Consonant Codes in 267 Sino-Tibetan Languages

Thirteen branches and two consonants: the family with the best structure and the worst instrument

Alejandro Toledo Martínez · ORCID 0009-0000-1277-9697 · Integrative Corpus · 2026

Abstract

We present a catalogue of 10,497 consonant codes grouping 99,061 lexical forms from 267 Sino-Tibetan languages, built with no etymology, no proto-forms and no cognacy judgements.

This family has at once the best and the worst profile in the series, and both for the same reason. Its thirteen branches — nine with nine or more languages, the largest at 38% — give the richest genealogical material of the eight families measured: 1,819 codes cross three or more branches. And its words are the shortest, 2.22 consonants per form, with the highest collision, 71.9%: shuffling concepts within each language already groups 42.0% of the forms, against 4.7% for Afro-Asiatic.

The widest-reaching codes reproduce the comparanda Sino-Tibetan comparison established a century ago, crossing six to ten branches and having reconstructed nothing: `m·n` for *name* in 53 languages — *miŋ* —, `l·n` for *stone* in 44 — *luŋ* —, `m·k` for *smoke* in 46.

And the figure of 1,819 is not published alone: 58% of those codes are a single consonant, and the crossing rate falls with arity — 28.0% for unary, 14.4% for binary, 3.3% for ternary. The shorter the code, the more it crosses, which is what one expects if crossing comes from availability rather than inheritance. The defensible figure is 762.

This work also carries a prediction registered before anything about this family was computed, to test out of sample a finding from the previous paper. It is confirmed in its principal part and fails in a secondary one, and the previous paper is corrected by what is measured here.

1. What is sought, and why this family

The procedure is described once for the whole series in ”The catalogue of consonantal codes” and is not repeated. The essentials: a code is a consonant skeleton appearing for the same concept in three or more distinct languages, and it is not claimed that its members are cognate, that they descend from a proto-form, or that the code means anything.

The series plan called this family ”the grand prize”, and also deferred it. It had the best branch structure in the corpus and the worst instrument cost at the same time, and was left for fourth place ”because by then the three decisions it needs — segmentation, tone and branches — will have been taken in families where they cost less”. They have been: automatic segmentation was diagnosed and left out of the catalogue, tone was declared a dimension the skeleton does not capture, and branch structure was worked first in Pama-Nyungan and Nuclear Trans New Guinea.

2. The population

267 languages · 200,339 forms · 2,737 concepts · 13 branches.

Branch	Languages	Forms
Burmo-Qiangic	102	91,002
Kuki-Chin-Naga	38	30,878
Sinitic	33	28,610
Bodic	23	15,684
Kho-Bwa	18	9,391
Macro-Bai	17	7,300
Karenic	14	2,555
Himalayish	10	5,020
Brahmaputran	9	5,984
Nungish · Digarish · Kman-Meyor · Macro-Tani	2+2+1+1	2,541

100% of the forms obtain a skeleton, and no language is excluded. The largest branch is 38% — the best balance in the series, against 51% for NTNG and 72% for Chadic — and nine branches have nine or more languages.

A single macroarea: the anomaly list of §6.1-bis does not exist here.

2.1 Thirteen branches, not fourteen

The corpus carried **`Kho-Bwa`** with seventeen languages and **`Khobwa`** with one — Miji Nafra, 1,520 forms. It is the same branch with and without a hyphen, and Miji belongs to the Kho-Bwa group.

This is not cosmetic: a branch written two ways counts as two, isolates a language in a branch of its own, and distorts any measure of ”codes crossing N branches” — which is precisely the most important section of this paper. Declared as a corpus fact. Verified by normalising the branch labels of all 225 families: it is the only case.

2.2 Two Sinitic languages flagged by distance, and none excluded

The geographic check proposed excluding **`Ha_erbin`** and **`Shenyang`**, at 3,258 and 2,793 km from the family’s centre. They are north-eastern Mandarin: fully Sinitic, and far because the family stretches from Pakistan to north-eastern China, with a median of 628 km from the centre.

It is the fourth time in the series this check has proposed too much — in Afro-Asiatic it wanted to delete Semitic entirely, 30 of 112 languages. The intruder it did get right, the Tai-Kadai language inside Uralic, was flagged by three checks at once: distance, tone digits, and its own forms. That is the difference, not a threshold: in an extensive family, being far from the centre is normal.

Corrected for the whole series: distance is now a signal that accompanies rather than decides, and only proposes when another independent check flags the same language. Verified that Uralic still catches its intruder.

3. What the instrument costs here

Mean arity: 2.22 consonants per form. The lowest of the eight families. Collision: 71.9%, the highest. More than two in three forms share a skeleton with another concept within their own language.

The two figures are the same thing said twice, and everything else follows. Extreme cases: k in Ng'ga Kant-thayung covers eleven concepts — STINKING, bitter, die, drink, eat…; chü in Ntenyi, nine; ʑi²¹ in Kua-nsi, nine.

3.1 Tone, which the skeleton does not capture

137 of the 267 languages mark tone. The skeleton discards it, as it discards vowels, and that must be said where the object is defined and not in a note: in a family where tone is contrastive, the consonant skeleton sees less than the language distinguishes.

And tone arrives written two ways: superscripts (ʑi²¹, ɳi³³) and ASCII digits (an4, sɔ42, te22_mok44) — 17,392 forms in 69 languages with the second convention. For the skeleton it makes no difference, since both are in the list of marks it ignores. For reading the family it does: "does this language have tone" is a question about the source before it is one about the language.

3.2 And a bug this tone uncovered, with consequences for the whole series

The normalisation table each paper publishes is derived by aligning each form with its skeleton. In this family only 42% of the forms aligned, discarding 115,204, and the footnote attributed them to "prenasals, glides and affricates".

The cause was different: the tone marks were being counted as consonant candidates. The skeleton ignores them correctly, but the alignment script did not have the same list and shifted the whole form from the first mark onward. Fixed by importing the real list instead of repeating it, alignment goes from 42% to 100%.

It is the same bug that already appeared with the compound marker + in the Turkic paper, and it recurred for the same reason: a list repeated in two places. The tables of all eight families in the series were regenerated.

4. The catalogue

10,497 codes over 2,737 concepts, grouping 99,061 of 198,742 forms (49.8%). Of these, 6,302 are binary or ternary and 3,770 are unary: 35.9%.

Arity	Codes	%
1	3,770	35.9%
2	4,979	47.4%
3	1,323	12.6%
4	361	3.4%
5+	64	0.6%

Precision: 3.90% against 0.108% chance = 36× [36–36× across seeds].

And on grouping without the three-language threshold: 61.7% observed against 42.0% null — an advantage of 19.8 points, 1.5× chance, $\kappa = 0.340$.

4.1 The widest-reaching codes are a single consonant, and cross thirteen branches

The twelve of widest reach, unfiltered:

Code	Concept	Languages	Arity	Branches
n	*thou*	165	1	13
s	*die*	163	1	13
n	*I*	154	1	11
m	*mother*	125	1	10
s	*eat*	125	1	9
n	*fish*	124	1	11
n	*five*	115	1	8
s	*ten*	113	1	9
s	*blood*	111	1	8
m	*fire*	105	1	8
n	*two*	105	1	9
s	*salt*	102	1	7

Twelve of twelve are unary, and the first two cross all thirteen branches of the family.

A code crossing thirteen branches looks like the best genealogical candidate imaginable. It is not, and the reason is measured in §6.1: a single consonant crosses branches trivially. The catalogue cannot distinguish a genuine monoconsonantal etymon from a coincidence of availability — the Afro-Asiatic paper showed this with Chadic `l|cow`, which is the lateral `ɬ` of a real etymon — and in this family that indistinction takes the entire top of the list.

5. A registered prediction, and its verdict

The previous paper in this series — Nuclear Trans New Guinea — claimed, over seven families, that the null of grouping explodes below three consonants per word. That finding was built with those seven: it is description, not prediction.

Sino-Tibetan offered the chance to turn it into a prediction, because its profile places it at arity 2.22 — lower than any of the seven — and nothing about it had yet been computed. The prediction was written and dated before the grouping was run (`runs/prediccion-escalon-aridad.md`), with a declared magnitude and not only a direction:

Predicted	Observed	
null > 30%	42.0%	☐
advantage < 2.5×	1.5×	☐
$\kappa < 0.30$	0.340	☐

The principal one is confirmed comfortably. The three families below arity 2.8 give nulls of 37%, 41% and 44%; the five above 3.0, between 4.7% and 15%.

Family	Arity	Groups	Null	κ
Sino-Tibetan	2.22	61.7%	42.0%	0.340
Nuclear Trans New Guinea	2.41	51.9%	37.4%	0.232
Austronesian	2.79	64.8%	44.0%	0.371
Turkic	3.03	56.9%	15.0%	0.493
Afro-Asiatic	3.16	29.7%	4.7%	0.262
Uralic	3.26	33.9%	5.9%	0.298
Indo-European	3.35	44.3%	12.9%	0.361
Pama-Nyungan	3.35	40.2%	14.0%	0.305

And the secondary one fails, which is what the prediction contributes. $\varkappa$ comes out at 0.340, above the declared 0.30. The reason: the null and $\varkappa$ do not move together. This family has a higher null than NTNG — 42.0% against 37.4% — and at the same time a higher $\varkappa$ — 0.340 against 0.232 — because its raw grouping is also higher. The prediction conflated them, and having written it beforehand is the only thing that made that visible: written as a direction alone, it would have been taken as confirmed.

The precision that follows: the step is a regularity of the null, not of performance. Read as "the method performs worse with short words", Sino-Tibetan disproves it — $\varkappa = 0.340$ is better than Afro-Asiatic (0.262) and Uralic (0.298), at the lowest arity of the eight.

5.1 And a claim in the previous paper is falsified

That same paper claimed that the drop from automatic segmentation follows the same axis as the null: the shorter the word, the more it destroys. With all eight families:

Family	Arity	$\Delta\varkappa$ from segmentation
Afro-Asiatic	3.16	−0.004
Turkic	2.85	−0.009
Sino-Tibetan	2.22	−0.020
Uralic	3.27	−0.035
Indo-European	3.35	−0.094
Pama-Nyungan	3.35	−0.102
Nuclear Trans New Guinea	2.41	−0.118
Austronesian	2.78	−0.151

The two shortest-word families sit at opposite ends — 2.22 at −0.020 and 2.41 at −0.118 — and the two longest-word ones in the middle. The association does not exist. With seven families a pattern was visible that the eighth breaks, and the NTNG paper is corrected in both its versions.

6. What the family gives

6.1 The best genealogical material in the series, and its honest decomposition

1,819 codes cross three or more branches, and 657 cross four. That is more than double what the series plan estimated and far above Pama-Nyungan's 349.

But it is not published alone. Decomposed by arity:

Arity	Codes	Cross $\geq$3 branches	Rate
1 (unary)	3,770	1,054	28.0%
2	4,979	719	14.4%
3	1,323	43	3.3%
4+	425	3	0.7%

58% of those that cross are unary, and the crossing rate falls by a factor of nine between unary and ternary. The shorter the code, the more it crosses — exactly what one expects if crossing comes from availability rather than inheritance, and the opposite of what would be expected if it measured genealogical depth.

The defensible figure is 762: the binary and ternary codes crossing three or more branches. It nearly matches the 755 the plan estimated, and it is the one this paper uses.

6.2 And the widest-reaching ones are the textbook comparanda

Concept	Code	Languages	Branches	The four largest, in languages
name	m·n	53	9	Kuki-Chin-Naga 18 · Bodic 18 · Brahmaputran 5 · Burmo-Qiangic 4
smoke	m·k	46	6	Burmo-Qiangic 31 · Kuki-Chin-Naga 9 · Karenic 3 · Digarish 1
stone	l·n	44	10	Kuki-Chin-Naga 19 · Kho-Bwa 7 · Himalayish 6 · Brahmaputran 3

m·n is Sino-Tibetan *miŋ* 'name'; l·n is *luŋ* 'stone'; m·k is *mVk* 'smoke'. The procedure did not look for them: they emerged from requiring a skeleton to repeat for the same concept across three languages. And they cross six, nine and ten branches while being binary, which is what separates them from the unary codes of §4.1: there the reach comes from one consonant being available everywhere; here it comes from two, and from the same two.

The first version of this table said "four branches", and that was my misreading of the run. The `codigos_hermanos.py` table showed only the four largest branches and counted forms rather than languages — which is why m·n appeared with 71 spread over four when it has 53 languages over nine. The script now prints the branch total alongside the four it shows, so the list cannot be read as complete again.

The family's highest-enrichment code spanning seven branches is **s·r·m** for *otter*, 18 languages at 596× — Tibeto-Burman *sram*. It is ternary, crosses seven branches, and appears in no reach ranking: the ordering by reach, which is the annexe's, does not see it.

6.3 Substitution pairs: 4,550 disjoint candidates

Of 5,200 pairs of codes for the same concept differing in exactly one position, 4,550 are disjoint in languages and only those can be proposed as an isogloss. That is triple NTNG's and ten times Afro-Asiatic's, and the reason is structure: thirteen branches.

The widest-reaching ones cluster around the same etyma: m·j ↔ m·n and b·n ↔ m·n in *name*, l·m ↔ l·n in *stone*, b·k ↔ m·k in *smoke*. The candidate substitutions are **j ↔ n**, **b ↔ m**, **m ↔ n**, **l ↔ m** — nasals and labials.

6.4 Emergent classes: the family produces none, and here is why that is said

Pair	Times (discover)	Ratio	Times (check)	Ratio
s ↔ θ	340	3.2×	665	8.8×
v ↔ x	3,230	9.8×	1,652	6.9×
b ↔ pf	178	12.8×	63	6.3×
b ↔ p	6,344	4.2×	5,399	4.9×
q ↔ x	504	4.2×	417	4.8×
f ↔ pf	71	21.6×	11	4.7×
b ↔ ɖ	229	18.2×	41	4.5×

The pairs survive validation on languages never looked at: nine of them exceed 3× in the check set, and the three that rise most — s↔θ, v↔x, b↔pf — spread across different places of articulation. They are not regular correspondences — environment and direction are missing — but candidates, with the place to look.

And with all that, this family is published WITH NO emergent classes. The closure of the pairs surviving above 2× puts 23 of the 39 symbols into a single class, and a class of twenty-three segments informs of nothing. It is the same result as Turkic and Uralic, and it is declared the same way: `ci/clases.py` carries an empty dictionary for Sino-Tibetan, which is a statement and not a gap.

6.4-bis The class that did come out was a fault of ours, and this review withdrew it

The first version of this section published **n** ↔ **ɳ** at 8.1× and 8.5× as the family's second class, and it was the only one surviving the closure: {n, ɳ}, two clean segments against a magma of twenty-three.

It was not a finding about the family. ɳ is the alveolo-palatal nasal in Sinological notation, the same sound as IPA ɲ. `recompute_skeleton.py` merges ɲ into n — and merges ɕ and ʑ, the fricatives of that same series, into s — but its canonical table had no ɳ, ʈ, ɖ or ɭ. The same file declares the equivalence in its OAS class layer, in a comment reading literally "+Sinological alveolo-palatals: ɭ→L, ɳ→N, ʈ/ɖ→T". The decision was taken in one layer and not applied in the other.

Fixed — 3,923 forms across five families, backup `skeleton_bak_alvpal_20260909` — the pair disappears and with it the only clean class. No published family moves: the 20 Uralic forms affected belong to the doculect phase 0 already excluded.

What changed in this paper's figures on re-running every phase: 10,511 → 10,497 codes, coverage 49.5% → 49.8%, grouping 61.4% → 61.7%, null 41.5% → 42.0%, and $\varkappa = 0.340$ in both. The verdict on the §5 prediction does not move.

7. Limits

Geography is weak, for a structural reason:

Codes shared	Pairs	Mean distance	vs chance
1	8,471	1,192 km	1.07
2	4,122	1,124 km	1.01
3–5	4,178	931 km	0.84
6–10	1,433	643 km	0.58
11–20	776	468 km	0.42
21+	1,146	431 km	0.39

Monotone, but falling only to 0.39 against NTNG's 0.09. This family stretches from Pakistan to north-eastern China; NTNG occupies a continuous area with a median of 95 km from its centre. Compactness and geographic signal are the same property seen twice, and this family confirms it from the other side.

A single macroarea: §6.1-bis does not exist.

Tone does not enter the skeleton, and 137 of 267 languages mark it. In a family where it is contrastive, the skeleton sees less than the language distinguishes, and no figure in this paper compensates for that.

35.9% of the catalogue is a single consonant, and the catalogue cannot distinguish the genuine mono-consonantal etymon from coincidence. The unary codes are counted and not read.

Burmo-Qiangic is 38% of the languages. It is the best balance in the series, and it is still a branch that weighs more than the others.

8. Conclusion

The series plan called this family "the grand prize" and deferred it for its cost. Both were true, and both are now measured: thirteen well-populated branches give 762 binary and ternary codes crossing three or more — the best genealogical material of the eight families — and 2.22 consonants per word give the highest collision and a null of 42.0%.

That the widest-reaching codes are `m·n` for *name* across nine branches and `l·n` for *stone* across ten, with no reconstruction consulted, is the same kind of convergence Afro-Asiatic gave with the triliteral root: the catalogue contains the established correspondences where there are any.

And the methodological lesson comes from the parts that did not go as expected, of which there were three. The prediction registered before computing confirmed the principal claim and failed on a secondary one, and that failure separated two things the previous paper conflated: the null follows arity, κ does not. This family falsifies a claim in that same paper about segmentation. And its own adversarial review withdrew an emergent class that was an omission in our normalisation table, and corrected a "four branches" that came from reading a truncated list as complete.

None of the three would have been visible without doing the eighth family. A series of eight is not eight times a series of one: it is the only way for the seventh to be wrong about something checkable.

References

Brown, C. H., Holman, E. W., Wichmann, S. and Velupillai, V. (2008). Automated classification of the world's languages. *Sprachtypologie und Universalienforschung* 61.

Dolgopolsky, A. B. (1964). Gipoteza drevnejšego rodstva jazykovyx semej Severnoj Evrazii s verojatnostnoj točki zrenija. *Voprosy Jazykoznanija* 2.

Hammarström, H., Forkel, R., Haspelmath, M. and Bank, S. *Glottolog*. Max Planck Institute for Evolutionary Anthropology.

Hubert, L. and Arabie, P. (1985). Comparing partitions. *Journal of Classification* 2(1). — [the chance-correction this paper uses as κ]

Jäger, G. (2018). Global-scale phylogenetic linguistic inference from lexical resources. *Scientific Data* 5.

List, J.-M. (2012). SCA: phonetic alignment based on sound classes. In *New Directions in Logic, Language and Computation*. Springer.

List, J.-M. (2014). *Sequence Comparison in Historical Linguistics*. Düsseldorf University Press.

List, J.-M., Forkel, R., Greenhill, S. J. *et al.* (2022). Lexibank, a public repository of standardized wordlists. *Scientific Data* 9.

List, J.-M., Cysouw, M. and Forkel, R. (2016). Concepticon: a resource for the linking of concept lists. *LREC*.

Matisoff, J. A. (2003). *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. University of California Press. — [the comparanda of §6.2, which this work did not consult]

Sagart, L., Jacques, G., Lai, Y. *et al.* (2019). Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *PNAS* 116(21).

Toledo Martínez, A. (2026). *The catalogue of consonantal codes: a comparative method without genealogy, cognacy or reconstruction*. — [the method of the series, described once]

Toledo Martínez, A. (2026). *Consonant Codes in 196 Trans-New Guinea Languages.* — [the previous paper, whose §5 this work tests and whose §5.1 it corrects]

Toledo Martínez, A. (2026). *Integrative Corpus: a multi-layer lexical database of 225 macrosystems, with row-level provenance and its verification apparatus.* — [the data] · https://doi.org/10.13140/RG.2.2.31244.68486

Turchin, P., Peiros, I. and Gell-Mann, M. (2010). Analyzing genetic connections between languages by matching consonant classes. *Journal of Language Relationship* 3.

Data and reproducibility

Catalogue	`familias/sino-tibetan/output/catalogo-sino-tibetan.tsv` — 10,497 rows
Per-form detail	`…-lenguas.tsv` — 89,007 rows
Hygiene (§2.2)	`pipeline/fases/fase0_higiene.py` → `runs/fase0-higiene.md`
Branch label (§2.1)	`corpus_integrativo/ci/decisiones.py::SUBGRUPOS`
Normalisation table (§3.2)	`pipeline/fases/tabla_normalizacion.py` → `runs/normalizacion.md`
Loading, population, warnings	`cod/familia.py`
Catalogue (§4)	`pipeline/fases/catalogo_familia.py`
Grouping and nulls (§4, §5)	`pipeline/fases/agrupamiento_familia.py` → `runs/agrupamiento.md`
Registered prediction (§5)	`runs/prediccion-escalon-aridad.md` — dated before the grouping
κ (§5)	`cod/metricas.py::kappa`
Automatic segmentation (§5.1)	`pipeline/fases/fase2_despiece.py` → `runs/fase2-despiece.md`
Branch crossing (§6.1)	`pipeline/fases/fase6_por_rama.py` → `runs/fase6-por-rama.md`
Sibling codes (§6.3)	`pipeline/fases/codigos_hermanos.py` → `runs/codigos-hermanos.md`
Emergent classes (§6.4)	`pipeline/fases/fase6_clases_emergentes.py` → `runs/fase6-clases.md`
Geography (§7)	`pipeline/fases/codigos_vs_parentesco.py` → `runs/codigos-parentesco.md`

Every figure comes from those scripts over the same population declared in §2, with the same exclusions — which here are none. None was written by hand.

Appendix A · One hundred codes, with commentary

The catalogue holds 10,497 codes and the complete data are in `familias/sino-tibetan/output/`. What follows is a readable sample: the hundred binary and ternary codes of widest reach, each with up to sixteen languages and their words.

Three things about the presentation. Languages are spread along the alphabetical list, not taken from its head. Long names are trimmed; the forms are not altered. And the thematic grouping is a reading order declared by hand, not a semantic classification.

This appendix carries no emergent-code column, and here that is a result rather than a precaution. The phase-6 closure puts twenty-three of the family's thirty-nine symbols into a single class, and a class of twenty-three segments informs of nothing. §6.4 publishes it that way: this family produces no emergent classes.

What this appendix selects, measured

Sorting by reach is not neutral, and in this family it biases in four directions at once. All four come from `pipeline/tools/audita_anexo.py`:

	catalogue	this appendix
binary codes	47.4%	97%
ternary codes	12.6%	3%
Burmo-Qiangic as dominant branch	45.0%	55%
Sinitic as dominant branch	18.3%	22%
crossing four branches or more	6.3%	32%
mean enrichment	172×	87×

Ninety-seven of a hundred are binary, the opposite of what happened in the Afro-Asiatic appendix, where ternaries were over-represented. At 2.22 consonants per word, reach lives in the binaries and ternaries barely reach ten languages.

Seven of the thirteen branches have not a single code here — Karenic, Brahmaputran, Himalayish, Digarish, Nungish, Kman-Meyor, Macro-Tani. A paper whose central argument is thirteen branches publishes an appendix showing six.

And at the same time it over-represents the genealogically interesting: codes crossing four or more branches are 6.3% of the catalogue and 32% of the appendix, five times more. The bias runs both ways and for the same cause: reach.

Enrichment halves. Reach and enrichment are bought with one another: a code reaching many languages is usually made of frequent consonants, which is exactly what enrichment penalises. The hundred widest-reaching codes on this list are, by construction, among the least distinctive in the catalogue.

Audit: the normalisation table holds up two thirds of this appendix

Sixty-eight of the hundred codes have at least one position where the table merges two segments that three or more languages write differently. In the Afro-Asiatic appendix it was nine of ninety-one. The difference is not one of method but of family: here `s` gathers the entire sibilant-affricate complex — `ts tsʰ tɕ tɕʰ ɕ ʃ ʂ tʂ dʑ dʒ ʐ` — and at 80,694 positions is the most loaded symbol in the corpus, while `n` gathers `ŋ ɲ ɳ ɴ`.

Three cases need separating, because they are not the same:

1 · Aspiration. `n·k|nose` is `k` in 18 languages and `kʰ` in 15; `t·n|drink`, `t` in 10 and `tʰ` in 9. The skeleton has never distinguished aspiration, in any family of the series. It is an instrument decision declared from the start, not an accident of this family.

2 · Nasal place. `m·n|name` — the emblematic code of §6.2 — carries `ŋ` in 35 languages and `n` in 19. This is the case to read slowly: the final alternation *-ŋ* *-n* in 'name' is a real Sino-Tibetan correspondence, and a comparatist would unify those forms just as we do. The difference is how: they by demonstrating the correspondence; we because the table writes them the same before looking. The result coincides and the grounding does not, and that distinction is the whole paper.

3 · Sibilant place, which is the serious case. `n·s|medicine` is `tsʰ` in 34 languages and `tɕʰ` in 15; `s·n|onion`, `tsʰ` in 15 and `tʃ` in 11; `s·n|sing`, `tsʰ` in 9 and `tʂʰ` in 9. The dental / palatal / retroflex opposition is what Sinitic dialectology is built on, and the table erases it entirely. These codes are not false — they group forms a specialist would also relate — but their reach is inflated by a distinction the skeleton does not carry.

And there is a fourth case that this very review created. `n·s|twenty` now comes out with `n` in 32 languages and `ɳ` in 16, and `s·n|twelve` with `n` in 23 and `ɳ` in 16. Before, they did not: `ɳ` was its own canonical symbol and those codes were split in two. The correction of §6.4-bis joined them, and in joining them made them dependent on the table. Fixing an inconsistency does not make the normalisation innocent: it makes it consistent, which is another thing.

And what this appendix shows without meaning to: enrichment does not rank by antiquity

The ten highest-enrichment codes of this appendix, in order:

Code	Concept	Languages	Enrich.	What it is
`n·s·t`	twenty one	24	1545×	compound numeral: "two-ten-one"
`b·d·m`	groundnut	18	1086×	*badam*, Persian 'almond' — a loan
`s·r·m`	otter	18	596×	*sram, a genuine Tibeto-Burman etymon
`p·t`	grape	20	380×	Chinese ?? *pútáo*, itself an Iranian loan
`p·l`	glass	19	330×	Chinese ?? *bōli*, from Sanskrit *sphaṭika*
`s·g`	plough	19	258×	material culture
`d·l`	wing	30	170×	—
`m·l`	sorghum	18	156×	a crop
`z·m`	daughter	23	148×	—
`n·m`	younger sister (of man)	24	130×	—

The first four are a compound numeral, a Persian loan, a real etymon and an Iranian loan via Chinese, in that order and with nothing separating them. Enrichment measures how distinctive a skeleton is, and a word that has just arrived is maximally distinctive: it is the same everywhere because it arrived whole. It is the cassava case of the Afro-Asiatic appendix, multiplied here.

`s·r·m|otter` is the one to look at twice. It is *sram* — Old Tibetan *sram*, Jingpho *saram*, Lushai *sa+ṛam* — a fully-fledged Tibeto-Burman etymon, ternary, crossing seven branches. And it sits at position 94 of this appendix by reach, with eighteen languages. Ordering by reach barely sees it.

The compound numerals, which are not root codes

Four of the first forty codes are numerals of the decade: `n·s|twenty` (58 languages, the first of the whole appendix), `s·n|twelve` (50), `s·t|eleven` (34), `n·s·t|twenty one` (24). Read their blocks: *ni-tsi*, *tsʰe-ɳi*, *tsʰi-ti*, *ni-tsi-thi*. They are "two-ten", "ten-two", "ten-one", "two-ten-one".

What those languages share is not a root but an arithmetic, and the roots composing it — 'two', 'ten', 'one' — are already in the catalogue on their own account as unary codes. The procedure counts the compound numeral as one more code, and has no way of knowing it is the sum of two others. It is the same boundary §5.1 measures for automatic segmentation, seen from the other side.

And one last thing, visible in the section index

Fifty-five of the hundred codes fall under "Other". This appendix's thematic list was declared by hand for families with Swadesh-type basic vocabulary, and the Sino-Tibetan Lexibank lists are not that: they carry medicine, cotton, onion, gold, shoe, chili pepper, walnut, board, bracelet, copper, ring, glass, steel, plough, sorghum, groundnut. More than half of this appendix is material-culture and trade vocabulary, and that is the concrete shape the loan problem takes here. It is not a defect of the corpus: it is what has been collected for this part of the world, and it must be read knowing so.

Reverse index: codes recurring across concepts

Twenty-three codes of the appendix cover seventy-nine of its hundred entries. It is by far the highest figure in the series — Turkic had ten concepts under a single code — and it is §3's 71.6% collision put on display: four out of five blocks that follow repeat a skeleton already seen.

`s·n` takes sixteen concepts on its own: twelve, onion, gold, shoe, new, firewood, long, sing, seven, day before yesterday, wood, heart, heavy, mountain, near, dirty. No semantic reading sustains that, and none is attempted: `s` and `n` are the two most frequent symbols of the inventory, and their combination is the most available skeleton in the family.

It is worth separating from `d·z`, which covers eat, cooked rice, ride, raw and thorn with enrichments from 58× to 119×. There the skeleton is rare and the concepts touch — eat, rice, raw — so it is plausible they are one family of words. The same row format and two different phenomena, and the code alone does not say which.

Code	Concepts it appears with, and their enrichment
`s·n`	twelve (50, 38×) · onion (32, 24×) · gold (30, 12×) · shoe (30, 16×) · new (27, 7×) · firewood (24, 10×) · long (24, 6×) · sing (24, 10×) · seven (21, 6×) · day before yesterday (20, 18×) · wood (20, 9×) · heart (20, 5×) · heavy (20, 5×) · mountain (20, 5×) · near (20, 5×) · dirty (18, 7×)
`d·z`	eat (34, 58×) · cooked rice (27, 86×) · ride (24, 119×) · raw (22, 112×) · thorn (18, 54×)
`s·m`	three (34, 21×) · dream (26, 23×) · old person (25, 56×) · walnut (23, 42×) · flour (21, 27×)
`s·s`	sand (28, 10×) · tree (26, 9×) · claw (25, 16×) · dew (19, 12×) · flea (18, 11×)
`m·n`	name (53, 44×) · cat (22, 26×) · dream (22, 29×) · no (21, 47×)
`n·m`	sun (40, 41×) · heart (39, 42×) · younger sister (of man) (24, 130×) · sky (23, 29×)
`k·n`	root (21, 13×) · steel (20, 111×) · empty (19, 34×) · dry (18, 12×)
`n·s`	twenty (58, 57×) · medicine (45, 48×) · rainbow (18, 14×)
`n·n`	thou (54, 34×) · breast (30, 24×) · silver (20, 20×)
`s·t`	eleven (34, 45×) · kill (27, 20×) · ashamed (24, 79×)
`s·p`	board (23, 58×) · host (20, 69×) · lung (20, 20×)

Numerals

The appendix's four numerals are of the decade, not of the unit, which sets them apart: `n·s|twenty`, `s·n|twelve`, `s·t|eleven` and — under "Other", because of its concept label — `n·s·t|twenty one`.

The family's unary codes do carry the low numerals — `n|five`, `s|ten`, `s|three` — and do not appear here because the appendix takes only binaries and ternaries. The result is a numeral section made entirely of compounds, with this appendix's header explaining why.

`s·m|three` is the exception: *sum* in Achang, Bisu, Chepang, Tibetan, Tamang, Old Burmese. It is the real Tibeto-Burman numeral, across five branches.

`n·s|TWENTY` · binary · 58 languages · enrichment 57×

Alo ȵi⁵⁵tse³³	Black Lahu ni⁵³tshi³³	Guanyinqiao Khros nɛ⁵⁵sə³³
Hlepho n²¹ʦɪ³³	Kazhuo ŋ³¹tsi³³	Lavu ȵi²¹tsi⁵⁵
Lisu ȵi³¹◦tsi̱³³	Mondzi ni⁴⁴si⁴⁴	Nersu ȵi²¹tsɤ¹³
Njorogs Khroskyab nei⁵⁵sə³³	S. Muji n²¹ʦʰə³³	Sherpa ɲiɕju
TibetoKinauri_Bun nidza	Xinlong Queyu nyɕy	Yi ni²¹tsɯ⁵⁵
Yijiacun ŋ²¹tɕi²⁴		

`s·m|three` · binary · 34 languages · enrichment 21×

Achang sum³¹	Bisu sum55	Chepang sum	Gyalsumdo sum
Jirel sum	Kaike sum	Laowu Nisu sə³³mu³³	Lowa sum
Monpa Kalaktang sam	Monpa Tembang sam	Old Burmese sum:	S.kong sem55
Tamang 'som	Tibetan_Lhasa sum⁵⁵	TibetoKinauri_Bya səm	Tsum sum

`p·n|five` · binary · 26 languages · enrichment 29×

Angami Khonoma pengu	Angami Kohima pengou	Ao Mongsen phanga	Chepang po+ŋa	
Chokri püngu	Kaang Kruk pɑŋɑ	Kezhama pangu	Lushai pa+ŋa	
Maring phanga	Miji Nafra puŋu	Mkaang pa̤ŋa	Mro-Khimi Areong	pə◦ŋ
Nruanghmei pang	Rawngtu Weilong pa̤ŋa	Sema pongu	Tangkhul pha◦nga	

`s·n|SEVEN` · binary · 21 languages · enrichment 6×

Dimasa sini	Garo_Garo sini	Guanyinqiao Khros sɲi⁵⁵	Kayaw sɔ³³ne⁵⁵
Liangmai chania	Mao cha◦ni	Maram sina	Nruanghmei cünei
Rengma tsanü	Sema tsi◦ni	Stau sɲe	Suansu ts_o_n_i
Tangkhul shini	Xinlong Queyu sna⁵⁵	Zeme sena	Zhaba ʂnɛ⁵⁵

Body parts

Ten codes, and the distribution is revealing. `m·k|eye` crosses eight branches; `n·p|ear`, two, at 84× enrichment; `n·k|nose`, four; `d·l|wing`, one alone at 170×.

High enrichment with a single branch is the signature of the local, not of the ancient. `d·l|wing` is entirely Burmo-Qiangic: inside that branch the skeleton is distinctive, and outside it does not appear. `m·k|eye`, at 49×, crosses Kuki-Chin-Naga, Himalayish, Bodic, Sinitic, Brahmaputran, Burmo-Qiangic, Macro-Tani and Karenic — half the family.

`n·p|ear` · binary · 41 languages · enrichment 84×

Alo nu²¹pɒ³³	Azha no55pu33	Diaocao na²¹pa⁵⁵	Eryuan ɲɤ³³pi²¹
Hlepho na⁵⁵pə³³	Kua-nsi na21po33	Lipo no55pɑ33	Lolopo no55pɑ33
Narsu no⁵⁵pa⁵⁵	Neshu no⁵⁵pa⁵⁵	Polo nɒ33pɒ33	Sanni na²¹pu³³
Sodi nɯ²¹po³³	Wadu nɐ⁵⁵_pɐ⁵³	Yi no⁵⁵p□ ³³	Yijiacun na²¹pa²⁴

n·m|heart · binary · 39 languages · enrichment 42×

Alo ȵi²¹mu²¹	Azha ni⁴⁴mɔ³³	Black Lahu ɔ³¹ni³³m□ ³³	Gepu ni³³mo²¹
Haoni nɯ33mɔ33	Kua-nsi ni³³mo³³	Lavu ni33mu33	Lolopo ȵi33mo33
Namuyi ȵi³³_mi⁵⁵	Nasu ȵi2mu2	Phuza ȵi⁵⁵mɔ³³	S. Muji ȵi³³ma²¹
Sanni ni⁵¹ma⁵¹	Shuizhuping ni⁴⁴ma³³	Xinlong Queyu nəmi	Yi ni³³mo³³

m·k|eye · binary · 33 languages · enrichment 49×

Asho mɛiˌkʰu	Chaozhou mak⁵	Daai Kal-im mik	Kaang Kruk mik
Kiranti_Hayu mek	Konyak mek	Maram a◦mek	Mkaang mik
Mzieme miak	Ng'ga Kant-thayun mik	Nruanghmei mik	Phom mük
Rondu Balti mik	Sinitic_Jieyang mak⁵	Tangkhul mik	TibetoKinauri_Bun mik

n·k|nose · binary · 33 languages · enrichment 47×

Burmese n̥ɑ²²khɑũ⁵⁵	Guiqiong ȵo⁵⁵_kũ⁵³	Hlepho na³³kɯ⁵⁵	Kayah Monu nɔ́-kʰì
Kezhama nhu◦ka	Lavu nu55khu55	Mangdi no⁵⁵kʰu³³	Monpa Tembang nak
Neshu no⁵⁵kǫ²¹	Phuza na⁵⁵kɯ³³	Rouruo na35ka35	Samu nau53khau21
Shuizhuping na⁴⁴kʰ꒡³³	WB nɑ²²_khɑũ⁵⁵	Xuzhang nɨ꒡⁴⁵kʰɤ꒡³³	Yi no⁵⁵kǫ²¹

n·n|BREAST · binary · 30 languages · enrichment 24×

Azha na⁴⁴na²¹	Burmish_Xiandao ɲɔn³⁵	Ersu ȵo⁵⁵_ȵo⁵⁵
Guangzhou nin⁵⁵	Guiqiong ni⁵⁵_ni⁵⁵	Kangding Minyag 2nønø
Mazur Stau nene	Njorogs Khroskyab nu³³nu⁵⁵	Nusu nɯ⁵⁵_nɯ³¹
Nyagrong Minyag nɯ³³_nø⁵³	Rahung nyŋ	Sartang Khoina nyŋ
Shergaon ɲuŋ	Siyuewu Khroskyab nunu	Suzhou nɑ¹³⁻²² nɑ¹³⁻³¹
Xinlong Queyu nənə		

s·w|blood · binary · 30 languages · enrichment 31×

Achang sui³¹	Bai suɑ⁴⁴	Bai suɑ⁴⁴	Eastern Kayah Li swì
Enqi sua55	Heqing sʰua⁴⁴	Lahta Kayan ʃwì	Loudi sue¹³
Northern Pwo swi³³	Northern Sgaw swi³¹˜	Old Burmese swe:	Qiliqiao ʃua⁴⁴
Tuoluo suɔ42	Xiangyun sua⁴⁴	Yunlong sua⁴⁴	Zhaozhuang Bai ɕua⁴⁴

d·l|wing · binary · 30 languages · enrichment 170×

Alo du²¹lą⁵⁵	Azhe do22lɪ21	E Lolo du³³lą²¹	Jinggu Limi du²⁴le²¹
Laowu Nisu du²¹la¹³	Lipo du33le33	Lolopo du33le21	Mondzi do¹³lɔ⁵³
Naso du²¹lą⁵⁵	Nasu du²¹lą⁵⁵	Neshu do²¹lę²¹	Phuza ɛ³³dɯ³¹li⁵⁵
Sanni du³³lę⁵⁵	Toloza du²¹lę²¹	Yangliu dɤ³³lai⁵³	Yi do³³lę²¹

l·p|hand · binary · 20 languages · enrichment 38×

Axi le21pu55	Azha le33phe44	Batang Tibetan la¹³◦pa⁵³
Bisu la31pu31	Jinggu Limi lɛ³³pʰɛ⁵⁵	Jinuo la55pu44
Kazhuo la53pha55	Kua-nsi lą²¹pʰa³³	Laowu Nisu la¹³pʰa⁵⁵
Lavu la21pha33	Lawu lą³¹pa³³	Lisu lɛ̱³¹◦phɛ³⁵
Sani le2phe44	Xinlong Queyu le⁵⁵pa⁵⁵	Yi li²¹_phi³³
Yijiacun lą²¹pʰe³³		

s·n|heart · binary · 20 languages · enrichment 5×

Beijing ɕin⁵⁵	Caijia saŋ33	Changsha ɕin³³	Chengdu ɕin⁴⁵
Fuzhou siŋ⁵⁵	Guilin ɕiŋ⁴²	Ha_erbin ɕin⁴⁴	Loudi ɕin⁴⁴
Nanchang ɕin⁴²	Nanjing ɕin³¹	Rongcheng sin⁵²	Sinitic_Beijing ɕiən⁵⁵
Suzhou sin⁴⁴	Taiyuan ɕin²¹	Tshobdun sni	Wenzhou saŋ⁴⁴

n·b|nose · binary · 20 languages · enrichment 53×

Axi no33bo21	Azhe nɯ33bɯ21	Bantawa nabu	Gepu no³³bo̱²¹
Kua-nsi na21bi55	Laowu Nisu nu³³bə²¹	Lipo nu55bi33	Lisu nɑ̱³³◦bi³³
Mondzi nɔ¹³bɪ²¹	Nasu nu³b	Nersu no¹³bɪ²¹	Neshu no⁵⁵bɤ̱²¹
Sanni na³³bɯ³³	Sodi na²¹bɪ̱⁵⁵	Yi no³³be̱³³	Yi n̥a²¹bi⁵⁵

Kinship and person

Four codes, the first two being those the paper uses: `n·n|thou` in 54 languages and eight branches, and `m·n|name` in 53 and nine. The other two are `z·m|daughter` (148×) and `m·m|mother` (75×).

`m·m|mother` deserves the standing warning. A repeated labial nasal for 'mother' is the textbook example of what proves nothing: it appears in unrelated families for reasons of language acquisition. The catalogue groups it by the same criterion as `m·n|name` and has nothing with which to separate them.

n·n|THOU · binary · 54 languages · enrichment 34×

Achang nuaŋ⁵⁵	Bisu naŋ33	Chin_Hakha n.ă/a.ŋ	Dimasa nung
Jingpho naŋ³³	Kangding Minyag 1nəne	Lashi naŋ³¹	Lushai naŋ
Maring nang	Monpa Kalaktang nan	Mro-Khimi Areong naŋ	Nruanghmei nang
Rahung naŋ	S.kong naŋ55	Shanghai noŋ¹³	Uppu naŋ

m·n|name · binary · 53 languages · enrichment 44×

Asho miŋ	Daai Kal-im miŋ	Ekai Dalet Ywa Ha ɑ.miŋ	Jirel min
Khumi əm□ n	Konyak min	Manange 1miŋ	Maru maŋ³¹
Monpa Kalaktang min	Ng'ga Kant-thayun _miŋ	Old Tibetan miŋ	Rabha muŋ
Sinitic_Longgang miaŋ²¹	Tani_Bokar a+min	TibetoKinauri_Bun min	Uppu amiŋ

z·m|DAUGHTER · binary · 23 languages · enrichment 148×

Azha ze44me33	Black Lahu ɔ³¹z□ ⁵³mi⁵³	Diaocao zɔ²¹mɛ²¹	Hani za31mi31
Haoni zɔ31mi31	Kazhuo za31m31	Kua-nsi zu²¹ma²¹	Lawu za²¹mi²¹
Lisu zɑ̱³¹mɯ̱³¹	Lolopo zo21mæ21	Samu zɑ21mi21	Shuizhuping za²¹mɛ²¹
Xuzhang za²¹mɛ̃²¹	Yangliu za²¹mɛ̃²¹	Yi zɛ²¹mɛ²¹	Yijiacun zo³³mɛ²¹

m·m|MOTHER · binary · 23 languages · enrichment 75×

Beijing ma⁵⁵ma⁰	Changsha ma³³ ma³³	Chengdu ma⁴⁴ma⁴⁴	Hangzhou mɑ³³mɑ³³
Jixi mɔ³¹ mɔ⁰	Khaling mɛm	Kiranti_Bahing mama	Kunming ma⁴⁴ma⁴⁴
Loudi m̩¹¹ ma⁴²⁻¹	Mandarin Chinese ma1ma	Nanjing mɑ³¹ mɑ⁰	Shanghai m̩⁵³ma¹³
Shexian m̩³⁵ma³¹	Sinitic_Beijing ma⁵⁵ma⁰	Xiamen ma⁵⁵ma⁵⁵	Yangzhou mɑ³¹mɑ³¹

Nature and sky

Fifteen codes, the second most populated section, and the one carrying the two comparanda of §6.2: `l·n|stone` across ten branches and `m·k|smoke` across six.

It also carries two different codes for 'fire' — `m·t` in 30 languages, all Burmo-Qiangic, and `x·w` in 21 spread between Macro-Bai and Sinitic — which are not variants of one another: they are two etyma holding different zones of the family.

`m·k|smoke` · binary · 46 languages · enrichment 83×

```
Asho mei.kʰɯ     Bisu mi31khau31            Daai Kal-im mɑi.kʰu
Gepu mə55khə33   Kangding Minyag mu³³khə⁵³   Kua-nsi mu33ku33
Lisu mu³¹◦khu³¹  Maang mi33khɑu35           Nasu mu33khɯ33
Naxi mɯ⁵⁵_khɯ³¹  Nyagrong Minyag mŭ³⁵_khɯ⁵⁵  Rouruo mi33khɯ33
Stau mkʰə        Thakali mi◦ku              WB mi⁵⁵+kʰo⁵⁵
Yi mu³³ku³³
```

`l·n|stone` · binary · 44 languages · enrichment 41×

```
Ao Chungli lung      Asho hliŋ   Chin_Hakha l.û/u.ŋ     Deng lɑuŋ³⁵
Jerigaon lyŋ         Kaike lung  Kiranti_Kulung luŋ     Lahta Kayan lòu̯ŋ
Lushai luŋ           Mkaang lúŋ  Mro-Khimi Areong  luŋ  Nruanghmei lung
Nungic_Dulong luŋ⁵⁵  Rupa lyŋ    Sherdukpen Rupa luŋ    Uppu lï:ŋ
```

`n·m|sun` · binary · 40 languages · enrichment 41×

```
Achang ni³¹+mɔ³¹         Bantawa nam             Chug Parchu na-mi
Gyalsumdo ɲima           Jirel ɲima              Kaike nhi◦mā
Kiranti_Hayu no+mo       Kiranti_Limbu nam       Monpa Kalaktang ŋam
Monpa Tembang ŋam        Naxi ɳi³³_me³³          Nungic_Dulong nɑm⁵³
Rondu Balti ñimɑ         Sartang Khoitam nə-mi:  Tibetan_Lhasa ɳi¹³◦mə⁵⁵
Xinlong Queyu ɳi⁵⁵_mɯ³³
```

`m·t|fire` · binary · 30 languages · enrichment 49×

```
Alo mu³³tu̱⁵⁵     Axi mu33tɯ55      Azhe mɯ55tɯ55        Gepu m³³tu̱²¹
Hlepho mi⁵⁵to³³  Kua-nsi mu21tu55  Lope mu55tu55        Narsu mɛ³³tu̱²¹
Nasu mo³³tɑ̱⁵⁵    Nersu mi³³te¹³    Nipu mu³³_tɔ⁵⁵       Polo mu13tu55
Sani m²¹tu⁵⁵     Sodi m²¹tɑ̱⁵⁵      Up: Phola mi³¹tʰɛ³³  Yi mu²¹tu⁵⁵
```

`l·m|ROAD` · binary · 28 languages · enrichment 67×

```
Alike Tibetan lam   Bantawa lam          Bugun Kaspi ləmua     Bugun Singchung ləmua
Chin_Hakha l.â/a.m  Daai Kal-im lɑm      Jingpho lam³³         Khaling lēm
Kiranti_Hayu lom    Kiranti_Limbu lam    Monpa Tembang lam     Ng'ga Kant-thayun lɑm
Old Tibetan lam     Rawngtu Weilong lɑm  Sherdukpen Rupa ləmu  Uppu lɑm
```

`s·s|sand` · binary · 28 languages · enrichment 10×

```
Achang si⁵ ʃɔ³¹    Bai so⁵⁵_tsi⁴⁴          Black Lahu s□ i³⁵si¹¹  Cebu sa⁵³tsi²¹
Chengdu sa⁴⁵ sa⁵⁵  Chokri tsüshe           Guilin sɑ⁴² tse²³⁻³³   Ha_erbin ʂa⁴⁴ tsɪ⁰
Jixi so⁵⁵ tsɪ⁰     Lahu s□ i³⁵_si¹¹        Loudi sɔ⁴⁴ tsɪ⁰        Nanchang sa⁴² tsɪ⁰
Qiang ʂɑ⁵⁵_tsɪ³¹   Shuizhuping ʂa³³tsɪ³³  Xi_an sɑ²¹ tsɪ⁰        Yi sɑ³³tsɪ³³
```

n·m|sky · binary · 23 languages · enrichment 29×

Chug Parchu na+mu	Gyalsumdo nam	Jerigaon nə+my:	Kagate nam
Kiranti_Kulung nam	Lowa nam	Monpa Kalaktang ŋam	Monpa Tembang ŋam
Rahung nə+my:	Rondu Balti n̥ʌm	Sartang Darbu 1 nymy	Sartang Khoina na+my:
Shergaon ni+mi:	Sherpa nam	TibetoKinauri_Bun nam	Tsum n̪am

p·n|ice · binary · 22 languages · enrichment 57×

Beijing piŋ⁵⁵	Changsha pin³³	Chengdu pin⁴⁴	Guangzhou p.ɪ.ŋ.⁵³₅₅
Guilin piŋ⁴²	Ha_erbin piŋ⁴⁴	Mandarin Chinese bing1	Meixian pen³³
Nanjing pin³¹	Rongcheng piŋ⁵²	Shenyang piŋ³³	Sinitic_Longgang pien⁴⁵
Suzhou pin⁴⁴	Taiyuan piŋ²¹	Xi_an piŋ²¹	Xiamen piŋ⁵⁵

l·n|cold · binary · 21 languages · enrichment 22×

Beijing ləŋ²¹⁴	Changsha lən⁴¹	Guangzhou laŋ²³	Ha_erbin ləŋ²¹³
Hefei lən²⁴	Jinan ləŋ⁴⁵	Mandarin Chinese leng3	Nanchang laŋ²¹³
Nanjing lən¹¹	Rongcheng ləŋ²¹⁴	Sinitic_Beijing l.ɤ.ŋ.²¹⁴	Sinitic_Longgang laŋ²¹
Suzhou laŋ³¹	Xi_an ləŋ⁵³	Xiamen liŋ⁵³	Yangjiang laŋ²¹

x·w|fire · binary · 21 languages · enrichment 76×

Bai xui³³	Bai xui³³	Beijing xuo²¹⁴	Enqi xui22
Guilin xuo²³	Ha_erbin xuo²¹³	Heqing xʰue³³	Jinan xuɤ⁴⁵
Lanping xue⁴⁴	Mandarin Chinese huo3	Rongcheng xuɔ²¹⁴	Taiyuan xuɤ⁵³
Tuoluo xue33	Xiangyun xue³³	Yunlong xue³³	Zhaozhuang Bai xui³³

s·n|WOOD · binary · 20 languages · enrichment 9×

Ao Chungli süng	Ao Mongsen a◦sung	Bantawa sɨŋ	Chug Parchu ɕiŋ
Khoirao a◦sing	Kiranti_Limbu siŋ	Kiranti_Thulung sʌŋ	Lotha o◦tsang
Monpa Tembang ɕiŋ	Old Tibetan ɕiŋ	Rengma süng	Sangtam sing
Tibetan_Lhasa ɕiŋ⁵⁵	TibetoKinauri_Bun ɕiŋ	TibetoKinauri_Bya ʃin	Tibeto_Kinauri_Ro siŋ

s·n|mountain · binary · 20 languages · enrichment 5×

Asho swŋ	Beijing ʂan⁵⁵	Chang shong
Changsha san³³	Ekai Dalet Ywa Ha swŋ	Fuzhou saŋ⁵⁵
Guangzhou s.a.n.⁵³₅₅	Ha_erbin ʂan⁴⁴	Lahta Kayan sòu̯ŋ
Mandarin Chinese ʂaɪn	Maring ching	Meixian san³³
Nanchang san⁴²	Nruanghmei ching	Rongcheng ʂen⁵²
Sinitic_Beijing ʂan⁵⁵		

f·n|wind · binary · 20 languages · enrichment 101×

Beijing fəŋ⁵⁵	Chengdu foŋ⁴⁴	Guangzhou fʊŋ⁵³	Guilin fɤn⁴²
Hefei fəŋ²¹²	Jinan fəŋ²¹³	Kunming foŋ⁴⁴	Mandarin Chinese fəŋ
Nanchang fuŋ⁴²	Nanjing fən³¹	Rongcheng fəŋ⁵²	Shenyang fəŋ³³
Sinitic_Longgang fʊŋ³³	Suzhou foŋ⁴⁴	Taiyuan fəŋ²¹	Xi_an fəŋ²¹

m·h|RAIN (PRECIPITATION) · binary · 19 languages · enrichment 58×

```
Alo mɔ³³hu³³      E Lolo a⁵⁵me²¹ho³³  Gepu mu44ho33     Jinggu Limi a⁵⁵mu²¹ho³³
Kua-nsi ma33ha33  Lisu mɯ³¹◦hɑ̃~³³    Mondzi mo⁴⁴hɔ¹³   Naso mu³³ho³³
Nasu mo³³ho³³     Nersu mi³³ho¹³      Nipu mu³³_ho³³    Shuizhuping a⁵⁵mu²¹hã³³
Toloza mu³³hɑ³³   Yangliu a⁵⁵m²¹ha³³  Yi ɑ⁵⁵_mu̱²¹_hɑ⁵⁵  Yi ma³³ha³³
```

n·s|RAINBOW · binary · 18 languages · enrichment 14×

```
Alike Tibetan ndʑæ   Batang Tibetan ndʑ.a.⁵³  Guanyinqiao Khros ndʑa⁵³
Japhug ndʑa          Kangding Minyag ɳdʑæ⁵³   Mazur Stau ndʑæ
Mbarkhams Situ ndʒɑ  Ngyaltsu Zbu ndʑê       Siyuewu Khroskyab ndʑɑ
Stau ndʑa            Tshobdun ᴺʤe22           Wobzi Khroskyabs ndʑô
Xinlong Queyu ndʑo   Zhaba ndʑa⁵⁵             Zlarong ndʑe53
rGyalrong nʤa
```

Animals

A single code reaches this section: `k·w|dog`, 24 languages and five branches. The family's animals — otter, rabbit, cat, flea — fall under "Other" because their glosses are not on the thematic list, and that is where `s·r·m|otter` sits, the best genealogical code in the appendix.

k·w|dog · binary · 24 languages · enrichment 36×

```
Achang kʰui⁴⁵  Bai khuɑ³³              Bela kʰui²³        Burmese khwe⁵⁵
Deng kui⁵⁵     Deng_Darang_Tarao kuɑɯ⁵³ Jianchuan kʰũã³³  Jingpho kui³¹
Maang khui³⁵   Mondzi khui⁴⁴           Old Burmese khwe:  Phon kʰui³¹
WB khwe3       Xiangyun kʰũã³³         Zaiwa khui²¹       Zhaozhuang Bai khua⁴⁴
```

Plants and food

Four codes, headed by `d·z|eat` (34 languages, 58×) and `t·n|drink` (22, seven branches). Both are verbs, not plants: this family's crops — sorghum, groundnut, walnut, grape, chili — are under "Other".

d·z|eat · binary · 34 languages · enrichment 58×

```
Achang dzɔ³       Diaocao dza²¹  Gepu dzo²¹  Jinggu Limi dzo²¹
Kiranti_Hayu dza  Lashi dzɔ³    Lisu dzɑ³¹  Maru dzo⁴
Nasu dzo³³        Neshu dzo³³   Sani dzɑ³³  Shuizhuping dza²¹
Toloza dza³³      Yangliu dza²¹ Yi dza²¹    Yijiacun dzo²¹
```

s·s|tree · binary · 26 languages · enrichment 9×

```
Axi si33ʦɛ33              Azha si⁴⁴ʦɛ²¹       Diaocao ɕi³³dzei²⁴  Jinggu Limi ɕi⁵⁵tɕi³³
Kiranti_Kulung s.i:.tʃʰ.ɔ Lavu sɪ33ʨɪ55      Lipo si³³tsi̱³³      Lolopo se³³_dzɯ³³
Naxi sɪ³³_dzɪ³³           Neshu si̱³³dʐe²¹    Old Burmese sac     Rouruo sɛ53ʦe33
Toloza sɪ55ʨɪ55           Xuzhang sɪ⁴⁵³dzɪ²⁴  Yi ɕi̱³³dzɛ²¹        Yi sɪ³³_dzɪ⁵⁵
```

t·n|drink · binary · 22 languages · enrichment 15×

```
Alike Tibetan thoŋ  Batang Tibetan thuŋ⁵³  Bisu taŋ55          Deng tɑuŋ⁵⁵
Deng_Yidu tioŋ⁵⁵    Gyalsumdo t̪ʰuŋ        Khaling tuŋ◦        Kiranti_Hayu tun
Lowa thuŋ⁵⁵         Manange 2tʰuŋ         Miji Nafra thoŋ     Newar ton-
Old Tibetan 'thuṅ   S.kong taŋ55          Tibetan_Lhasa thuŋ⁵⁵ TibetoKinauri_Bun tuŋ
```

k·n|root · binary · 21 languages · enrichment 13×

Changsha kən^{33}	Chaozhou kn̩$^{\sim 33}$	Chengdu kən^{45}	Eastern Puroik a^{31}–koŋ^{55}
Guangzhou k.ɐ.n.$^{53}{}_{55}$	Guilin kɤn^{42}	Ha_erbin kən^{44}	Meixian kin^{33}
Miji Nafra ukhən	Nanjing kən^{31}	Rongcheng kən^{52}	Sinitic_Beijing kən^{55}
Sinitic_Jieyang kɛŋ^{33}	Suzhou kən^{44}	Taiyuan kəŋ^{21}	Xiamen kun^{55}

Verbs of action and state

A single code: s·j|laugh. This family's verbs are short and collide heavily, so they reach the catalogue as unaries and do not enter this appendix.

s·j|laugh · binary · 22 languages · enrichment 21×

Beijing ɕiɑu^{51}	Changsha ɕiau^{55}	Chengdu ɕiau^{213}	Fuzhou tsʰieu^{212}
Guangzhou siu^{33}	Guilin ɕiɑu^{25}	Jinan ɕiɔ^{31}	Jixi ɕie^{35}
Loudi siɤ^{35}	Mandarin Chinese xiào	Meixian siau^{53}	Nanjing ɕiɔ^{44}
Rongcheng siau^{334}	Suzhou siæ^{412}	Wenzhou ɕiɛ^{51}	Xi_an ɕiɑu^{44}

Objects and culture

A single code carries this label — l·m|path or road — and yet material culture is half the appendix. That is an artefact of the thematic list, not of the family: the Lexibank glosses for this part of the world do not match the labels the list declares, and they all end up under "Other".

l·m|PATH OR ROAD · binary · 24 languages · enrichment 86×

Chang lam	Daai Kal–im lám	Jerigaon lə+mu	Jingpho lam
Khaling lāem	Khumi lām	Maram lam	Maring lam
Ng'ga Kant–thayun lám	Nocte lam	Old Tibetan lam	Phom lem
Rupa nə+mu	Sartang Khoitam lym	Sherpa lām	Uppu lam

Qualities

Five codes, and among them are the lowest-enrichment codes of the whole appendix: s·n|new 7×, s·n|long 6×, s·n|heavy 5×, alongside k·n|clean 24× and k·n|dry 12×.

Three of the five are s·n. An enrichment of 5× in a code covering twenty languages says exactly what it looks like: the code is picking up the frequent, not the shared. They are the floor of this list, and they are here because reach lifted them.

s·n|new · binary · 27 languages · enrichment 7×

Beijing ɕin^{55}	Caijia ɕin33	Chaozhou siŋ^{33}	Fuzhou siŋ^{55}
Guangzhou sen^{53}	Ha_erbin ɕin^{44}	Kua–nsi ɕi^{55}nu^{33}	Kuansi ɕi^{55}_nu^{33}
Manange 1tsaŋ	Meixian sin^{33}	Nanchang ɕin^{42}	Rondu Balti soŋɑ
Sinitic_Beijing ɕiən^{55}	Sinitic_Jieyang sɛŋ^{33}	Suzhou sin^{44}	Wenzhou saŋ^{44}

s·n|long · binary · 24 languages · enrichment 6×

Achang səŋ⁵⁵	Beijing tʂʰaŋ³⁵	Changsha tsan¹³
Chengdu tsʰaŋ²¹	Guilin tsɑŋ²²	Ha_erbin tʂʰaŋ²⁴
Jinan tsʰɑŋ⁵³	Lashi ʃ.ə̰ː.ŋ.³³	Mandarin Chinese chang2
Meixian tsʰoŋ¹¹	Nruanghmei seinu	Rawngtu Weilong ʂ̊sʰɛːn
Shenyang tsʰaŋ³⁵	Sinitic_Beijing tʂʰ.ɑ.ŋ.³⁵	Xi_an tʂʰɑŋ²⁴
Yangzhou tsʰɑŋ³⁴		

s·n|heavy · binary · 20 languages · enrichment 5×

Beijing tʂuŋ⁵¹	Changsha tsən²¹	Chengdu tsoŋ²¹³	Guangzhou tsoŋ²²
Hefei tʂəŋ⁵³	Jinan tʂuŋ²¹	Kunming tʂoŋ¹³	Mandarin Chinese chen2
Nanchang tsʰuŋ²¹	Nanjing tsoŋ⁴⁴	Rongcheng tʂoŋ³³⁴	Shenyang tsuŋ⁴¹
Sinitic_Longgang tsʰ.ʊ.ŋ.³³	Taiyuan tsuŋ⁴⁵	Tuoluo ʐvŋ33	Yangjiang tʃʰʊŋ²¹

k·s|CLEAN · binary · 18 languages · enrichment 24×

Diaocao ka⁵⁵tɕi²¹	Eryuan ka³⁵tɕɯ⁴²	Heqing kã⁵⁵tɕɯ⁴²	Jianchuan kã⁵⁵tɕw̃⁴²
Kunming kã⁴⁴tɕĩ¹³	Lanping kã⁵⁵tɕw̃³¹	Luobenzhuo ka⁵⁵tɕi⁴⁴	Nasu kɔ³³tsi²¹
Phuza ka⁵⁵ʨi¹³	Qiliqiao ka³⁵tɕɯ⁴⁴	Sanni ka³³tsi⁵⁵	Xiangyun kã⁵⁵tɕw̃⁴²
Yi k□ ³³tɕi²¹	Yunlong ka³⁵tɕɯ⁴²	Zhoucheng ka³⁵tɕɯ⁴²	rGyalrong kə−ʂo

k·n|dry · binary · 18 languages · enrichment 12×

Beijing kan⁵⁵	Changsha kan³³	Chengdu kan⁴⁴	Fuzhou kaŋ⁴⁴
Guangzhou kɔn⁵⁵	Ha_erbin kan⁴⁴	Kiranti_Limbu kaŋ◦	Mandarin Chinese gan1
Nanchang kon⁴²	Rengma aken	Rongcheng kɐn³³⁴	Sangtam a◦kong
Shenyang kan³³	Sinitic_Beijing kan⁵⁵	Sinitic_Longgang kɔn³³	Xiamen kan⁵⁵

Other

Fifty-five codes, more than half the appendix. It is not a residual category: it is the part of the Sino-Tibetan vocabulary collected in Lexibank that this appendix's thematic list — written for basic vocabulary — does not contemplate.

Here are the loans from the header — b·d·m|groundnut, p·t|grape, p·l|glass, l·s|chili pepper, k·n|steel, s·g|plough, m·l|sorghum, t·n|copper, n·n|silver — the highest-enrichment compound numeral, n·s·t|twenty one at 1545×, and also the best genealogical code on the list, s·r·m|otter, which crosses seven branches and sits at position 94.

Read this section knowing it mixes the two things and that the catalogue does not separate them: s·r·m and b·d·m look alike, share the same block format, and have enrichments of the same order. One is an etymon of millennia and the other arrived with the plant. The difference is not in the data this procedure looks at.

s·n|TWELVE · binary · 50 languages · enrichment 38×

Alo tsʰe²¹ɳ̊i⁵⁵	Diaocao tsʰei⁵⁵ŋ²¹	Gepu tsʰɪ³³n²¹	Hlepho ʧʰə²¹n³³
Kazhuo tshi³³ŋ³¹	Laowu Nisu tɕʰi̇²¹ni⁵⁵	Lisu tshɪ⁴⁴ɲi³¹	Mangdi tɕʰɛ⁵⁵ŋ²¹
Nasu tsʰe²¹ɳ̊i⁵⁵	Nipu tsʰɯ²¹ɳ̊i⁵⁵	Sani tsʰɪ³³n²¹	Shuizhuping tsʰe²⁴ni²¹
Tsum tɕuɲi	Xuzhang tɕʰĩ⁴⁵³n²¹	Yi tshi³³ni²¹	Yijiacun tɕʰɛ⁵⁵ŋ²¹

n·s|MEDICINE · binary · 45 languages · enrichment 48×

Alo nu²¹tsʰi³³	Azha na44ʈʂhɪ33	E Lolo nɥ³³tɕʰi²¹	Haoni na33ʈʂhi31
Kazhuo na24ʈʂhɪ31	Lahu nɑ33ʈʂhi33	Lavu no33ʈʂhɪ21	Lolopo no33ʈʂhi21
Narsu no⁵⁵tɕʰi³³	Nersu no²¹tsʰi³³	Nusu nɛ33ʈʂhi31	Polo nu21ʈʂhɛ13
Sani nɑ³³tsʰi²¹	Toloza na²¹tsʰi³³	Yangliu nɛ̃⁵⁵tsʰɪ⁴¹	Yi ne³³tshɪ²¹

s·l|COTTON · binary · 34 languages · enrichment 87×

Alo su³³lu³³	Black Lahu s□ ³³l□ ⁵³	Hani sa³¹la³¹	Hlepho sa³³la³³
Lavu su⁵⁵lu⁵⁵	Lisu sɑ⁴⁴lɑ³¹	Mondzi sɔ¹³lɔ¹³	Naso so³³lo³³
Neshu so⁵⁵lo⁵⁵	Nungic_Dulong sɑ⁵⁵lɑ⁵³	S. Muji sa³³la³³	Sanni sa³³la³³
Sodi sa³³lɯ³³	Xuzhang so⁴⁵lo²¹	Yi so⁵⁵lo²¹	Yi sa³⁴lɯ³³

s·t|ELEVEN · binary · 34 languages · enrichment 45×

Achang tɕhe⁵⁵ta³¹	Axi tshi³³ti⁵⁵	Diaocao tsʰei⁵⁵ti²¹	Gepu tsʰɪ³³ti⁵⁵
Haoni tshɛ⁵⁵thi³¹	Jinggu Limi tsʰe³³tʰi²¹	Laowu Nisu tɕʰi²¹ti³³	Lisu tshɪ⁴⁴ti⁵⁵
Narsu tɕʰe²¹ti³³	Nasu tsʰe²¹ti³³	Neshu tsʰɤ²¹ti³³	Phuza ʈʂʰə³¹ti³³
Sanni tɕʰi³³tʰi⁵⁵	Toloza tsʰi²¹tʰɤ⁵⁵	Xuzhang tɕʰi⁴⁵³tʰi³³	Yi tshi³³ti⁵⁵

s·n|ONION · binary · 32 languages · enrichment 24×

Beijing tsʰuŋ⁵⁵	Bugun Kaspi tʃoŋ	Bugun Singchung tʃoŋ	Changsha tsʰən³³
Chug Parchu tʃoŋ	Fuzhou tsʰøyŋ⁴⁴	Hefei tsʰəŋ²¹²	Kunming tsʰoŋ⁴⁴
Lowa tsoŋ	Miji Nafra tʃoŋ	Monpa Tawang Mona tʃoŋ	Nanchang tsʰuŋ⁴²
Sema atsüna	Sherdukpen Rupa tʃoŋ	Wenzhou tsʰoŋ⁴⁴	Yangjiang tʃʰʊŋ³³

s·n|GOLD · binary · 30 languages · enrichment 12×

Angami Khonoma sona	Angami Kohima suna	Bugun Kaspi san	Bugun Singchung san
Chokri suna	Khoirao sana	Lawu tɕiŋ⁵⁵	Mandarin Chinese jin1
Mao shana	Maring sana	Mzieme suna	Phon ʃaiŋ⁵⁵
Rupa sin	Sema suna	Shergaon sin	Wancho sien

s·n|SHOE · binary · 30 languages · enrichment 16×

Alo tɕʰi³³nɐ̯²¹	Axi tɕhɐ̯³³nɑ̯³³	Diaocao tɕʰi⁵⁵nə³³	Gepu tɕʰi³³nɤ̯³³
Hlepho ʈʂʰɪ²¹nɔ²¹	Kazhuo tshɪ³³nɛ³⁵	Kuansi tsʰɪ⁵⁵nɨ̯³³	Lavu tɕʰi⁵⁵nɤ̯³³
Lisu tʃhi⁴⁴ne³³	Narsu tɕʰi³³nɤ̯³³	Nasu tɕʰɐ̯⁵⁵nɛ̃²	Neshu tshɪ⁵⁵nɤ̯³³
Phuza ʈʂʰi³¹nʌ⁵⁵	Sani tʂʰi⁴⁴nɤ⁴⁴	Yi tɕhi⁵⁵nɯ̯³³	Yi ɕi³⁴n̥ɪ̯³³

d·z|COOKED RICE · binary · 27 languages · enrichment 86×

Alo dzu²¹	Diaocao dza²⁴	Jinggu Limi dzo²⁴	Lavu dzu⁵⁵
Mangdi dzo⁵⁵	Mondzi dzɔ¹³	Naso dzo²¹	Nasu dzo²¹
Neshu dzo²¹	Sanni dza³³	Shuizhuping dza²⁴	Toloza dzɑ²¹
Yangliu dza³³	Yi dzo²¹	Yi dza⁵⁵	Yijiacun dzo²⁴

s·t|kill · binary · 27 languages · enrichment 20×

Achang sat⁵⁵	Bantawa set	Deng sɑt⁵⁵
Garo_Garo si÷et÷a	Guangzhou sat³	Jingpho sat³¹
Kiranti_Bahing sat	Kiranti_Hayu sɪt	Kiranti_Thulung set
Mandarin Chinese sha1diao4	Mbarkhams Situ siɛ̃t	Monpa Tawang Mona soto
Njorogs Khroskyab sat⁵⁵	Nungic_Dulong sɑt⁵⁵	Yangjiang ʃat²¹
Zeme chat		

s·m|DREAM · binary · 26 languages · enrichment 23×

Alo ʑi̱⁵⁵mą̱²¹	Axi ʑe21me33	Azhe ʑi21mɛ22	Dimasa saomi
Khaling semɛ	Lavu ʐɹ21ma33	Lipo ʑɹ21mi33	Mondzi ʑi13mɔ44
Naso ʑi̱⁵⁵mą̱³³	Nasu ʑi̱⁵⁵mą̱³³	Neshu ʑi̱²¹mę̱²¹	Nipu ʑe¹³ma⁵⁵
Samu ʥɹ33mɔ33	Shuizhuping ʐl⁴⁵³mɨ³³	Up: Phola ʑi³³mɔ³³	Yi ʐl²¹ʔmę̱³³

l·s|CHILI PEPPER · binary · 25 languages · enrichment 54×

Alo la̱⁵⁵tse²¹	Axi l□ □ ⁵⁵tsi̱²¹	Hlepho la⁵⁵ʫɹ²¹	Jianchuan la³⁵tsɹ³³
Kua-nsi la⁵⁵dzɹ²¹	Kunming la³¹tsɹ⁵³	Lavu la⁵⁵tsi³³	Mandarin Chinese la4jiao1
Nasu la̱⁵⁵tse²¹	Phuza la³¹ʥi³³	Qiliqiao la³⁵tsɹ³¹	Toloza la³³tsi⁵⁵
Xiangyun la³⁵tsɹ³¹	Yi l□ ⁵⁵tsi²¹	Yi l□ □ ⁵⁵tsi³³	Yunlong la³⁵tsɹ³¹

s·m|OLD PERSON · binary · 25 languages · enrichment 56×

Axi tshu³³mu²¹	Azha ʫʰɯ²¹mu²²	Diaocao tsʰa⁵⁵ma²¹	Gepu tsʰɯ³³mɯ²¹
Hlepho ʫʰə²¹mə³³	Jinggu Limi tsʰo³³mo²¹	Lavu tsʰɑ⁵⁵me²¹	Lawu tsʰu⁵⁵ma³³
Mondzi tsʰɔ¹³mɔ²¹	Neshu tʂh□ ²¹m□ ²¹	Phuza ʫʰʌ³¹mɔ³¹	Sanni tʂʰu³³mo²¹
Shuizhuping tsʰa²⁴ma²¹	Xuzhang tsʰa⁴⁵³ma²¹	Yangliu tsʰu³³mu²¹	Yi tshu⁵⁵mu²¹

s·s|claw · binary · 25 languages · enrichment 16×

Angami Khonoma jetse	Axi tɕhi³³si²¹	Bai tso³¹_tsi⁴⁴
Changsha tsau⁴¹ tsɹ⁰	Ersu dzɹ⁵⁵_dzɹ⁵⁵	Guilin tsɑu²³ tse²³⁻³³
Jixi tsɤ²¹³ tsɹ⁰	Kua-nsi tsʰɹ⁵⁵tɕʰi³³	Lanping sdʒã¹³
Loudi tsɤ⁴² tsɹ⁰	Mandarin Chinese zhua3zi	Nasu tɕʰi⁵⁵tɕa²¹
Nersu tɕʰi¹³sɤ³³	Sodi ɕi³³sɪ³³	Toloza tɕʰi⁵⁵se̱³³
Yi ɕi³⁴sɪ³³		

m·s|BROOM · binary · 24 languages · enrichment 34×

Alo mi³³sɯ²¹	Axi mi³³si̱³³	Black Lahu mi³¹si⁵⁴	Diaocao mə⁵⁵sɹ³³
Hlepho mi⁵⁵sɯ²¹	Kazhuo mi³³sa³⁵	Kuansi ɱɨ⁵⁵_sɹ³³	Lavu my⁵⁵su̱³³
Narsu mi⁵⁵si̱³³	Naso mi³³si̱²¹	Neshu mi⁵⁵si̱³³	Phuza mi³³si⁵⁵
Sani mi⁴⁴si⁴⁴	Sanni mi³³si̱⁵¹	Toloza mi²¹si̱³³	Yi mi⁵⁵sɻ̱³³

s·n|FIREWOOD · binary · 24 languages · enrichment 10×

Bantawa sɨŋ	Daai Kal-im ʃiŋ	Gyalsumdo ɕiŋ	Jirel ɕiŋ
Kagate siŋ	Khaling sʌŋ	Kiranti_Limbu siŋ	Lawu ɕiŋ⁵⁵
Manange 2ʃiŋ	Monpa Tembang ɕiŋ	Nungic_Dulong ɕiŋ⁵⁵	Old Chinese sin
Rondu Balti šiŋ	Sherpa ɕiŋ	Tsum ɕiŋ	Wadu sĕŋ

d·z|RIDE · binary · 24 languages · enrichment 119×

Alo dzɔ³³	Diaocao dzɛ²¹	Gepu dzɛ²¹	Hani dzɹ³¹
Lavu dzɑ²¹	Lisu dzɯ³¹	Naso dzɔ³³	Nasu dzǫ̱³³
Neshu dzE³³	Nipu dze³³	Shuizhuping dzɛ²¹	Sodi dzi³³
Xuzhang dzɛ̃²¹	Yangliu dzɛ̴²¹	Yi dzɛ²¹	Yi dzi³³

n·s·t|TWENTY ONE · ternary · 24 languages · enrichment 1545×

Alo ȵi⁵⁵tse³³tʰa̱²¹	Axi ni²¹tsi³³thi²¹	Black Lahu ni⁵³tshi³³te⁵³
Hani ɲi³¹tshe⁵⁵ti̱	Kua-nsi ȵi²¹tɕi⁵⁵ti²¹	Lawu n²¹tɕʰi⁵⁵tʰi³³
Mondzi ni⁴⁴si⁴⁴tɔ⁴⁴	Narsu ni²¹tse⁵⁵tʰi²¹	Nasu ȵi⁵⁵tse³³tʰa̱²¹
Nersu ni²¹tsɤ¹³ti³³	Nipu n²¹tsɯ³³tʰi⁵⁵	Phuza ȵi³³ʦɛ⁵⁵tʰa³³
Sanni n²¹tɕi³³tʰi²¹	Shuizhuping ni²¹tse²⁴tʰi⁵⁵	Yi ni²¹tsi³³thi²¹
Yi n²¹tsi⁵⁵ti̱²¹		

s·n|sing · binary · 24 languages · enrichment 10×

Achang tʃʰuŋ⁵	Beijing tʂʰɑŋ⁵¹	Changsha tsʰan⁵⁵
Chaozhou tsʰ.i.a.ŋ.¹³	Deng_Yidu ɑ⁵⁵◦ɕeŋ⁵⁵	Guangzhou tsʰ.œ.ŋ.³³
Ha_erbin tʂʰaŋ⁵³	Jinan tʂʰɑŋ³¹	Maru tʃʰæuŋ⁴³
Meixian tsʰoŋ⁵³	Nanchang tsʰɔŋ²¹³	Nruanghmei son
Shuizhuping tʂʰaŋ⁵⁵	Sinitic_Beijing tʂʰ.ɑ.ŋ.⁵¹	Sinitic_Longgang tsʰ.ɔ.ŋ.⁵⁴
Wancho shun		

s·t|ASHAMED · binary · 24 languages · enrichment 79×

Alo ʂa̱⁵⁵tɒ³³	Diaocao ɕi⁴²ta⁵⁵	Hani sa⁴⁴to⁵³	Jinggu Limi ɕi⁵⁵to³³
Laowu Nisu sa¹³tɔ³³	Lavu ʂa²¹ta⁵⁵	Mangdi ɕɛ̱²¹ta⁵⁵	Mondzi ʂa⁴⁴tɔ⁴⁴
Naso ʂa̱⁵⁵tɔ³³	Nasu ʂɑ̱⁵⁵tɔ³³	Neshu ɕe̱²¹ta⁵⁵	Nipu ʂa³³tɔ³³
Shuizhuping ɕi⁵⁵ta³³	Xuzhang ɕi³³ta³³	Yi ɕe̱⁵⁵ta³³	Yi ɕɛ̱²¹tu⁵⁵

n·m|YOUNGER SISTER (OF MAN) · binary · 24 languages · enrichment 130×

Black Lahu ɔ³¹ni³³m□ ³³	E Lolo nɯ⁵⁵mo³³	Hlepho ɛ⁵⁵ȵa⁵⁵ma³³	Jinggu Limi na³³mɔ³³
Lipo ni³³mo³³	Lisu ɲi³³mɑ³³	Narsu nɛ²¹mo²¹	Nasu ni²¹mɔ⁵⁵
Phuza ȵi⁵⁵mɔ³³	Sani nɤ⁵⁵mɑ³³	Shuizhuping ne²⁴ma³³	Sodi ȵi̱³³mɑ̱⁵⁵
Xuzhang ɲɛ̆²⁴mo³³	Yi nɯ⁵⁵mo³³	Yi n̥i²¹mo²¹	Yijiacun nɛ²⁴mo³³

s·p|BOARD · binary · 23 languages · enrichment 58×

Alo si̱²¹pʰe⁵⁵	Axi si̱³³_phi²¹	Azha si⁴⁴pʰi²²	Gepu si̱³³pʰɪ²¹
Jinggu Limi ɕi³³pʰi²¹	Mondzi si⁴⁴pʰi²¹	Narsu ɕi̱³³pʰe²¹	Nasu si̱²¹pʰe⁵⁵
Nersu si̱³³pʰe⁵⁵	Neshu si̱³³pʰɪ²¹	Sani si⁴⁴pʰɪ²¹	Sanni si̱⁵¹pʰi²¹
Toloza si̱³³pʰi⁵⁵	Xuzhang sɪ³⁵³pɛ̆³³	Yi si̱³³pʰɪ²¹	Yino si̱³³pʰɪ²¹

l·s|BRACELET · binary · 23 languages · enrichment 88×

Alo la̱⁵⁵dʑu²¹	Axi le̱²¹tʂo³³	Azha lɛ²²ʨœ²¹	Gepu lɪ̱²¹tʂo³³
Hlepho lɛ³³ʥo²¹	Kazhuo la⁵³tsɪ³⁵	Laowu Nisu la¹³dʑy²¹	Lipo li̱³³tʂɯ³³
Lisu lɛ³¹dʒu³³	Mondzi lɔ⁵³dʑo¹³	Nasu la̱⁵⁵dʑu²¹	Nersu la¹³dʑu²¹
Nipu la¹³dʑu²¹	Phuza li⁵⁵tʂɯ³¹	Toloza le̱²¹dʐu²¹	Yi le̱²¹dzu̱²¹

t·n|COPPER · binary · 23 languages · enrichment 46×

Achang tɔŋ55	Beijing tʰuŋ³⁵	Bisu toŋ55	Chaozhou taŋ⁵⁵
Chengdu tʰoŋ³¹	Guangzhou tʰʊŋ²¹	Hefei tʰəŋ⁵⁵	Kunming tʰoŋ³¹
Lavu thoŋ13	Maang thoŋ33	Mbarkhams Situ □ tʰoŋ	Meixian tʰuŋ¹²
Nanchang tʰuŋ²⁴	S.kong toŋ31	Xiamen taŋ²⁴	Yangjiang tʰʊŋ⁴⁴³

s·m|WALNUT · binary · 23 languages · enrichment 42×

Alo sɒ²¹me²¹	Diaocao sa²¹mei²⁴	Gepu sa²¹mɪ³³	Laowu Nisu sa²¹mei²¹
Mondzi sa²¹mi¹³	Naso sa²¹me²¹	Nasu sa²¹me²¹	Neshu sa²¹me²¹
Nipu se²¹mi²¹	Phuza sa³³mi³¹	Sanni sa²¹mi³³	Shuizhuping sa²¹me²⁴
Xuzhang sa²¹my²⁴	Yangliu sa²¹mi³³	Yi sɛ²¹mi⁵⁵	Yi si²¹m̥i³³

m·n|CAT · binary · 22 languages · enrichment 26×

Axi mɛ55nɛ33	Azha ma44ne33	Black Lahu mɛ³⁵ni³³	Gepu mɛ⁵⁵nɛ³³
Jinggu Limi mi⁵⁵ne⁵⁵	Kaang Kruk min	Lanping mu¹¹_n̥i⁵⁵	Laowu Nisu mi⁵⁵nɛ³³
Lope mæ˜33næ˜33	Mikir meng	Nasu mɛ˜⁵⁵nɛ˜³³	Nocte miang
Nruanghmei meo◦na	Nusu mɯ33n̥e31	Sani mae⁴⁴nae³³	Up: Phola mæ³¹nɨ³³

m·n|DREAM · binary · 22 languages · enrichment 29×

Ao Mongsen a◦mang	Caijia məŋ22	Chang mang	Chin_Hakha m.â/a.ŋ
Guangzhou moŋ²²	Konyak mang	Lushai maŋ	Mandarin Chinese mə̂ŋ
Mikir māng	Mzieme mang	Naga_Ukhrul mɐŋ	Phom mang
Sinitic_Beijing mɤŋ⁵¹	Sinitic_Jieyang maŋ¹¹	Tangkhul mang	TibetoKinauri_Bya maŋ

m·s|NOT HAVE · binary · 22 languages · enrichment 101×

Alo ma²¹dʐɒ²¹	Black Lahu m□ ⁵³tsɔ³¹	Burmese mɑ⁵³ɕi⁵³	Hani ma³¹dʑa̱³³
Haoni ma³¹tʃa³³	Kazhuo ma³¹tso³²³	Mangdi ma²¹dʒa⁵⁵	Mondzi ma²¹dʐɔ¹³
Nasu ma²¹dʐɒ²¹	Nersu ma²¹dʑo²¹	Neshu ma²¹dʐa²¹	Sanni ma²¹dʐu³³
Shuizhuping ma²¹dʐu²⁴	Toloza mɤ³³dʐo²¹	Yi m□ ²¹dz□ ²¹	Yi ma²¹dʐu⁵⁵

d·z|RAW · binary · 22 languages · enrichment 112×

Alo dzi³³	Diaocao dzɪ²¹	Gepu dzɛ²¹	Lavu dzi²¹
Lisu e⁵⁵dzɪ³¹	Narsu dzɛ³³	Nasu dzi³³	Nersu dze³³
Nipu dzɤ³³	Sani dzi²¹	Shuizhuping dzɪ²¹	Toloza dzɤ⁵³
Xuzhang dzɪ²¹	Yangliu dzɪ²¹	Yi dzɪ²¹	Yi dzɪ³³

l·p|RING · binary · 22 languages · enrichment 96×

Alo la̱⁵⁵pe²¹	Axi le̱²¹pi⁵⁵	Azha lɛ²²pi̱⁴⁵	Gepu lɪ̱²¹pɪ⁵⁵
Hlepho lɛ³³pi⁵⁵	Jinggu Limi le²¹pe⁵⁵	Narsu la̱²¹pe³³	Naso la̱⁵⁵pe³³
Nersu la¹³pe⁵⁵	Neshu le̱²¹pe²¹	Nipu la¹³pi⁵⁵	Sanni le̱⁵⁵pi⁵¹
Sodi lɪ̱⁵⁵pɪ³³	Toloza le̱²¹pu⁵⁵	Yi le̱²¹pe³³	Yi lo̱⁵⁵pɪ³³

s·j|small · binary · 22 languages · enrichment 17×

Beijing ɕiau²¹⁴	Caijia ɕje22	Changsha ɕiau⁴¹	Guangzhou siu³⁵
Ha_erbin ɕiau²¹³	Hefei ɕiɔ²⁴	Jixi ɕie²¹³	Kunming ɕiau⁵³
Mandarin Chinese xiǎo	Mbarkhams Situ tsêj	Nanchang ɕiɛu²¹³	Rongcheng siau²¹⁴
Shenyang ɕiau²¹³	Suzhou siæ⁴¹	Tangut tsəj¹	Xi_an ɕiɑu⁵³

t·j|JUMP · binary · 21 languages · enrichment 76×

Beijing tʰiau⁵¹	Changsha tʰiau⁵⁵	Chaozhou tʰiəu²¹³	Chengdu tʰiau¹³
Fuzhou tʰiu²¹³	Guangzhou tʰiu³³	Jinan tʰiɔ²¹	Kuansi tʰiɑu³³
Kunming tʰiau¹³	Mandarin Chinese tiào	Nanchang tʰiɛu²¹³	Shenyang tʰiau⁴¹
Suzhou tʰiæ⁵¹³	Xi_an tʰiɑu⁴⁵	Xiamen tʰiau¹¹	Yangjiang tʰiu²⁴

m·n|NO · binary · 21 languages · enrichment 47×

Alo ma²¹ɳe²¹	Diaocao ma²¹ŋa²⁴	Gepu ma²¹ŋɯ³³	Hani ma³¹ŋɤ⁵⁵
Mangdi ma²¹ŋo⁵⁵	Mondzi ma²¹ŋe⁵³	Narsu ma²¹ŋɤ²¹	Nasu ma²¹ɳe²¹
Nersu ma²¹ɳe¹³	Neshu ma²¹ŋɤ²¹	Sani mɑ²¹ŋae³³	Sanni ma²¹ni³³
Shuizhuping ma²¹ŋa²⁴	Yangliu m²¹ŋa³³	Yi m□ ²¹ŋɯ²¹	Yi ma²¹ŋa⁵⁵

s·m|FLOUR · binary · 21 languages · enrichment 27×

Alo ʂu³³mu²¹	Burmese dʑõ²²m̥õ⁵³	Diaocao sa⁵⁵mə³³	Kazhuo sa²⁴m³³
Mondzi ʂɔ¹³mo¹³	Narsu so⁵⁵mu²¹	Naso ʂo³³mu²¹	Nersu ʂu¹³mu²¹
Neshu so⁵⁵mu²¹	Nipu ʂu³³mu²¹	Sodi ʂa³³mo³³	Toloza ʂɤ³³mu²¹
Up: Phola ɕɑ⁵⁵mu³³	Yi so⁵⁵mu²¹	Yi sɑ⁵⁵mu³³	Yi ʂa³³mo³⁴

p·l|CHEAP · binary · 20 languages · enrichment 83×

Alo pʰu³³lɒ²¹	Axi phu²¹lu³³	Azha pʰy²²lu²¹	Gepu pʰu²¹lu³³
Lavu pʰu²¹la⁵⁵	Mondzi pʰo⁴⁴lɔ¹³	Narsu pʰu³³la²¹	Naso pʰu³³lɔ²¹
Nersu pʰi³³lo²¹	Neshu phu³³l□ ²¹	Nipu pʰu³³lɔ²¹	Sani pʰu²¹lo³³
Sodi pʰu³³lo³³	Toloza pʰu³³lo²¹	Up: Phola pʰu³³lɛ³³	Yi phu³³lɒ²¹

s·n|DAY BEFORE YESTERDAY · binary · 20 languages · enrichment 18×

Black Lahu □ ³⁵sɪ³³ni³³	Diaocao sɪ³³nɛ³³	E Lolo sɪ³³ni³³	Eryuan tsɔ³³ɲi⁴⁴
Jinggu Limi ʂi³³ɳi³³	Kazhuo sɪ³³ŋ³²³	Kua-nsi sɪ²¹ɳi³³	Laowu Nisu sɪ⁵⁵ni²¹
Lisu ʃɪ⁴⁴ɲi³³	Mangdi ɕi³³ni³³	Neshu ʂɻ³³ni²¹	Shuizhuping ʂɛ̃n³³
Xuzhang ʂɻ³³ɲi³³	Yangliu sɪ³³ni³³	Yi si³³ni²¹	Yi ʂi³³ɳi³³

p·t|GRAPE · binary · 20 languages · enrichment 380×

Achang phu³¹thau³¹	Beijing pʰu³⁵tʰau⁰	Changsha pu¹³tau¹³
Chaozhou pʰu⁵⁵₂₁₃tʰo⁵⁵	Guangzhou pʰou²¹tʰou²¹	Hefei pʰu⁵⁵tʰɔ⁵⁵
Jinan pʰu⁴²tʰɔ⁰	Kunming pʰu³¹tʰau³¹₄₄	Meixian pʰu¹²tʰau¹²
Nanchang pʰu²⁴tʰau²⁴	Nungic_Dulong pu⁵³tɔ⁵⁵	Shenyang pʰu³⁵tʰau⁰
Xiamen pʰu²⁴₃₃to²⁴	Yangzhou pʰu³⁴tʰɔ³⁴	Yi phu²¹thɒ³³
Yi phu²¹thɑo³³		

s·p|HOST · binary · 20 languages · enrichment 69×

Alo si²¹pʰu³³	Axi s□ ²¹pho²¹	Azha i³³sɛ³³pʰɔ²²	Black Lahu ɔ³¹sɛ³¹ph□ ⁵³
Lisu sɪ⁴⁴phɑ³¹	Mangdi ɕi⁵⁵pʰo²¹	Narsu sɛ²¹pʰo³³	Naso si²¹pʰo³³
Neshu sE²¹pho³³	Nipu ɕe²¹pʰu³³	Phuza ɛ⁵⁵ɕɛ³³pʰa³³	S. Muji i²¹so³³pʰa⁵⁵
Sanni ɕi²¹pʰa²¹	Up: Phola sæ³³pʰɑ̰²³	Yi sɛ²¹pho³³	Yi sɛ²¹pho²¹

s·p|LUNG · binary · 20 languages · enrichment 20×

Axi tsi³³pu³³	Azha i²²ʧʰy³³pɔ³³	Black Lahu ɔ³¹tshɪ⁵³pho⁵⁴	Chin_Hakha tsuap
Gepu tsʰi²¹pɤ²¹	Hlepho ʧʰɯ³³pʰɔ²¹	Jinggu Limi si²¹pɨ⁵⁵	Lahu ɔ31ʧhi53pho54
Lushai chuap	Rouruo ʧhue55phu33	S. Muji ɛ²¹ʧʰa²¹pʰi³³a²¹	Samu sɑ25pɯ22
Up: Phola ʧʰi⁵⁵pɨ³¹	WB ɑ1shup4	Yi tshi²¹pɯ⁵⁵	Zhaba tse⁵⁵◦pe⁵⁵

n·n|SILVER · binary · 20 languages · enrichment 20×

Chaozhou ŋɯŋ⁵⁵	Daai Kal-im ŋun	Fuzhou ŋyŋ⁵²	Guangzhou ŋɐn²¹
Konyak ngin	Mro-Khimi Areong ŋun	Nocte ngun	Phom ngin
Rawngtu Weilong ŋun	Rupa ɲin	Sartang Darbu 1 niŋ	Sartang Khoina niŋ
Sherdukpen Rupa niŋ	Shergaon ɲin	Wancho ngung	Yangjiang ŋɐn⁴⁴³

k·n|STEEL · binary · 20 languages · enrichment 111×

Achang kaŋ⁵⁵	Beijing kaŋ⁵⁵	Changsha kan³³	Chaozhou kɯŋ²¹³
Fuzhou kauŋ²¹³	Guangzhou kɔŋ³³	Jinan kaŋ²¹³	Meixian kɔŋ⁴²
Nungic_Dulong kɑŋ⁵⁵	Shenyang kaŋ³³	Suzhou kɒŋ⁴⁴	Xi_an kɑŋ²¹
Yangjiang kɔŋ²⁴	Yangzhou kɑŋ³¹	Yi kɑŋ³³	Zaiwa kaŋ²¹

s·n|near · binary · 20 languages · enrichment 5×

Asho seŋ	Beijing tɕin⁵¹	Caijia tɕin33	Changsha tɕin⁴¹
Guilin tɕiŋ²⁵	Ha_erbin tɕin⁵³	Kua-nsi dʑu⁵⁵na²¹	Kuansi dʑu⁵⁵_nu²¹
Mandarin Chinese jin4	Nanchang tɕʰin²¹	Nanjing tɕin⁴⁴	Shuizhuping tɕin⁴⁴
Stau tɕæ-ne	Suzhou dʑin²³¹	Taiyuan tɕiŋ⁴⁵	Xi_an tɕin⁴⁴

s·s|DEW · binary · 19 languages · enrichment 12×

Alo tʂi⁵⁵ʑi²¹	Axi tɕi⁵⁵ʑi³³	Azha ʦi̱⁴⁵ʑi²¹
Batang Tibetan s.i.¹³.◦.tɕʰ.□ .⁵³	Laowu Nisu tsɪ¹³ʑi²¹	Lavu tʂɯ²¹ʑi⁵⁵
Narsu tsɤ²¹ʑi²¹	Nersu tʂy⁵⁵ʑi²¹	Neshu tsɤ²¹ʑi²¹
S. Muji ʦi²¹ɕi³³	Sani tʂi⁵⁵ʐi³³	Toloza tʂo⁵⁵ʑi²¹
Up: Phola ʦi̱²³ʑi̱²³	Yi tsɯ²¹ʑi²¹	Yi tʂi⁵⁵ʑi³³
Yi tʂɯ⁵⁵ʑɪ³³		

n·k|LAST YEAR · binary · 19 languages · enrichment 95×

Alo a²¹ȵi̱²¹kʰu̱⁵⁵	Axi o⁵⁵ni³³khu̱²¹	Azha ɛ⁴⁴nɛ³³kʰɯ²²	E Lolo a²¹ni̱³³kʰo̱²¹
Gepu a²¹ȵɪ³³kʰu̱²¹	Hlepho a³³nɛ³³kʰoa⁵⁵	Mondzi ni⁴⁴kʰu⁵³	Narsu a²¹ȵɪ³³kʰu̱²¹
Naso a²¹ȵi²¹kʰu̱⁵⁵	Nasu a²¹ȵi̱³³kʰo̱⁵⁵	Neshu a²¹ne³³kʰu̱²¹	Phuza a³³ȵi³³kʰɯ⁵⁵
S. Muji a²¹ȵi³³kʰɯ²¹	Sanni a²¹ni⁵¹kʰo̱⁵⁵	Toloza a⁵⁵nɤ³³kʰo̱²¹	Yi □ ²¹ne̱³³kʰu̱²¹

s·g|PLOUGH (INSTRUMENT) · binary · 19 languages · enrichment 258×

Alo si̱²¹gu̱⁵⁵	Diaocao ɕi³³gu⁴²	E Lolo ɕi̱³³go̱²¹	Jinggu Limi ɕi³³go²¹
Kua-nsi sɪ³³gu²¹	Kuansi sɪ³³_gu²¹	Lavu si̱³³go̱²¹	Lipo si̱³³go̱²¹
Lisu sɪ³⁵go⁴¹	Naso si̱²¹gu̱⁵⁵	Nasu si̱²¹go⁵⁵	Sodi si̱³³go̱⁵⁵
Toloza si̱³³go̱²¹	Xuzhang sɻ³³gu³³	Yi ɕi̱³³gu̱²¹	Yi si̱³³gu̱⁵⁵

n·b|DEAF · binary · 19 languages · enrichment 116×

Gepu no⁵⁵bu²¹	Hani na³¹bo³¹	Hlepho ɛ³³na⁵⁵bə⁵⁵	Lavu nu²¹ba²¹
Lisu nɑ⁵⁵bo³¹	Naso no³³bɔ³³	Neshu no⁵⁵ba³³	Nocte naba
Phuza na³³bɔ³³	S. Muji na³³bə⁵⁵	Sanni na²¹bu²¹	Sodi nɤ³³bo³³
Toloza na²¹bo³³	Wancho naba	Xuzhang na²¹ba²¹	Yangliu na²¹bu²¹

k·n|EMPTY · binary · 19 languages · enrichment 34×

Beijing kʰuŋ⁵⁵	Changsha kʰən³³	Chaozhou kʰaŋ³³	Chengdu kʰoŋ⁴⁴
Fuzhou kʰøyŋ⁴⁴	Hefei kʰəŋ²¹²	Konyak kiang	Kunming kʰoŋ⁴⁴
Mandarin Chinese kong1	Meixian kʰuŋ⁴⁴	Nanchang kʰuŋ⁴²	Shenyang kʰuŋ³³
Suzhou kʰoŋ⁴⁴	Wenzhou kʰoŋ⁴⁴	Xiamen kʰaŋ⁵⁵	Yangzhou kʰoŋ³¹

p·l|GLASS · binary · 19 languages · enrichment 330×

Beijing pɔ⁵⁵li⁰	Changsha pɔ³³li⁰	Chaozhou pɔ³³₂₃li⁵⁵	Fuzhou pɔ⁴⁴lɛ⁵²
Guangzhou pɔ⁵⁵lei²¹₅₅	Jinan pə²¹³li⁰	Mandarin Chinese bo1li	Meixian pɔ⁴⁴li¹²
Nanchang pɔ⁴²li⁵⁵	Shenyang pə³³li⁰	Suzhou pu⁴⁴li²⁴₂₁	Xi_an po²¹li⁰
Xiamen po⁵⁵₃₃le²⁴	Yangjiang pɔ³³lei⁴⁴³	Yangzhou po³¹li⁴²	rGyalrong po_li

s·s|FLEA · binary · 18 languages · enrichment 11×

Alo tsʰi³³ɕe²¹	Azha ʨʰi³³ɕɛ²¹	Gepu tɕʰi²¹ɕɪ³³	Hlepho ʨʰɪ⁵⁵sɪ³³
Kangding Minyag 2ʐʌtshi	Kazhuo tshɪ³¹sɛ³³	Lavu tʂʰi²¹ɕe⁵⁵	Narsu tɕʰi³³ɕe²¹
Nersu tsʰi³³ɕi²¹	Neshu tɕhi³³ɕɪ²¹	Phuza ʨʰi⁵⁵ɕi³³	S. Muji ʨʰɪ⁵⁵ɕɛ³³
Sani tʂʰi²¹ɕɪ³³	Sodi tsʰi³³ɕɪ³³	Up: Phola ʨʰi³³ɕi⁵⁵	Xia Tujia sɪ⁵⁴ sɪ³³

s·r|GOLD · binary · 18 languages · enrichment 51×

Chug Parchu ser	Gyalsumdo seɹ	Jirel seɹ	Kagate ser
Lish Lish ser	Lowa seɾ	Mikir ser	Monpa Kalaktang ser
Monpa Tembang ser	Rahung sɛɹ	Rondu Balti sɛr	Sartang Darbu 1 ser
Sartang Khoina ser	Sartang Khoitam ser	Sherpa ser	Tsum seɾ

b·d·m|GROUNDNUT (PEANUT) · ternary · 18 languages · enrichment 1086×

Bugun Bichom badam	Bugun Kaspi badam	Bugun Namphri badam
Bugun Singchung badam	Bugun Wangho badam	Chug Parchu badam
Jirel mbəd̪am	Kagate bəd̪am	Monpa Kalaktang badam
Monpa Tawang Mona badam	Monpa Tembang badam	Sartang Darbu 1 badam
Sartang Khoina badam	Sherdukpen Rupa badam	Sherpa bəd̪əm
Tsum bəd̪am		

s·r·m|OTTER · ternary · 18 languages · enrichment 596×

Ao Chungli shirem	Ao Mongsen chiram	Chin_Hakha sàay+hrěm
Chug Parchu ɕa+ram	Guanyinqiao Khros sram⁵³	Jingpho saram
Lushai sa+r̥am	Mbarkhams Situ srem	Naga_Ukhrul si+rɐm
Ngyaltsu Zbu srâm	Nungic_Dulong sɯ³¹|ɹăm⁵³	Old Tibetan sram
Rawngtu Weilong ʃɑ.r̥ɑm	Siyuewu Khroskyab srə̂m	Tani_Bokar ɕə+ram
Tshobdun srɐm		

t·l|RABBIT · binary · 18 languages · enrichment 93×

Diaocao tʰɔ³³lɔ³³	Hani tho31la33	Jinggu Limi tʰɔ²¹lɔ³³	Kua-nsi tʰa²¹lo³³
Lavu thɑ21lu33	Lipo thɑu21lo33	Lisu tho31lɑ33	Narsu tʰa⁵⁵ɬo̱³³
Neshu th□ ⁵⁵ɬo̱²¹	Nusu tha33la31	Polo ta33lo55	Rouruo tha33lɔ35
Sanni tʰo⁵⁵lo̱⁵¹	Toloza tho21lɑ33	Yi th□ ²¹|lo³³	Yi tha³³|la³³

m·l|SORGHUM · binary · 18 languages · enrichment 156×

Alo mu²¹ɬu³³	Axi m□ ⁵⁵ɬo²¹	E Lolo mo²¹lu⁵⁵	Gepu ma⁵⁵ɬo²¹	
Jinggu Limi mu²¹lu⁵⁵	Lavu mu²¹lu³³	Lipo mə̱²²lə̱⁵⁵	Lisu mu⁴¹lu⁵⁵	
Naso mu²¹ɬu³³	Nasu mo²¹ɬu³³	Nersu mu²¹ɬu³³	Neshu mo⁵⁵ɬo³³	
S. Muji mo³³ɬɯ⁵⁵	Sani mo⁵⁵ɬo⁵⁵	Sanni ma⁵⁵lu²¹	Toloza mu²¹lu³³	

d·z|THORN · binary · 18 languages · enrichment 54×

Achang dzu³⁴	Alo dzɯ⁵⁵	Gepu dzɯ²¹	Lashi dzu⁴
Maru dzau⁴	Mondzi dzo²¹	Narsu dzɯ²¹	Naso dzɯ⁵⁵
Nersu dzɤ⁵⁵	Neshu dzɯ²¹	Nipu dzɯ⁵⁵	Sani dzi²¹
Sanni dzu²¹	Xuzhang a⁴⁵dzy²¹	Yangliu dzə²¹	Yi a⁵⁵dzy²¹

s·n|dirty · binary · 18 languages · enrichment 7×

Ao Chungli jana	Beijing tsaŋ⁵⁵	Caijia tsʊŋ55	Changsha tsan³³
Chengdu tsaŋ⁴⁴	Ekai Dalet Ywa Ha ɑ.so.hnɑ	Ha_erbin tsaŋ⁴⁴	Hani ʂẹ³³nɔ³¹
Lashi a³¹+san⁵⁵	Manange 1a-tsaŋ	Mandarin Chinese zang1	Nanchang tsɔŋ⁴²
Rongcheng tsaŋ⁵²	Shuizhuping tsaŋ³³	Sinitic_Beijing tsɑŋ⁵⁵	Xi_an tsɑŋ²¹