Coding Procedures for the
J. P. Harrington Database Project

You are about to begin transcribing J. P. Harrington's notes on Native American languages and cultures. The work you will be doing is the heart of this project. Because of your efforts native people and scholars will benefit from better access to this unique collection of information.

Please read this through once carefully, and then refer to it as needed. We expect to update these procedures, so please let us know if you find any instructions incorrect or unclear.

Level 1 coding is the most important, and the most challenging. It requires the coder to read and to type the actual handwritten notes, along with citation information. Even with this first level of coding, you will have created a searchable document, greatly increasing access to the material these notes contain.

Level 2 coding includes additional data required to automatically generate word lists from the database. Here the coder must add an Indian word, if there is one, as well as Harrington's gloss in either English or Spanish (sometimes both). With this level of coding you have further increased access to the material, allowing us to print out rough word lists sorted by the Indian language, or by the English and Spanish meanings, very valuable to native communities.

Level 3 coding includes additional historical, linguistic, and cultural information. This level of coding will enable you to search the database by categories, such as *birds, kin terms, place names,* etc. Since the catagories of genre, language, dialect, etc. are often repetative, it is easier for the project staff to add this information to the computer file once you are finished.

Level 4 coding will be done by those who have a background in linguistics. With this level of coding, roots, grammatical morphemes, and other sorts of linguistic analysis will be noted. Level 4 coding is not absolutely essential, and will probably not be done for every language.

The computer file you are working on may not have all of the field names listed below, or it may have them in a slightly different order.

The database format we are using does not allow more than one font within a data cell, so we have designed keyboard equivalents in Times or Times New Roman font for typing various phonetic symbols found in Harrington's notes (see chart of keyboard equivalents). Likewise, the database format does not allow any tabs or returns within a data cell. If used they will cause data to be lost in a transfer. So please use avoid typing tabs and returns.

LEVEL 1 (transcription with minimal coding)

Citation
Type the *volume* number, *reel* number, *frame* number, *paragraph* number, and *sentence* number each separated by a colon, with no spaces. For example:

2:005:358a:7:1

indicates volume 2, reel 5, frame 358, side a, paragraph 7, sentence 1. J. P. Harrington usually writes in paragraphs, so numbering follows his format. He usually begins a sentence with a capital letter. If there is a lists, figure, or map on a page, count each line or label as a separate paragraph of one sentence.

Volume is always written with one digit, but *reel* and *frame* must be **three** digits. *Paragraph* and *sentence* must be two digits only if there are more than 10 paragraphs or sentences on a page (for example a list of words or a map with more than 10 labels). The initial zeros allow the computer to reorder the records accurately after they have been sorted alphabetically by word, consultant, language, etc. *Frames* generally have two pages, labeled here as *a* and *b*. The *a* frame is the page on the left and the *b* frame is on the right. The frame information is written on the back of your hard copy in the lower right hand corner. Unnumbered frames have 000 plus a letter, e.g., 000a, 000b, 000c. Enter this identifying information for each record (entry) that you create.

Text
Text is entered exactly as it appears on the page. Even if words are abbreviated, misspelled, or sentences unfinished, type the text without revision. If a word is misspelled type {sic} after it, indicating that the word is a misspelling by Harrington and not a typing error on your part. Wavy brackets {} are also used to describe words or sentences Harrington may

have inserted with an arrow or other mark. Wavy brackets are also used for brief descriptions of drawings or maps.

As mentioned above, paragraphs are usually indicated in Harrington's notes by indention. Headings, like lists and map labels, are considered separate paragraphs of one sentence. In the absence of a period at the end of what is obviously a sentence, write {sic}.

Break a sentence into a new record (entry) every time there is an Indian word. This allows us to generate dictionaries automatically from the database. Create a new record with the Indian word and continue with the sentence. The first portion of the sentence is labeled with *a*, for example: **01a**. The next portion of the sentence is labeled **01b**, etc.

Do not break the sentence if two or more Indian words are written together and have a single meaning. Usually such items have a comma following the second word followed directly by the English gloss. If you are unsure, do not divide the phrase. We will check this in level 3 coding.

Notes
This is where you can add information to clarify what you have typed, reference a previous entry, or make a comment that may help with interpretation of the notes.

LEVEL 2 (transcription plus dictionary coding)

Consultant
This field is for the name of the consultant Harrington was working with. He may use initials, a nickname, or sometimes a full name. It is usually written in the beginning of the paragraph or sentence.

Indian Word
This field is used to isolate the Indian word from the text to enable the project to create word lists and dictionaries. You can fill in this field by copying it from the *text* field.

English Gloss (Harrington)
This is Harrington's English translation for the Indian word as it appears in the notes. This gloss comes directly from the text, although the information may be in the previous or following sentence. Do not, however, try to guess what he means. It is better to leave the field blank.

Spanish Gloss (Harrington)
Some of Harrington's fieldwork was conducted in Spanish. This gloss comes directly from the text, although the information may be in the

previous or following sentence. Do not, however, try to guess what he means. It is better to leave the field blank.

The dialects of Spanish Harrington encountered among native people include Old California Spanish as well as more standard varieties. Please do not attempt to translate these in the English Gloss field, even if you think you know what they mean.

LEVEL 3 (additional linguistic, historical, and cultural information)

Semantic Domain
If there is an Indian word in this sentence, choose the best category from the list provided to describe it. If you are not sure of the best category, leave this field blank.

Genre
This refers to the type of information being collected in the notes. Story, dance, song, or word list are some of the categories. Choose the best one from the list provided. If you are not sure, leave this field blank.

Text Notes
Most of Harrington's work is hand written, so text notes identify anything that needs to be explained about the texts, such as whether it is typed, written on a card or slip note, etc.

Page Notes
Here the coder identifies anything out of the norm in the notes such as a page number or other significant markings on the page that are not part of the text, or a map or drawing that needs to be described.

English Gloss of Harrington's Spanish
Translation of Harrington's Spanish by someone familiar with the varieties of Spanish he recorded.

Scientific Name (Harrington)
Sometimes Harrington gives the Latin name for various flora or fauna

Scientific Name (current)
Here the coder would put the corrected or current scientific name for the flora or fauna referenced by Harrington.

Language Family
This is the language family of the language in these notes. This will be filled in by project personnel.

Language
This is the language name used by Harrington. It may not be a current or widely used name. Many times he uses regional names for languages. **Fill in this record** if you are doing Level 3 coding. This is how we find words in one language commented on while Harrington is working on another.

Dialect
Many times Harrington provides the name of the dialect he is writing about. Other times it can be identified by the family, or town or region of the person being interviewed.

Location of Interview
Generally the location of the interview is on the first page of a set of notes. Enter the state abbreviation, usually *CA*, followed by a space followed by a county or town separated by a comma, followed by the immediate location, for example: **CA Orange County, Santa Ana, Miller's Ranch**.

Date of Interview
Occasionally the date of an interview can be found on the first page of a set of notes. For some interviews the date is provided in the guides prepared by Elaine Mills. Leave the field blank if you have no information. Please enter all dates in the following format: YY/MM/DD with zeros before single digits, for example: **23/09/26** for September 23, 1923. This will allow us to sort records by date automatically.

Reference to Personal Communication
This is for references to personal communication between Harrington and other scholars. The usual format is full name followed by a date, for example:
personal communication, Alfred Kroeber 1938.

Reference to Written or Published Communication
This is for references to published materials. Please use the following format:
Kroeber 1903 **or** *Kroeber 1903:283.*

Photograph
Harrington took many photographs of the people he worked with and the places he traveled. This field references photographs that relate to the notes. The photographs are housed in the Smithsonian National Anthropological Archives, and will be incorporated into this database at a later time.

Ethnographic/Archaeological Data
Ethnographic or archaeological information related to the notes.