New Kusunda data: A list of 250 concepts

**This is a joint post by Uday Raj Aaley (independent researcher, Dang, Nepal) and Timotheus A. Bodt.**

Between 29th July 2019 and 12th August 2019, we invited the then remaining two speakers of the Kusunda language to Kathmandu, where we interviewed them. One of these speakers, Gyani Maiya Sen Kusunda, unfortunately passed away early 2020. At the moment of writing this, there is only one Kusunda speaker left, Kamala Sen Kusunda.

We made a total of around 20 hours of video- and audio recordings. Part of our research involved the triple-repeated recording of a 250-concept word list from both speakers. The concepts themselves were taken from the original concept lists underlying the study of Sagart et al. (2019) on Sino-Tibetan languages. This means that comparisons of the new data for Kusunda can be conveniently done with any of the 50 languages in the sample reported by Sagart et al.

We make both this list and the original sound recordings available along with this contribution, and invite everyone to work with these data and communicate their findings. Any comments on the phonology of Kusunda are welcome, so that these can be used to further refine both the Roman and Devanāgarī orthographies used for writing and teaching Kusunda to the next generation of Kusunda speakers.

Out of the 250 concepts, the speakers did not remember 9 concepts and thought that 20 concepts did not exist. They had descriptive compounds for 3 concepts but did not agree on the exact form and they did not have consensus over the form in 7 other cases. A total of 10 concepts are confirmed Nepali loans (including all numerals 5-10). This left 200 concepts for comparison.

The 250-concept list consists of the 250 concepts with an ID in the first column, an English gloss in the second column, the identifier of Huáng’s Tibeto-Burman Lexicon (1992), a link to the Concepticon project (List et al. 2020), a preliminary Kusunda reconstruction based on the available forms, comments on this reconstruction, a phonetic transcription of the form attested from Gyani Maiya with comments, and a phonetic transcription of the form attested from Kamala with comments.

The following table shows a small excerpt of the spreadsheet (note that columns and rows have been transposed here):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **ID** | 33 | 34 | 35 | 36 | 37 |
| **ENGLISH** | the dew | to die | to dig | dirty | the dog |
| **Kusunda** | ∅ | oɢ.da | mek | hu.wi.gɐn | ɐ.gəj |
| **Comments1** | does not exist, new compounds are made based on FALL + WATER, STAY + WATER etc. |   |   | < hu.wiː aː.gɐn, hu.wiː = ‘dirt, dust’, also verb ‘be dirty’ |   |
| **Gyani Maiya** | ∅ | ɔk.d̪aː | mʲɛk.tɔː | huj.gɐn | ɐ.gəj |
| **Comments2** | tɐŋ ʣɐk.ʣi water + stay i.e. ‘water which has accumulated on / in leaves’ | ɔk.daː = die, ɔk.daː aː.gɔː = kill | mʲɛχ.tɔː, mʲɛʔ aː.gɔː |   | ~ aː.gəj |
| **Kamala** | ʤʲun t̪ɐŋ | ɔ̰.d̪aː | mʲɛk ɐ.gɔː | huj.gɐn | ɐ.gəj |
| **Comments3** | KGG\_310719\_A1; lit. FALL + WATER | ‘to die (imp)’, ‘to kill’ ɔ̰ˁ.daː ɐ.gɔː (imp) | ~ mʲɛːʔ.ɐɔː | < hu.wiː ɐ.gɐn, ~ huj.ʤiː; hu.wiː = dirt (n), dust | ~ aː.gəj |

The Kusunda reconstruction is based on the following observations and analyses:

* In general, word-initial affricates are palatalised in K, but non-palatalised in GM, whereas word-medial affricates are always non-palatalised in GM and sometimes non-palatalised in K. When preceding rhymes with vowel /i/, even GM *may* palatalise the word-initial and -medial affricates. Because all previous sources list only a dental affricate series [ʦ, ʦʰ, ʣ, ʣʰ], all affricates are thought to derive from this dental series, even when both speakers realise a palatal affricate [ʧ, ʧʰ, ʤ, ʤʰ].
* GM initial uvular stops [q, qʰ] corresponds to K initial uvular stops [q, qʰ]. Uvular stop [ɢ] has been completely lost word-initially in both speakers.
* Where a GM final [q] or [χ] corresponds to a K final [q] or [χ], this is thought to derive from underlying final \*q.
* Where a GM final [χ], marginally [k] (sometimes preceded by a creaky vowel [v̰]) corresponds to a K creaky vowel [v̰ː] this is thought to derive from underlying final \*ɢ.
* Where a GM intervocalic [χ] corresponds to a K intervocalic [qʰ], this is thought to derive from underlying initial \*qʰ.
* Where a GM intervocalic [kʰ] corresponds to a K sequence of creaky vowels [v̰.v̰], this is thought to derive from underlying initial \*ɢ.
* Where a GM final [ŋ] corresponds to a K open nasalised vowel [ṽː] this is thought to derive from underlying final \*ɴ.
* Where there is variation between vowel /o/, realised as [ɔ], and vowel /u/ in open syllables, this is thought to derive from an underlying vowel \*u, with varying realisations depending on speaker. In closed syllables, variation between a [ɔ] and an [u] is thought to derive from an underlying vowel \*o.
* Where a vowel /i/ varies with vowel /e/, the choice has been made to represent this in the ground form as \*e, because vowel \*i is preserved in both speakers.
* Where a rhyme [ɛ, ɛCf] is preceded by a palatalised onset this is thought to derive from underlying form \*e, \*eCf, respectively.
* Vowels /ɐ, ə/ are always short and hence vowel length has not been indicated in column 3.
* Length of vowels /i, ɔ, u/ is predictable on phonotactic position (long in open syllables, short in closed syllables) and has hence not been indicated in column 3.
* Vowel /a/ is always long and hence vowel length has not been indicated in column 3.
* Labialised onsets (e.g. sw-, gw-) are thought to be old and are hence indicated in the ground form.
* The off-glide in rhyme -ej is thought to be epenthetic and derive from an open rhyme \*-e.
* Syllable-initially, both K’s and GM’s speech may show a simplification of diphthongs /əj/ and /ɐj/ to /ə/ and /ɐ/. Syllable-finally, both K’s and GM’s speech may show a simplification of diphthongs /əj/ to /a/. Where there is variation between K’s [-əj] and GM’s [-ɐj], -ɐj has been taken as underlying because K’s [-əj] generally corresponds to GM’s [-əj].
* Where GM’s rhyme [-ɔw] varies in realisation with [-ɔː] and corresponds to K’s rhyme [-uː], this is thought to derive from rhyme \*-ow.

The data has been submitted to Zenodo, where it can be accessed in its version 1.0 via its DOI [10.5281/zenodo.3377537](http://doi.org/10.5281/zenodo.3377537). We will be very happy for any kind of comments or suggestions, for which contact details can be found in the data we archived with Zenodo.

**References**

Huáng, Bùfán 黃布凡 (1992): {Z}àngmiǎn yǔzú yǔyán cíhuì 藏緬語族語言詞匯 [A Tibeto-Burman lexicon]. Běijīng 北京: Zhōngyāng Mínzú Dàxué 中央民族大学 [Central Institute of Minorities].(Available online at [https://stedt.berkeley.edu/ stedt-cgi/rootcanal.pl/source/TBL](https://stedt.berkeley.edu/%20stedt-cgi/rootcanal.pl/source/TBL))

List, Johann Mattis & Rzymski, Christoph & Greenhill, Simon & Schweikhard, Nathanael & Pianykh, Kristina & Tjuka, Annika & Wu, Mei-Shin & Forkel, Robert (eds.) 2020. Concepticon 2.3.0. Jena: Max Planck Institute for the Science of Human History. (Available online at [http://concepticon.clld.org](http://concepticon.clld.org/))

Sagart, Laurent and Jacques, Guillaume and Lai, Yunfan and Ryder, Robin and Thouzeau, Valentin and Greenhill, Simon J. and List, Johann-Mattis (2019): Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Science of the United States of America* 116. 10317-10322. DOI: [10.1073/pnas.1817972116.](https://doi.org/10.1073/pnas.1817972116)