word>
    <word id="2162">ni</word>
    <word id="4399">'oolem'</word>
    <word id="1124">kue</word>
    <word id="1248">kwegeruer'</word>
</parsetx>
<tx>Weet nee 'oolem' kue kwegeruer'.</tx>
<tr>Pigs live there.</tr>
</s>
```

2.3 *Lexicon*

The 1958 grammar of R. H. Robins contains a short lexicon presented in a traditional style. In 2003-2004 its content was combined with vocabulary from a word list published by a Yurok elder (Exline n.d.), and with vocabulary encountered during YLP fieldwork, in a database using the Linguist's Shoebox dictionary program created by SIL International.[2] This was exported as an XML file, still bearing the recognizable imprint of its Shoebox origins, which was used to generate a printed volume (Garrett et al. 2005). This volume has a Yurok-English lexicon, an English-Yurok finderlist, and indices of Yurok words belonging to various morphological and semantic categories. The database underlying this volume also now underlies the online dictionary of the Yurok Language Project.

The YLP lexicon database has about 4,500 entries in 72,902 lines, as of mid-2011, and is illustrated with two fragments in (3a-b). The fragment in (3a) shows a noun (database ID number 3907): the various tags show the form of the word (*wohkelo'*), the part of speech (n = noun), the vernacular gloss and English and Latin scientific names, a set of source references, a semantic domain (sd 40 = plants and trees), and encyclopedic information (quoted from Baker 1981). The fragment in (3b) shows a typical verb, with some details pruned. Shown here, in addition to tags like those in (1a), are a more complex part of speech tag (this is an intransitive verb belonging to the *oo*-class) and information about two irregular paradigm forms: the 3rd person singular form is *neskwechokw'* and the collective stem is *nuuem'*. Some lexicon entries are large because there are many irregular paradigm forms, and in some cases other information is also included.

(3) Lexicon: Two database entries (XML fragment)

a. A representative noun

```
<lxGroup id="3907">
<lx>wohkelo'</lx>
<ps>n</ps>
<ge>pepperwood</ge>
<taxon>California laurel</taxon>
<sci>Umbellularia californica</sci>
<rf>R264</rf>
<rf>JE82</rf>
<rf>JE102</rf>
<rf>MAB60</rf>
<sd>40</sd>
<encyclopedia>MAB59-60: "Fruit are eaten, the seeds are
picked after the pericarp has rotted off, but birds usually eat
them first. They can be gathered and buried until the shells
```

```
        rot off. Once the shells were removed, the seeds were baked in
        the sand with a fire made above. It is also a medicinal plant.
        The leaves are burned in the house to take bad luck away or the
        smoke waved over people as they leave for the same reason.
        It was put under the bed to rid it of fleas."</encyclopedia>
        </lxGroup>
    b. A representative verb
        <lxGroup id="2144">
        <lx>neskwechok'</lx>
        <ps>vi oo-class</ps>
        <ge>I come</ge>
        <ge>I arrive</ge>
        <ge>I return</ge>
        <rf>R229</rf>
        <rf>JE125</rf>
        <pdGroup pd="3sg">
            <pdf>neskwechokw'</pdf>
            <rf spkr="MM">YT1019</rf>
            <rf>R34</rf>
        </pdGroup>
        <pdGroup pd="collective">
            <pdf>nuuem'</pdf>
            <rf>R229</rf>
        </pdGroup>
        </lxGroup>
```

2.4 *Other material*

In addition to its primary content (audio, texts, lexicon), the YLP database includes a set of digital photographs, selected to illustrate toponyms, animal and plant names, and other culturally specific terms in the lexicon. It is worth adding here too that most of the YLP database is open to the public, but some audio recordings have access conditions based on cultural norms (e.g. songs regarded as private property, ceremonial texts regarded as sacred) and the preferences of families of individuals recorded.

3. Using the YLP database

3.1 *Dictionary searches*

From the web interface shown in Figure 1, users may search the dictionary just by typing in the 'Quick dictionary search' box at the top right of the screen, or

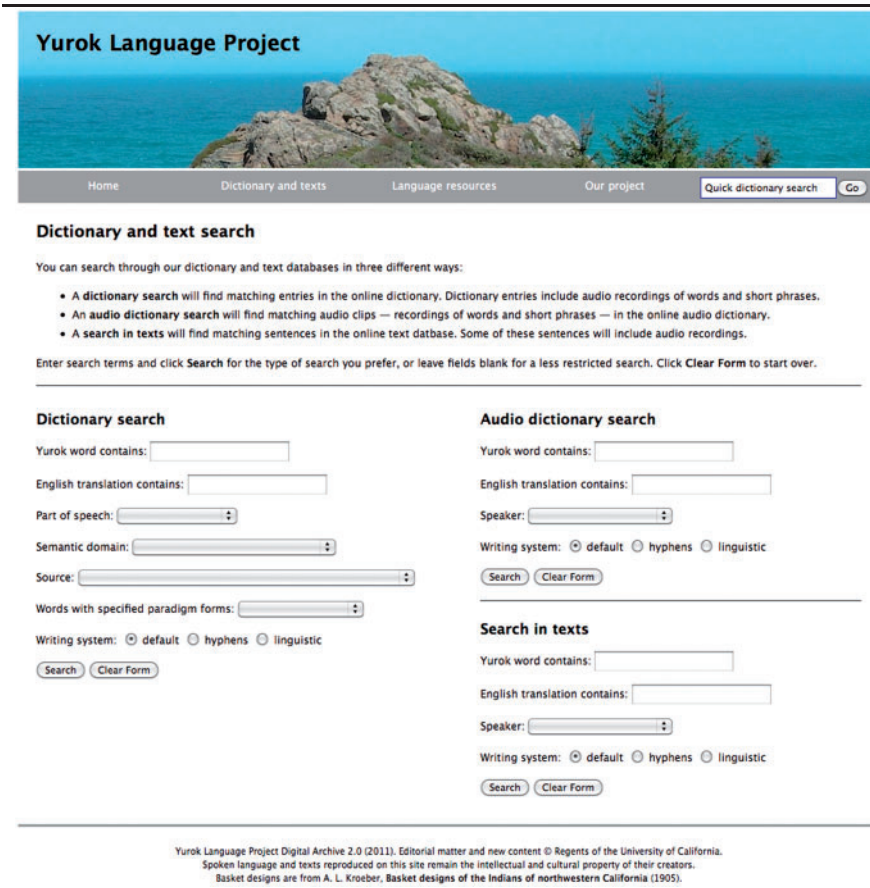

Figure 1: YLP dictionary and text search page. This figure appears in colour in the online version of the *International Journal of Lexicography*.

by entering more detail below. Three search options are available: a 'Dictionary search' will query the lexicon database, an 'Audio dictionary search' will query the audio database of secondary word recordings as in (1) above, and a 'Search in texts' will query the text corpus. Shown in Figure 2 is the result of a quick dictionary search for the term 'pepperwood', with one of the two results highlighted. (Such result pages are generated via XSL style sheets housed on an Apache server, using PHP to pass user query terms to the style sheets; the database itself contains digital audio files, photos, and text files in valid, DTD-constrained XML.)

The search result in Figure 2 should be compared with the corresponding lexicon database content in (3a) and the audio metadata in (1a). Elements of the lexical entry are displayed and spelled out, e.g. '<sd>40</sd>' as 'Semantic domain: plants and trees'. In addition, a lexicon search always also involves a search through the text corpus and the corpus of secondary word recordings. As Figure 2 shows, the latter includes recordings of the word

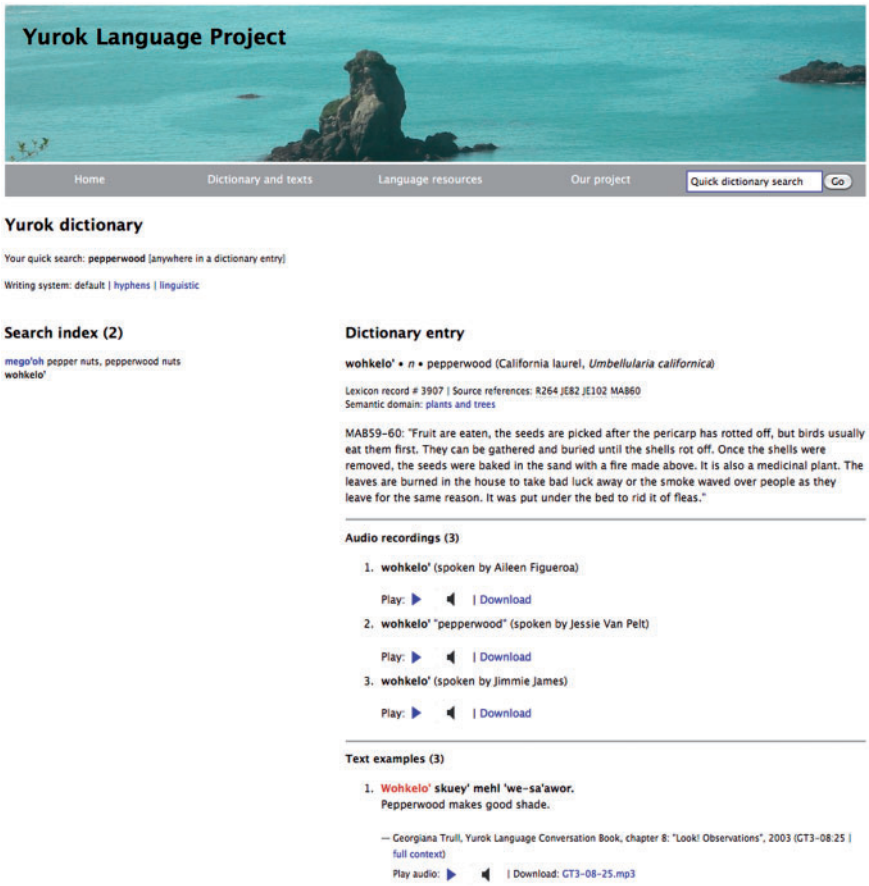

Figure 2: Yurok dictionary entry *wohkelo'* 'pepperwood' (only one text example shown). This figure appears in colour in the online version of the *International Journal of Lexicography*.

wohkelo' by three Yurok speakers; users can download the recordings or listen to them in an audio player. Finally, the text corpus yielded three matches for this word, one of which is shown. The audio clip can again be downloaded or played, and by clicking on 'full context' users can navigate to a page that displays the whole text. These elements are possible because the lexicon ID number, 3907, has been entered into audio metadata like (1a) and also into text transcriptions.

3.2 *Reading texts*

It is possible to read (and listen to) texts online through the YLP website. Users can come to texts through dictionary searches (e.g. clicking on the 'full context' link in a search result page as in Figure 2) or directly through a web page that

Figure 3: YLP text display: Florence Shaughnessy, 'The Mourning Dove' (1951). This figure appears in colour in the online version of the *International Journal of Lexicography*.

allows them to specify search parameters like date, speaker, and genre. Shown in Figure 3 is part of the display of a narrative text; compare the transcription of the second sentence of this text in (2a). Text metadata is presented at the top of the page, followed by audio options (for the Yurok text and a separately recorded translation) and the text itself.

Below the title, the first line gives several 'display styles'; 'sentence' style is selected. If a user chooses 'paragraph' style, the display is reformatted with the Yurok sentences displayed continuously, followed by the translation. If a user chooses 'look-up' style, all the Yurok words in the text are displayed as clickable links, and clicking on any link opens the associated lexicon page. This is possible because, as shown in Figure 2, all texts include 'parsetx' and 'word' tags in which the lexicon ID numbers of individual words are entered. Since complex sentences can be hard for learners to understand, direct access to relevant dictionary entries is invaluable.

4. The Yurok lexicon and language revitalization

Several YLP lexicon and website features are meant to assist learners and teachers. Here I describe one feature that integrates the lexicon and text database, and one involving the text database.

4.1 *Collocational and frequency data*

A display like that in Figure 2 reflects a search for the term 'pepperwood'; there are only two matches, and the meaning difference between them is obvious: *mego'oh* refers to the nuts and *wohkelo'* refers to the tree. But a search for general verbal meanings like 'go', 'run', and 'see' may yield dozens of matches, and to a language learner it is not always obvious how the resulting words differ in meaning or usage. This may not be a problem where language-learning infrastructure is well developed and there are many fluent speakers; it is more significant for an endangered-language community where the teachers are themselves advanced learners. A YLP dictionary entry has two features that may be helpful.

First, dictionary searches allow results to be sorted either alphabetically or by frequency in the texts. This makes it easy for users to learn which verbs with the meaning 'run' are common and which are rare, even if their literal translations might not make this obvious. For example, several Yurok verbs have a general meaning 'I run', including *ro'opek'* and *ro'onek'*, but the first is common while the second is rare. The second most frequent verb meaning 'run' is a verb *raayor'* 'run past'; again, its relative frequency would be not be guessed from meaning alone.

Second, where a dictionary entry is inexplicit about the syntax of a word (especially a verb), users can make reasonable inferences about usage from the complete dossier of text examples displayed with search results. In some cases hundreds of examples are shown, but learners find the wide range helpful. For example, the objects of transitive verbs are sometimes omitted, with an unspecified-object interpretation: in context *negeee'nowok'* can mean 'I look for her/him/it' or it can mean 'I look (around)'. But this never happens with *nepek'* 'I eat it', as users of the online dictionary have verified through texts, which clearly show that the only way to say 'I eat (something unspecified)' is *kol' nepek'*, literally 'I eat something'. This syntactic difference between the two verbal patterns is not currently represented in the lexicon, but can be inferred from an accessible corpus.

Of course, since these two features require (naive) users to work as on-the-fly lexicographers, they cannot replace careful analysis and exposition, and they cannot result in usage as precise as might be found for languages with a tradition of sophisticated dictionaries. But it is important to stress that dictionaries available in typical endangered-language revitalization settings are often necessarily simplistic: they are word lists, in effect, providing little guidance about subtleties of usage or details of meaning. The presentation of an easy-to-use corpus in tandem with a simple lexicon encourages learners to become proactive, and allows them to make at least some usage discoveries that will benefit language learning. Eventually, in the case of Yurok and other such languages, we may hope for detailed lexica that express such generalizations clearly.

4.2 *Language learning tools*

With a database of transcribed texts and audio clips of words and sentences, it is possible to develop online tools to assist language learning. Here I briefly describe one such tool, designed in collaboration with the Yurok Tribe.

A 2009 California law gives Indian tribes the right to certify indigenous language teachers for public schools, provided that they develop appropriate standards and certification procedures; teachers so certified are given state teaching credentials. The Yurok Tribe has developed a set of standards that incorporate levels of language competence ('basic', 'intermediate', etc.) defined by communicative skill and grammatical knowledge. To help teachers and learners, sentences in the YLP text corpus are now tagged for 'level'. For example, the grammatically simple sentence in (2a) above is tagged as level 1. The sentence in (2b) is level 2, i.e. 'intermediate', because it has an uninflected collective verb form *oolem'* and on account of the syntactic construction *weet nee* 'in that (area)'; both are defined elements of the 'intermediate' grammatical curriculum.

Sentence tagging by level is then incorporated in the language-learning tool illustrated in Figure 4. YLP website users can generate pages with content like Figure 4a, including an audio player and a text box. A random sentence has been selected from the YLP database, at whatever pedagogical level the user chooses ('basic' in Figure 4a), the user may listen to this as often as she likes and may then type in the text box. After clicking 'Check My Answer', the screen in Figure 4b is generated, showing what was typed as well as the actual Yurok sentence and its English translation. The user may listen again, to understand any mistakes, and may try another sentence. Since the database is reasonably large, a large variety of sentences is encountered; learners have the freedom to select the level they prefer and to listen as often as they wish.

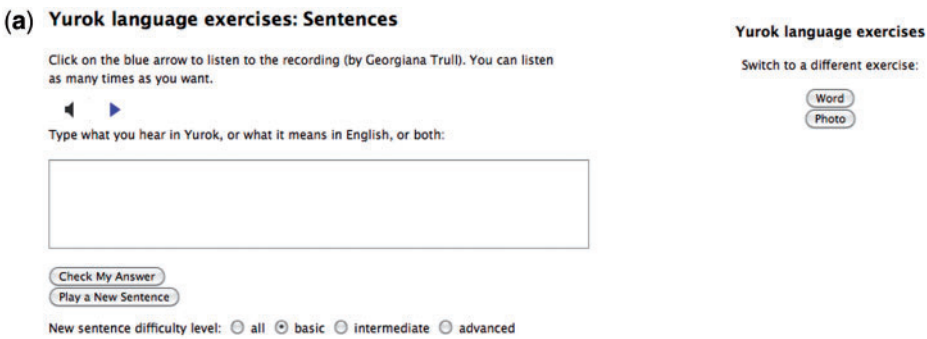

Figure 4a: Sentence exercise screenshot. A random sentence has been chosen from among the 'basic' level sentences in the YLP database; users listen and then type what they hear.

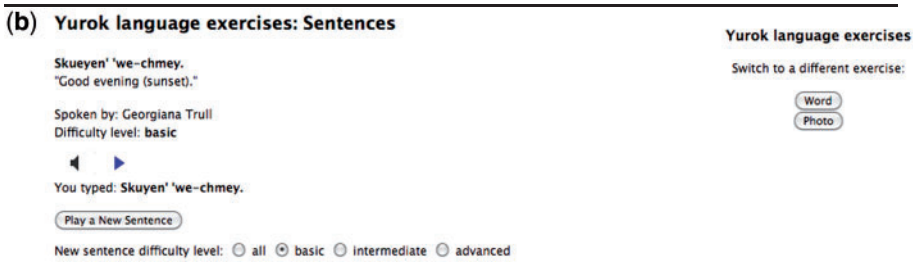

Figure 4b: The 'answer' for the exercise in Figure 4a. The screen displays what the user typed and what the sentence actually says and means; users can listen again or try a new sentence.

5. Conclusion

In this paper I have described an online dictionary project meant to be useful for scholars and an endangered-language community, with integration of lexicon and texts and several features supporting language teaching and learning. Of course, no dictionary or internet resource is ever an optimal substitute for direct contact with fluent speakers. But in a setting where severe language endangerment makes this temporarily impossible, elements of the YLP online lexicon may be feasible given suitable technological infrastructure, and might play some role in helping a language's teachers and learners.

Notes

1 The digital content of the YLP database is housed on the server of the UC Berkeley Department of Linguistics, with automatic off-site backups. It is worth emphasizing that the YLP database is not an *archive*: recordings and other documentary materials are housed elsewhere in preservation archives (including, at UC Berkeley, the Hearst Museum of Anthropology and the Berkeley Language Center); the YLP database includes only digital copies.

2 For this program see http://www.sil.org/computing/shoebox/; it was replaced by a tool called the Field Linguist's Toolbox, still widely used (http://www.sil.org/computing/catalog/show_software.asp?id=79) though there are now a number of competing off-the-shelf tools for language documentation.

References

Baker, Marc A. 1981. *The ethnobotany of the Yurok, Tolowa and Karok Indians of northwest California*. Unpublished M.A. thesis, Humboldt State University.

Exline, Jessie. [n.d.]. Yurok dictionary. [Eureka, Calif.:] Yurok Tribe.

Garrett, Andrew, Juliette Blevins and Lisa Conathan (compilers) 2005. *Preliminary Yurok dictionary*. Berkeley: Yurok Language Project, Department of Linguistics, University of California.

Kroeber, A. L. 1925. *Handbook of the Indians of California*. (Bulletin of the Bureau of American Ethnology, 78.) Washington: Smithsonian Institution.

Mithun, Marianne. 1999. *The languages of Native North America*. Cambridge: Cambridge University Press.

Robins, R. H. 1958. *The Yurok language: Grammar, texts, lexicon*. (University of California Publications in Linguistics, 15.) Berkeley: University of California Press.

Sapir, Edward. 2001. Yurok texts. Edited by Howard Berman. *Collected works of Edward Sapir*, vol. 14, Northwest California linguistics, ed. by Victor K. Golla and Sean O'Neill, pp. 1015–1038. Berlin: de Gruyter.

Teeter, Karl V. 1964. *The Wiyot language*. (University of California Publications in Linguistics, 37.) Berkeley: University of California Press.

Trull, Georgiana. 2003. *Georgiana Trull's Yurok language conversation book*. [Klamath, Calif.: Yurok Tribe.].

Waterman, T. T. 1920. Yurok geography. *University of California Publications in American Archaeology and Ethnology*, 16: 177–314.