Language Documentation in the Age of Smartphones

EUGENE BUCKLEY
UNIVERSITY OF PENNSYLVANIA

Native American Studies at Penn: Indigenous Knowledge in the Academy

7–8 November 2014
Indigenous languages in the US
UNESCO endangerment criteria

<table>
<thead>
<tr>
<th>Degree of endangerment</th>
<th>Intergenerational Language Transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>safe</td>
<td>language is spoken by all generations; intergenerational transmission is uninterrupted &gt;&gt; <em>not included in the Atlas</em></td>
</tr>
<tr>
<td>vulnerable</td>
<td>most children speak the language, but it may be restricted to certain domains (e.g., home)</td>
</tr>
<tr>
<td>definitely endangered</td>
<td>children no longer learn the language as mother tongue in the home</td>
</tr>
<tr>
<td>severely endangered</td>
<td>language is spoken by grandparents and older generations; while the parent generation may understand it, they do not speak it to children or among themselves</td>
</tr>
<tr>
<td>critically endangered</td>
<td>the youngest speakers are grandparents and older, and they speak the language partially and infrequently</td>
</tr>
<tr>
<td>extinct</td>
<td>there are no speakers left &gt;&gt; <em>included in the Atlas if presumably extinct since the 1950s</em></td>
</tr>
</tbody>
</table>
Languages in the Southwestern US

- Kashaya
- NE Pomo
- SE Pomo
- Yavapai
- Havasupai
- Hualapai
- Navajo

Languages in the Southwestern US.
Documentation

• compiling information about a language
  – written transcription
  – audio recordings
  – video recordings

• for academic interest
  – to understand languages in general
  – their structure and how they work
  – how they’re historically related to one another

• in the community interest
  – to sustain and revitalize endangered languages
  – to revive languages with no native speakers

• these goals sometimes overlap
  – but often lead in quite different directions
Language families of California

**INDIGENOUS LANGUAGE FAMILIES**

**Algic**
- Wiyot
- Yurok

**Athabaskan**
- Bear River–Mattole
- Chetco–Tolowa
- Eel River
- Hupa
- Kato

**Chumashan**
- Barbareño
- Ineseño
- Interior Chumash
- Island Chumash
- Obispeño
- Purisimeño
- Ventureño

**Hokan**
- Chimariko
- Esselen
- Karuk
- Palaihnihan languages
- Pomoan languages
- Salinan
- Shastan languages
- Washo
- Yana
- Yuman languages

**Penutian**
- Klamath–Modoc
- Maiduan languages
- Miwokan languages
- Chilone languages
- Wintuian languages
- Yokutsan languages

**Uto-Aztecan**
- Numic languages
- Takic languages
- Tubatulabal

**Yuki–Wappo**
- Coast Yukian
- Northern Yukian
- Wappo
s’uʔnuʔnu šínał : Huckleberry Heights
Consulting with speakers

Essie Parrish & Robert Oswalt

at Stewart’s Point Rancheria

(s’uʔnuʔnu šínal, Huckleberry Heights)

1960
Words and sentences

• Kashaya “bear” by itself
  (provided by Essie Parrish)
  – [what is the word for] “bear”
  – buṭaqá ; buṭaqá

• Kashaya “bear” in a narrative
  (told by Herman James)
  – mulídom pʰiʔt’an duwéʔli buṭaqá ṭem ma·dúʔ.
  – “Suddenly, in the evening, the bear arrived home.”
Taking notes (1957)

- any bean
- black bean
- horse
- dog
- deer
- panther
- lion
- wild cat
- cat
- chicken

Note: Some entries are in another language, possibly indicating a non-English context.
Organizing information

butaca

gha butaca sealion or

butaca birka as an ant or

butaca kili black bear

gahle polar bear "white bear"

uncia butaca brown bear 6/69

animal
Use of information

• technical papers about language structure
  – typically necessary for academic progress
  – but not useful to communities
• grammars written for specialists
  – may contain many useful examples of the language
  – but mainly accessible only to specialists
  – sketches of the grammar can be written with more familiar terms
• dictionaries are an essential resource
  – but like grammars they may be very technical
  – most community members prefer simpler wordlists
  – audio recordings make pronunciation much clearer
A xeroxed wordlist (1975)

baby  naṭa qáwi, ná·na
back (of body)  baḍo·
bad  pišudu
bag, sack  mišukh
ball  phíʔko
bank, cliff  ʔamahchili
bank (undercut)  seʔpétmo
bar, rod (of metal)  wá·ra
bare, nude  qahqo
bare, without  laʔka
bark (of tree)  qahwa·
barrel  warí·i
barren, sterile  haʔdas

bat (animal)  qahca tínaʔna
beads
of clamshell  kuhnu
of magnesite  na·muli
seed beads  walehólyo
beans  yihó·l

bear (animal)  buṭaqa
brown bear  mima
human bear  khuy
beard  hasime
beautiful  čiškan
bed, floor  cahti
bee, bumble-  cehey
small species  holóm?lo
- hcha- ʊl V. fall over, push over; of stationary, vertical long object, usually inanimate. Pl.Act -hcha-Xt- : -chat-. Cf. ²c'eəl-, -dahchi-, -da?di-.

²-+ ba-+ ca-+ da-+ ha-+ ma-+ pha-+ qa-+ $a-+
bi-+ chi-+ di-+ hi-+ mi-+ phi-+ si-+
cu-+ du-+ mu-+ phu-+ $u-+
Examples with all prefixes are in KG 410-411.21.

²-hcha-w /chaw/ to fall over (from no apparent cause).
/sima' hchaw/ to fall asleep. Calque based on English.
²-hcha-m-? /cham?/ to fall onto something; (walls) to fall in.!!
/bahchaw/ to knock over with the mouth, snout, or beak (as a pig might do).
/bihchaw/ to put both arms around and throw over (as in wrestling).
/dahchaw/ to push over with the hands or paws.
/bu7taga@1 ya?@2 ?ahca@3 da'hchaw@4./ A bear@1 pushed@4 the house@3 down@4.

[Ps ?ahcha-, Pcy cha'ù-]
Practical problems

• even if we include cross references for all the possible forms, this becomes cumbersome in a printed volume
  – simply takes up too much space on paper
• also frustrating to find the root and then have to flip to the main entry many pages away
  – or vice versa, if organized by roots
• much easier to search *electronically* for a particular form and click to be taken directly to the relevant entry
  – similarly, click between various related entries
A computer database
A new wordlist (2014)

bare-legged  
adj. **puhṭi**
bark  
n. **qahwa**
bark house  
n. **hca s’i?da hca**
barked, peeled off  
n. **hca s’i?da**
barn  
n. **kawa’yuhca**
barn, cow  
n. **wá’kahca**
barnacle  
n. **qhac’o**
barrel  
n. **wari’l**
barren, sterile  
adj. **ha?dás**
bartender  
n. **tawené’ru**
baseball  
n. **p’hi?k’op’hik’o? ṭ’hahqaw**
basement  
n. **?ama? yó? ?ahca**

bean, lima  
n. **yihó’l baḥṭ’e**
bean, string  
n. **yihó’l s’amo?**
beans  
n. **yihó’l**
bear  
n. **buṭaqá**
bear, human  
n. **ʔaca? buṭaqá**
beard  
n. **hasimé**
beater, seed  
n. **yuhu šám?**
beautiful  
adj. **c’iškan**
because  
conj. **ʔé?**
bed  
n. **cahti**
bed, hot  
n. **honém?li**
bedding  
n. **šamát’h ma?**
<table>
<thead>
<tr>
<th>Kashaya-English</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bohóhso</strong></td>
</tr>
<tr>
<td>n. Sour clover, bull clover.</td>
</tr>
<tr>
<td><strong>bóṭ’</strong></td>
</tr>
<tr>
<td>adj. soft.</td>
</tr>
<tr>
<td><strong>boʔóm?</strong></td>
</tr>
<tr>
<td>n. 1 • purse, sewing bag, bundle, traditional buckskin bag for valuables.</td>
</tr>
<tr>
<td>2 • sack full of (songs): many, or enough, songs.</td>
</tr>
<tr>
<td><strong>budu</strong></td>
</tr>
<tr>
<td>n. toyon berry.</td>
</tr>
<tr>
<td><strong>budu qʰále</strong></td>
</tr>
<tr>
<td>n. toyon.</td>
</tr>
<tr>
<td><strong>buhkʰu</strong></td>
</tr>
<tr>
<td>adj. humped.</td>
</tr>
<tr>
<td><strong>buhlu</strong></td>
</tr>
<tr>
<td>adj. hunched; stooped.</td>
</tr>
<tr>
<td><strong>buhqʰál</strong></td>
</tr>
<tr>
<td>n. burden basket, coarse-woven basket; container (in certain constructions).</td>
</tr>
<tr>
<td><strong>buʔt’u</strong></td>
</tr>
</tbody>
</table>
A full web dictionary

<table>
<thead>
<tr>
<th>Kashaya</th>
<th></th>
</tr>
</thead>
</table>
| **bute** | Adverb. 1 • when. buté ?wa ṭa ya ʃuqʰæc’hk’e’ when are we going to rest? hic’i buté ?še ʔyowam mu’ let’s see, when was that now?  
2 • any time, ever. muʔnati bute’ ma’dúʔ tʰìn but still he never returned. bute’ men ṭama’ cic’i-du tʰìn she never did anything like that again. mu ṭo bute’ bahyeʔk’h e tʰìn she is not to stop singing ever.  
**but’uc** | Verb. sprout (above ground). but’ucí sprout! but’ú-tʰu’i don’t sprout! but’ucí-du keep sprouting. but’úhcime? sprout! (to several plants). ?ul but’ucqʰ it’s sprouted already. *synonyms: cubuš*; *compare: culuš*.

| **buṭaqa** | Noun. bear (general term, and grizzly).  
**buṭaqa biʔk’as** [buṭaqa biʔk’as] ⇒ buṭaqa, biʔk’as. Noun. carpenter ant, black ant.  
**buṭaqa ch’ilinta** [buṭaqa ch’ilinta] ⇒ buṭaqa. Noun. chickadee?  
**mima buṭaqa** [mima buṭaqa] ⇒ mima, buṭaqa. Noun. brown bear.  
**buṭuc** [buṭúʔ?] ⇒ buṭuc*. Adjective. skinny. |
Special characters

• a new problem is how to input special characters that are often needed
  – this didn’t happen in books, at least
• special keyboard, or letters to click for insertion
  – an issue for many national languages also
  – more challenging when the character is not found in common European languages, such as ֹ or ۇ
• another option: search under basic letters, be shown a range of matching options
  – such as typing ta and getting all the Kashaya t options
  – interesting use of this concept in (printed) Zapotec
Zapotec language family
Valley Zapotec
Zapotec vowels

• four kinds of PHONATION
  – modal (plain) [a]
  – creaky [à]
  – post-glottalized [a’]
  – breathy [ah]

• four tones
  – high, low, rising, falling

• in total, 33 different possible combinations
  – on any particular vowel or diphthong

• this leads to very complex spellings
  – a’, aa, àa, a’ah, a’àa, aàa, aàa’, àa’a, etc.
  – but necessary if the exact pronunciation is to be shown
“Once upon a time there was a man who had eight daughters.”
### Simplified headwords

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>becw</strong></td>
<td>dog</td>
</tr>
<tr>
<td><strong>bel</strong></td>
<td>meat: esp., beef</td>
</tr>
<tr>
<td><strong>Be'Il</strong></td>
<td>Abel</td>
</tr>
<tr>
<td><strong>bèe'Il</strong></td>
<td>naked</td>
</tr>
<tr>
<td><strong>bèe'Il, bèèe'lld</strong></td>
<td>sister (of a woman)</td>
</tr>
<tr>
<td><strong>bèèe'Il, bèèe'lld</strong></td>
<td>snake</td>
</tr>
<tr>
<td><strong>behll, behlld</strong></td>
<td>fish</td>
</tr>
</tbody>
</table>
Autocorrection

• Google can correct our spelling

• and suggest what we can’t remember
Autocorrection in Spanish

- it can supply special characters

- or correct typos
Autocorrection in Kashaya?

• it could suggest similar words

- Google search suggests:
  - ma·yu ‘dove’
  - maya ‘you’
  - mahyuṭʰmaw ‘poke or stir (a fire)’
  - maʔyuʔyuw ‘walk painfully (on gravel)’

• no word begins exactly with *mayu*, but there are some fairly close (“fuzzy”) matches
Match for Navajo dloo

dlöö’ in English

Translation and definition "dlöö’", Navajo-English Dictionary online

Translators into English:

- prairie dog

  (Noun ) (noun )

  burrowing rodent

  (Cynomys) A small, stout-bodied burrowing wild rodent with shallow cheek pouches, native to North America and Central America.

Other meanings:

- prairie dog

Similar phrases in dictionary Navajo English. (82)

- döö
  aadöö
  adâgâdâ
  ak’öö
  alohinidâ
  sego
## Smartphone dictionaries

<table>
<thead>
<tr>
<th><strong>deer</strong></th>
<th><strong>shaa nílé</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>n.</strong></td>
<td><strong>v.</strong></td>
</tr>
</tbody>
</table>

**Description**
- : a wild animal with brown fur

**Example:**
- Haigo bjjh bitsj' ayóó lıkán. During the winter, deer meat tastes so good.

**Synonym**
- good, bjjh

**Description**
- : to ask for an object with the characteristic of being a slender and/or stretchable cord and being able to be carried

**Example:**
- Tł’óól shaanílé. Hand me the rope.
- Béésh’ahédílí shaa nílé. Hand me the scissors.
- Shikee’ shaanílé.
Smartphone learning tools
Specialized keyboards
Conclusion

• digital technology creates new challenges for academics
  – it’s no longer enough to produce traditional printed materials
  – we have to learn how to use these technologies
• but it also creates new opportunities
  – modern dictionaries are created with computers anyway
  – we just need to take that information and create resources for the age of smartphones
Yahwiy!