Consonant Codes in 154 Pama-Nyungan Languages

The family where the instrument costs least

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

Abstract

We present a catalogue of 2,092 consonant codes grouping 11,171 lexical forms from 154 Pama-Nyungan languages. A code is a consonant skeleton shared by words for the same concept across distinct languages. No reconstruction, no expert cognacy judgements, and no hypothesis about Proto-Pama-Nyungan are used.

This is the fourth family in the series and the one that costs the instrument least. 34 % of its forms share a skeleton with another form in the same language — the lowest collision of the four — and its words are the longest, averaging 3.35 consonants. All 45,829 forms in the subset have a skeleton: none is lost.

Four results. The catalogue's widest-reaching code, **`m·r|hand`** with 69 languages, appears from one end of the continent to the other. The search for sibling codes yields 653 pairs of which only 5 % share languages, against 25 % in Uralic and 39 % in Turkic: the cleanest split of the series. Geography gives a monotonic gradient across all six well-populated buckets, from 1.05 to 0.15. And the closure of substitutions, which had produced nothing for two families running, here yields three clean classes that turn out to be places of articulation.

We also report the third consecutive negative result: automatic affix discovery raises grouping by twelve points and the null by twenty-six — that is, it worsens the advantage over chance.

1. What is being looked for, and why Pama-Nyungan

Historical linguistics has a powerful and expensive instrument: the comparative method. It has been applied with enormous unevenness — Indo-European has dense reconstruction and expert cognacy, most families have neither — so any study resting on the historical layer of the lexicon will measure, unintentionally, the density of the comparative literature rather than the past of the languages.

The series this work belongs to asks whether something useful can be done by giving up reconstruction: not improving or replacing it, but doing without it and seeing what remains.

Pama-Nyungan was chosen for three properties of the catalogue it produces, not for its classification: it is the corpus family with the lowest collision (34 %), the highest arity (3.35 consonants per form), and zero form loss — all 45,829 of its forms have a skeleton, against the 18 % lost in Uralic and 21 % in Turkic. That is: the place where the instrument works under the best conditions, and therefore where one best sees how far it reaches when nothing gets in its way.

1.1 Order matters: catalogue first, genealogy last

The catalogue is built without genealogy. Codes come from three things and only three: a form's consonant skeleton, the concept it expresses, and a threshold of three distinct languages. No branch

classification, no protoform, no cognate list enters. The procedure does not know this family has branches.

Genealogy enters afterwards, and only to read what came out. When §5 crosses codes with branch labels, it does so on an already-closed catalogue, and the question is not ”which codes does the classification predict?” but the reverse: ”does the catalogue, on its own, split in any recognisable way?”

That ordering is what makes the instrument usable on families where no reconstruction exists, which is what it was designed for.

1.2 What this work does not attempt

We do not reconstruct protoforms. We propose no ancestral form, not even implicitly.

We do not claim relatedness. Two words sharing a code does not mean they descend from a common one.

We do not distinguish inheritance from borrowing. Without protoforms this is impossible. A group of words sharing a code may be inheritance, borrowing, convergence, or coincidence. The catalogue does not choose.

We do not attribute meaning to a code. We write `m·r|hand` and never `m·r = hand`.

And we do not compare Pama-Nyungan with any other family: the comparisons in §5.6 are about the instrument’s behaviour, not about the languages.

Where this procedure sits in the automatic-comparison literature — Dolgopolsky, Turchin/Peiros/Gell-Mann, ASJP — is covered in the method paper.

2. What a code is, and what it costs here

2.1 The consonant skeleton

The skeleton is a word’s consonant sequence, with vowels removed — not lost: they are kept in a separate column — and each consonant written with a canonical representative. Pama-Nyungan uses 24 canonical symbols, 16 of which merge something. The full table is in `runs/normalizacion.md` and is derived by aligning each form with its skeleton, not written by hand.

The 24 symbols, needed to read §5:

`n r k l m p t j w c ŋ d ɲ ʈ q g b h ɳ s ɭ x ɾ ʊ`

What matters when reading the results: retroflexes and palatals are preserved (`ʈ`, `ɳ`, `ɭ`, `ɾ`, `ɲ` each have their own symbol), and that decision matters because these languages’ inventories are organised largely on those contrasts. Length variants and affricates are merged.

2.2 A code is a skeleton that recurs across languages

A code is a skeleton appearing, for the same concept, in three or more distinct languages. It is written `m·r|hand`, never `m·r = hand`: the code does not mean ”hand”, it is the shape ”hand” repeatedly takes.

2.3 What the instrument costs: collision

34 % of Pama-Nyungan forms share a skeleton, within their own language, with a form of another concept. This is the clean measure of what the skeleton fails to distinguish, and it is the lowest of the four families in the series (Turkic 59 %, Uralic 59 %, Austronesian 46 %).

The reason lies in arity: 3.35 consonants per form, also the highest. With one more consonant on average than Austronesian there is more room to distinguish.

Even so, 34 % weighs, and the appendix's reverse index shows it: `m·n` appears with ten concepts — "what", "neck", "night", "fat", "ear" — and `n·n` with nine. These are nasal combinations, which are abundant in these languages.

2.4 Arity

Arity	Codes	
1 (unary)	82	3.9 %
2 (binary)	835	39.9 %
3 (ternary)	728	34.8 %
4 or more	447	21.4 %

The figures this paper defends are restricted to the 1,563 binary and ternary codes.

Two things set this distribution apart from the other families. Unaries are 3.9 %, a third of what they were in Uralic or Turkic: with long words there are few one-consonant ones. And those with four or more are 21.4 %, twice Uralic's — which is what the same cause predicts, and §5.3 returns to it.

3. Method

3.1 Data

Integrative Corpus, Pama-Nyungan subset: 154 languages, 45,829 forms with transcription, assigned concept, and skeleton.

No form is lost. This is the only family in the series where the raw and published counts coincide: all 45,829 forms with a concept have a skeleton. Uralic lost 18 % and Turkic 21 %.

No language is excluded. Phase 0 (`runs/fase0-higiene.md`) found no geographic intruders, no tone marks, no adjacent names. It did find three pairs of very close languages — Dhuwal with Djambarrpuyngu (73 % identical skeletons) and with Djapu (72 %), and Mayi-Kulan with Ngawun (65 %) — but none crosses the 80 % threshold declared in the Turkic paper, and applying a lower threshold here would mean using two different criteria across families. They are declared and kept.

A single macroarea, so this paper has no geographic-anomaly section.

3.2 Enrichment

The code's coverage divided by its base frequency; the full definition, and the warning that it is not comparable across families, are in the method paper.

3.3 Nulls

Every grouping is compared against a permutation shuffling concepts within each language. Real grouping covers 38.8 % of forms; the null, 11.3 %.

3.4 One population for the whole paper

The seven reports in `runs/` that declare it all print the same line:

```
POPULATION · Pama–Nyungan: 154 languages · 45,829 forms · 1,160 concepts · 15 branches ·
             segmentation=yes(0) · excluded=0 · threshold=3 languages
```

This is checked mechanically (`analysis/tools/verifica_paper.py`, check 7).

4. Results

4.1 The catalogue

2,092 codes over 360 of the 1,160 concepts, grouping 11,171 forms.

Quantity	Value	Over which population
Forms inside a code	24.4 % (11,171 of 45,829)	3-language threshold, the catalogue's
Forms in any group	38.8 % (null: 11.3 %)	2-language threshold, grouping test
Precision	5.87 % (chance: 0.298 %)	pairs of forms from different languages
Times over chance	20× [19–20; chance is computed exactly]	same
Binary and ternary	1,563 (74.7 %)	of the 2,092 codes

Each row carries its population because there are three distinct ones.

On the interval, and on an error this family helped find. The preceding papers published wide ranges — [172–227] for Turkic, [73–88] for Uralic — and attributed them to the data. They were not the data's: chance was estimated with eighty thousand draws, and at a base rate around 0.1 % that is some eighty hits, so the figure swung on our own sampling. Chance is now computed exactly, in closed form over the distribution of concepts and languages, and the series' three figures were corrected: 225× → 216×, 77× → 79×, 39× → 37×.

With the null exact, the only remaining variation is the subsampling of large groups, and Pama-Nyungan is the one family of the four where that variation shows: [19–20], against a single point in the other three. It has the largest groups in the series, so it subsamples the most. This is the opposite of what an earlier version of this paragraph claimed.

![Distribution of codes by arity. Unaries are 3.9 %, a third of what they were in the preceding families: with long words there are few one-consonant ones.](../output/fig1-aridad.en.svg)

4.2 The widest-reaching codes

Concept	Code	Languages	Enrichment	Branches it appears in
hand	m·r	69	53×	Karnic, South-West, Southeastern, Greater Maric and others
FAECES	k·n	47	66×	six branches
WE TWO (INCL.)	n·l	46	109×	five branches
CAMP	n·r	42	42×	six branches
WHO	n·n	38	33×	seven branches
eye	m·l	34	43×	five branches
I	n·j	30	109×	five branches
BREAST	n·m·n	29	87×	four branches

m·r|hand, with 69 languages, is the widest-reaching code in the whole catalogue — more than twice the second. Its forms are *mara*, *marra*, *maRa*, *murra*, and they appear from one end of the continent to the other.

The branch column describes how each code's languages are distributed; it does not claim that distribution has a historical cause, which is a different question and not this method's.

Figure 1: The twenty widest-reaching binary and ternary codes.

5. What this family teaches

5.1 Ninety-six branch pairs

The Turkic paper established that the substitution matrix must be computed per branch pair, since aggregating over the family dilutes what distinguishes the groups; Turkic could only test this on one cut and Uralic on thirty-six. Here there are ninety-six pairs with data.

One example, the Karnic ↔ Paman cut:

Segment	Main	Cases	Share when crossing	Share when not	Diagnostic
q	n	39 of 86	45 %	4 %	12.0×
h	r	134 of 234	57 %	11 %	5.1×
m	n	245 of 552	44 %	24 %	1.9×
b	p	24 of 49	49 %	40 %	1.2×

The last column compares how much more concentrated an alternation is when crossing the cut than within a branch. **q** **n** comes out at 12.0×, the highest value the series has produced on this measure.

What it is and what it is not. It is a measure of where an alternation concentrates, computed on the catalogue. It is not a sound correspondence: a correspondence requires environment, direction, and regularity, and none of the three is present here. What that concentration means is a question for someone with the family's historical phonology in front of them.

Three caveats. The table is one cut of ninety-six and is trimmed to the highest-share rows; the rest is in `runs/fase6-por-rama.md`. And the "cases" are pairs of forms, not independent observations: the figure points at where to look, it is not a test.

5.2 Sibling codes: the cleanest split of the series

The search for pairs of codes for the same concept differing in a single position (`analysis/fases/codigos_hermanos.py`) yields 653 binary and ternary pairs here. And the filter that discarded a quarter of them in Uralic discards almost nothing here:

	Pama-Nyungan	Uralic	Turkic
Pairs	653	156	297
Overlap in language	35 (5 %)	39 (25 %)	113 (38 %)

95 % of pairs are disjoint: no language appears in both codes, so each pair splits the languages into two groups without exception. It is by far the cleanest result of the series, and the most likely reason is the one in §3.1: there are no sources here with incompatible transcriptions, because there are no lost forms and no duplicated languages.

"Hand" is the richest case in the catalogue. The concept has five codes:

Code	Languages	Where
`m·r`	69	Karnic, South-West, Southeastern, Greater Maric
`m·l`	6	Paman, Greater Maric, Desert Nyungic
`m·h`	4	Paman only
`m·n`	4	Southeastern only
`b·r`	3	Southeastern, Rockhampton-Gladstone

Four of the five share the initial `m` and differ in the second position. Two of them are confined to a single branch. That is the observation, and this paper goes no further.

`WHAT` gives another sharp split: `m·n` (23 languages, mostly Karnic) against `n·n` (29, mostly Paman and Desert).

5.3 Automatic segmentation makes things worse, for the third time

Step	In a group	Null	Advantage	κ	Codes of 4+ consonants
Whole word	38.8 %	11.3 %	27.5	0.310	38.3 %
+ discovered affixes	50.6 %	37.6 %	13.0	0.208	13.6 %

The κ column is the Hubert-Arabie correction, `(observed − null) /(1 − null)`: it normalises by the attainable ceiling and so allows comparison across families where an advantage in points does not — ten points over a 5% null and ten over a 44% null are not worth the same. Here $\Delta\kappa = -0.102$. All eight families in the series give the same sign: Turkic −0.009, Afro-Asiatic −0.004, Sino-Tibetan −0.020, Uralic −0.035, Indo-European −0.094, Pama-Nyungan −0.102, Trans-New-Guinea −0.118, Austronesian −0.151.

(This list had six families and the figures of the time. It was rebuilt with all eight on 2026-09-09 from `pipeline/tools/tabla_serie.py`. The magnitude does NOT grow with the raw gain: Sino-Tibetan, with the shortest word of the eight, falls less than Indo-European.)

Twelve points of signal, twenty-six of noise. The advantage over chance falls to less than half, however spectacular the quaternary column may be.

This is the third consecutive family where it happens, after Turkic and Uralic, and the three have different morphologies. That makes the finding a property of the procedure: discovering affixes by

shape similarity inflates the null more than the signal, because the same trimming that brings related words together brings unrelated ones together too.

Here the discoverer also finds 774 prefixes. Pama-Nyungan is suffixing, so the number does not describe these languages' morphology — it describes what an unsupervised segmenter does with a consonant inventory.

The published catalogue uses no automatic segmentation.

5.4 The substitution closure, silent for two families, produces here

In Austronesian, transitively closing the substitutions that survive the validation half gave three classes. In Turkic the segments collapsed into one; in Uralic likewise. We wrote then that the phase "worked in Austronesian and has not worked since."

It works again here, with a structure that reads itself:

	Segments
1	c d j t ʈ
2	g k
3	b p

Three classes, and all three group segments from the same place of articulation: coronals and palatals in the first, velars in the second, labials in the third. The procedure knows no articulatory feature — it only counts which positions two skeletons differ in — and the closure separates them anyway.

That it appears here and not in the previous two has a candidate explanation the catalogue itself suggests: the Pama-Nyungan inventory is organised on place-of-articulation contrasts — four or five coronal series — and §2.1 preserves them unmerged. Where the normalisation preserves the dimension the family uses, the closure finds it.

5.5 Geography, with the best gradient of the series

Shared codes	Pairs	Mean distance	÷ chance
1	4,035	1,743 km	1.05
2	1,957	1,764 km	1.06
3–5	2,196	1,583 km	0.95
6–10	599	987 km	0.59
11–20	202	459 km	0.28
21+	153	250 km	0.15
two random languages	40,000	1,660 km	1.00

Monotonic across all six buckets and with large samples: the smallest has 153 pairs and the largest 4,035, against the 2 and 6 pairs of Uralic's extreme buckets. Languages sharing more than twenty codes sit at a sixth of the expected distance.

It is the cleanest gradient the series has produced, and it is worth saying why it can be trusted more than the previous ones: it depends on no small bucket.

5.6 Four families

With the four measured by the same code and the same definition (`ci/metricas.py`):

Family	Lgs	Arity	Collision	Precision	Chance	× chance
Turkic	32	2.85	59 %	25.48 %	0.118 %	216×
Uralic	27	3.28	59 %	6.53 %	0.083 %	79×
Austronesian	582	2.78	46 %	13.41 %	0.363 %	37×
Pama-Nyungan	154	3.35	34 %	5.87 %	0.298 %	20×

Two observations, and both are about the instrument:

Arity predicts collision. The four families fall in order: 2.79 → 46 %, 2.85 → 59 %, 3.28 → 59 %, 3.35 → 34 %. It is not perfect — Turkic and Uralic tie on collision with different arities — but the two extremes are the expected ones. It is the relationship `docs/PERFIL-FAMILIAS.md` measured across sixty families, and here it is visible in the four we have worked in depth.

The ratio over chance orders by nothing obvious. Pama-Nyungan, where the instrument costs least, gives the lowest ratio. That is enough to discard the easy reading — "higher ratio, better family" — and to recall what the Turkic paper established: the ratio is a quotient, and its denominator depends on the family. Precisions and chances are published separately, in their own columns, and anyone wanting to compare has both halves.

6. Limits

Collision, lowest of the series though it is, is 34 %. A third of forms share a skeleton with another word in their own language.

There are no zero reflexes. The §5.1 procedure compares skeletons of equal length, so it cannot see any alternation involving the loss of a segment.

A single macroarea, so there is no geographic-anomaly list.

Only 360 of the 1,160 concepts produce any code. With 154 languages but wordlists of uneven coverage, most concepts do not gather three languages with the same skeleton.

21.4 % of codes have four or more consonants, twice Uralic's. §2.4 says this is grounds for doubting the analysis itself, and §5.3 shows automatic segmentation is not the way out. Morphological segmentation remains unresolved, and here it weighs more than in any other family of the series.

We do not calibrate against expert cognacy. Comparative work on this family exists and we have not used it. As in the previous families, this is a gap and not an impossibility.

7. Conclusion

The Pama-Nyungan catalogue groups 24.4 % of forms into 2,092 codes with a precision twenty times chance, with chance computed exactly.

What this family contributes is having been the place where the instrument works under the best conditions, and having shown what changes when it does. Sibling codes split cleanly: 95 % of the 653 pairs share not one language, against Turkic's 61 %. The geographic gradient is monotonic across all six buckets and none depends on a small sample. And the substitution closure, silent for two families, here gives three classes that turn out to be places of articulation — without the procedure knowing any.

And a confirmation that is no longer of one family but of three: automatic affix discovery worsens the advantage over chance, in Turkic, in Uralic, and here. Three different morphologies, the same result.

An instrument that claims little can be applied where there is no literature. It is worth applying it also where conditions favour it — because only then does one learn how much of what it fails to find is its own fault.

References

Bowern, C. and Atkinson, Q. (2012). Computational phylogenetics and the internal structure of Pama-Nyungan. *Language* 88(4).

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.

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

Key, M. R. and Comrie, B. (eds.). *Intercontinental Dictionary Series*.

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

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

List, J.-M., Greenhill, S. J. and Gray, R. D. (2017). The potential of automatic word comparison for historical linguistics. *PLOS ONE* 12(1).

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	`families/pama-nyungan/output/catalogo-pama-nyungan.tsv` — 2,092 rows
Per-form detail	`families/pama-nyungan/output/catalogo-pama-nyungan-lenguas.tsv` — 10,190 rows
Hygiene (§3.1)	`analysis/fases/fase0_higiene.py` → `runs/fase0-higiene.md`
Loading, population and warnings	`ci/familia.py`

Headline figures (§4.1, §5.6)	`ci/metricas.py`
Construction	`analysis/fases/catalogo_familia.py`
Grouping and nulls (§3.3, §4.1)	`analysis/fases/agrupamiento_familia.py` → `runs/agrupamiento.md`
Segmentation (§5.3)	`analysis/fases/fase2_despiece.py` → `runs/fase2-despiece.md`
Substitution closure (§5.4)	`analysis/fases/fase6_clases_emergentes.py` → `runs/fase6-clases.md`
Substitutions per branch pair (§5.1)	`analysis/fases/fase6_por_rama.py` → `runs/fase6-por-rama.md`
Sibling codes (§5.2)	`analysis/fases/codigos_hermanos.py` → `runs/codigos-hermanos.md`
Codes and geography (§5.5)	`analysis/fases/codigos_vs_parentesco.py` → `runs/codigos-parentesco.md`
Normalisation table (§2.1)	`analysis/fases/tabla_normalizacion.py` → `runs/normalizacion.md`
Appendix A	`analysis/fases/anexo_codigos.py`
Cross-family mixing audit	`analysis/tools/audita_familias.py`

Data sources: Lexibank (List, Forkel, Greenhill et al. 2022, CC-BY-4.0), IDS (Key & Comrie, CC-BY-4.0), Concepticon and Glottolog (CC-BY-4.0).

Appendix A · One hundred annotated codes

The catalogue has 2,092 codes and the complete data is in `families/pama-nyungan/output/`. What follows is a readable sample: the hundred widest-reaching binary and ternary codes, each with up to sixteen languages and their words.

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

This appendix does carry an emergent-code column, unlike the Turkic and Uralic ones. The substitution closure (§5.4) yields three classes here — 1 = c d j t ʈ, 2 = g k, 3 = b p — and the column rewrites each code with them: `m·r` is unchanged because neither `m` nor `r` falls in any class, but `k·n` reads `2·n`. The classes are of this family; the Austronesian ones are not valid here, nor the reverse.

The comments in this appendix describe what is in each block — how many languages, how they are distributed, what varies and what does not — and not what any form descends from. That last is the work of the comparative method, which this paper does not do.

Reverse index: codes recurring across several concepts

This table is longer than in the three preceding families, and it is worth understanding why before reading it as a defect.

The recurring codes are almost all nasal-plus-liquid combinations — `m·n`, `n·n`, `n·r`, `n·l`, `m·r`, `n·m` — covering between three and ten concepts each. That is not a peculiarity of the catalogue but of the inventory: Pama-Nyungan languages have consonant systems with many nasals and laterals and few fricatives, so the space of short skeletons is more crowded than in the other families.

Note, however, that the measured collision — 34 %, §2.3 — is the lowest of the series. The two coexist because the appendix shows the widest-reaching extreme, where short codes concentrate, while collision is measured over all forms.

Code	Concepts it appears with, and their enrichment
`m·n`	what (23, 22×) · neck (18, 19×) · night (14, 12×) · fat (organic substance) (13, 11×) · flesh or meat (11, 10×) · ear (11, 9×) · fly_v (11, 9×) · throat (10, 10×) · ant (10, 10×) · good (10, 8×)
`n·n`	who (38, 33×) · what (29, 23×) · we (exclusive) (25, 27×) · sit (24, 17×) · that (20, 18×) · i (18, 13×) · you (17, 13×) · this (16, 13×) · see (11, 8×)
`n·r`	camp (42, 42×) · thou (20, 23×) · nose (16, 14×) · you (10, 10×) · hear (10, 10×) · name (10, 12×)
`n·l`	we two (inclusive) (46, 109×) · forehead (20, 29×) · he or she (15, 25×) · belly (10, 13×)
`m·r`	hand (69, 53×) · spearthrower (15, 13×) · black (12, 10×)
`n·m`	breast (29, 51×) · mother (11, 20×) · see (11, 20×)
`m·j`	vegetables (22, 81×) · fruit (17, 56×) · house (15, 49×)
`j·r`	tooth (21, 41×) · thou (18, 46×) · sharp (11, 32×)
`w·n·k`	be alive (16, 52×) · speak (14, 41×) · say (10, 33×)
`k·l`	fire (13, 25×) · boomerang (11, 22×) · louse (11, 27×)

Numerals

A single numeral makes the cut: `b·l·r|two`, with 21 languages.

This is a fact about the corpus, not the instrument: the lists this data comes from record few numerals for these languages, and "three" and above do not gather three languages with the same skeleton. Where Turkic gave a complete series, here there is one.

b·l·r|two · ternary · 14 languages · enrichment 107× · emergent 3·l·r

Awabakal bulóara	Badjiri bulari	Biri bularu	Dharawal bullar
Dharuk Bool–air	Gamilaraay boolar	Gangulu bularu	Gooreng Gooreng boolar
Gudjal bulari	Gumbaynggir bulari	Guwa bulara	Warungu bulari
Yirandali boolooroo	Yuwaalaraay Bullara		

Body parts

The largest section of the appendix, and the one that best shows the distribution.

m·r|hand, with 69 languages, is the widest-reaching code in the whole catalogue — more than twice the second. The forms are *mara*, *marra*, *maRa*, *murra*, and they appear from one end of the continent to the other: Adnyamathanha in South Australia, Pitjantjatjara in the central desert, Nhanta on the west coast, Gunggari in Queensland. Sixty-nine languages from four branches saying nearly the same thing.

And §5.2 shows what this block cannot: "hand" has five codes in the catalogue. Alongside m·r (69) are m·l (6), m·h (4), m·n (4) and b·r (3), and the four small ones fall in different branches from the large one. The appendix shows the majority code by its reach criterion.

Three concepts appear with two codes each in this same section — "foot" with c·n, t·n and d·n, "nose" with m·l, m·l·j and n·r, "tongue" with d·l·n, t·l·n and t·l — and in all three the alternations are of place of articulation or initial voicing. This is the raw material of §5.2.

m·r|hand · binary · 69 languages · enrichment 53×

Adnyamathanha mara	Biri mara	Durubul mara	Gunggari mara
Kartujarra mara	Kukatja mara	Malgana mara	Narrungga mara
Ngamini maRa	Nhanta mara	Pallanganmiddang marra	Pitjantjatjara mara
Wangkangurru mara	Warriyangga mara	Wirangu murra	Yarluyandi mara

k·n|FAECES (EXCREMENT) · binary · 47 languages · enrichment 66× · emergent 2·n

Badjiri kunna	Dharuk kuni	Jiwarli kuna	Karuwali koona
Kurrama kuna	Mayi–Kulan kuna	Mithaka kuna	Ngaiawang kunna
Ngarrindjeri kuna	Payungu kuna	Pitjantjatjara kuna	Tharrgari kuna
Wangkayutyuru kuna	Warlpiri kuna	Warumungu kuna	Yarluyandi kuna

m·l|eye · binary · 34 languages · enrichment 43×

Bilinarra mila	Dhuwal mi:l	Djapu mi:l	Durubul mil
Gooreng Gooreng mil	Gumbaynggir mil	Gurindji mila	Kungkari miil
Mudburra mila	Ngadjumaya miil	Ngawun mili	Nhanta milu
Nyungar meal	Watjuk meal	Wirangu meil	Yagara mil

n·m|BREAST · binary · 29 languages · enrichment 51×

Adnyamathanha ngama	Arabana ngama	Durubul ngamu	Garlali ŋama
Guyani ngama	Martu Wangka ngama	Mithaka ngama	Narrungga ngammi
Ngiyambaa Nummu#	Nyangumarta ngama	Parnkala numma	Pirriya ngama
Warluwarra ngama	Yalarnnga ngama	Yarluyandi ngama	Yugambeh nguma

n·m·n|BREAST · ternary · 29 languages · enrichment 87×

Badjiri ngamun	Batyala ŋamuŋ	Dhangu ŋamini	Djabugay ngamun
Djapu ŋamini	Dyirbal ngamun	Gangulu ŋamun	Gumatj ŋamini
Gupapuyngu ŋamini	Jaru namana	Karajarri ngamarna	Margany ngamun
Pitta-Pitta ngamanya	Wargamay ngamun	Wiradjuri namong	Yan-nhangu ŋamun

c·n|foot · binary · 29 languages · enrichment 78× · emergent 1·n

Badimaya jina	Dyirbal jina	Jiwarli jina	Karajarri jina
Kartujarra jina	Kurrama jina	Malgana chinna	Ngarla cina
Northern Mangarla jina	Nyangumarta jina	Panyjima jina	Thalanyji jina
Wakaya jinna	Warnman jina	Warumungu jina	Watjuk jenna

k·m·p|URINE · ternary · 27 languages · enrichment 145× · emergent 2·m·3

Gurindji kumpu	Jiwarli kumpu	Kaanju kumpu	Kartujarra kumpu
Kugu Nganhcara kumpu	Martu Wangka kumpu	Ngarla kumpu	Ngarluma kumpu
Nyangumarta kumpu	Pakanh kumpu	Panyjima kumpu	Pintupi-Luritja kumpu
Umpila kumpu	Wajarri kumpu	Warnman kumpu	Wik Mungkan kump

g·n|FAECES (EXCREMENT) · binary · 24 languages · enrichment 40× · emergent 2·n

Biri guna	Bularnu guna	Dharuk goone	Dharumbal guna
Dyirbal guna	Gamilaraay goona	Gudjal guna	Gumbaynggir guuna
Kukatja guna	Margany guna	Ngiyambaa guna	Nyungar goonna
Wajarri gurna	Wargamay guna	Yabula Yabula gunu	Yindjibarndi guna

m·l|nose · binary · 22 languages · enrichment 28×

Diyari mulha	Jiwarli mulha	Karajarri milya	Malgana moola
Mirniny mulha	Mithaka mula	Ngarluma mulha	Northern Mangarla milya
Pitjantjatjara mu̱la	Pitta-Pitta milya	Thalanyji mulha	Wangkangurru moola
Wangkayutyuru milha	Wangkumara mulha	Wirangu moola	Yandruwandha mulha

t·n|foot · binary · 21 languages · enrichment 45× · emergent 1·n

Awabakal tinna	Bunganditj tena	Dharuk tunna	Diyari tina
Guwa thina	Karuwali tina	Kungkari thina	Mayi-Thakurti tunna
Mithaka tina	Muruwari thina	Nhirrpi thina	Paakantyi tinna
Pitta-Pitta thina	Wangkayutyuru thina	Wangkumara thina	Yandruwandha thina

j·r|tooth · binary · 21 languages · enrichment 41× · emergent 1·r

Badimaya yira	Darkinyung yira	Gamilaraay yira	Gundungurra Yerra
Jiwarli yirra	Kariyarra yirra	Kartujarra yirra	Martu Wangka yirra
Ngarluma jirra	Nyamal yirra	Payungu yirra	Pirriya yira
Thalanyji yirra	Wargamay yira	Warnman yirra	Warriyangga yirra

n·l|FOREHEAD · binary · 20 languages · enrichment 29×

Badimaya ngalha	Badjiri ngulu	Dharuk nu–la	Dhudhuroa ngulua
Djabugay ngulu	Gamilaraay ngulu	Gundungurra ŋulu	Kukatja nala
Muruwari ngulu	Ngadjumaya ŋarla	Ngiyambaa ŋulu	Pitta–Pitta ngurlu
Warungu ngulu	Wik Mungkan ngul	Wulguru ngulu	Yawarrawarrka ngurlu

n·n·k|BEARD · ternary · 18 languages · enrichment 52× · emergent n·n·2

Diyari nganka	Guwa nganka	Jiwarliny ngarnku	Kariyarra ngarnka
Kungkari nganka	Mudburra ngarnka	Ngarla ngarnka	Ngarluma ngarnka
Nyamal ngarnka	Nyangumarta ngarnka	Parnkala ngarnka	Payungu ngarnka
Pirriya nganka	Thalanyji ngarnka	Wangkangurru nganka	Wangkumara nganka

m·n|neck · binary · 18 languages · enrichment 19×

Biri manu	Dhuwal mayaŋ	Djabugay manu	Djambarrpuyngu mayaŋ
Djapu mayaŋ	Djinang mani	Gangulu manu	Gupapuyngu mayaŋ
Kuku Yalanji manu	Margany manno	Mayi–Kulan mana	Mayi–Yapi mana
Ngawun mana	Pakanh mannu	Parnkala moorni	Wik Mungkan man

m·l·j|nose · ternary · 18 languages · enrichment 128× · emergent m·l·1

Badimaya mulya	Karajarri milya	Kartujarra mulya	Kukatja mulya
Martu Wangka mulya	Ngaanyatjarra mulya	Northern Mangarla milya	Nyangumarta milya
Pitjantjatjara mulya	Pitta–Pitta milya	Wajarri mulya	Wangkatja mulya
Wardandi molya	Warlmanpa mulyu	Warlpiri mulyu	Warnman mulya

d·l·n|tongue · ternary · 17 languages · enrichment 141× · emergent 1·l·n

Badjiri dhalaŋ	Biri dalany	Dharawal dhallañ	Dharumbal dalang
Gangulu dalany	Gundungurra Darline	Gunggari dalany	Gunya dhalañ
Margany dhalañ	Narrungga dhallanya	Ngaiawang dallan	Ngarigu dhalany
Ngunawal dalling	Nyungar dallan	Ritharrngu dhiilng	Wulguru dhalany

n·r|nose · binary · 16 languages · enrichment 14×

Bularnu ngurru	Dhangu ngurru	Dhayyi ŋurru	Dhuwal ngurru
Djambarrpuyngu ngurru	Djapu ngurru	Djinang ngurri	Gupapuyngu ngurru
Muruwari ŋuru	Ritharrngu ŋurru	Wakaya ngurru	Wargamay ŋuɽu
Warumungu ngurru	Yan–nhangu ŋurru	Yanyuwa ngurru	Yindjilandji ŋʊɾo

t·l·n|tongue · ternary · 16 languages · enrichment 154× · emergent 1·l·n

Dharuk Telang	Guwa thalaña	Kungkari thalany	Malgana tallanu
Mirniny thalany	Nhanta thalanyu	Paakantyi therlunnia	Payungu thalany
Pitjantjatjara t̪al̪i̪n̪	Thalanyji thalany	Wajarri thalany	Wangkumara thalinya
Wardandi talling	Warriyangga thalany	Watjuk talling	Wiradjuri talline

d·n|foot · binary · 15 languages · enrichment 35× · emergent 1·n

Adnyamathanha idna	Badjiri dhina	Biri dhina	Darkinyung dhana
Dharumbal dhina	Gamilaraay dinna	Gangulu dina	Gunggari dina
Gunya dhina	Margany dhina	Ngamini dina	Ngiyambaa dhina
Wulguru dhina	Yiningay dinna	Yirandali dinna	

m·n|FAT (ORGANIC SUBSTANCE) · binary · 13 languages · enrichment 11×

Arabana marni	Diyari mani	Garlali marni	Guyani mani
Karuwali murne	Malyangapa marni	Nhirrpi marni	Paakantyi murni
Parnkala munni	Wakaya manu	Wangkangurru mani	Wirangu mirnu
Yandruwandha marni			

d·r|UPPER LEG (THIGH) · binary · 12 languages · enrichment 34× · emergent 1·r

Badjiri dhurra	Biri dhara	Darkinyung dhara	Dharuk dara
Durubul dara	Gamilaraay dharra	Gundungurra darra	Gunya dharra
Margany dharra	Ngarigu dhara	Wulguru dharra	Yuwaalaraay dharra

t·l|tongue · binary · 12 languages · enrichment 38× · emergent 1·l

Bunganditj thali	Diyari tharli	Gamilaraay thulle	Guyani tali
Mithaka tharli	Ngamini tali	Ngarluma ṭali	Ngiyambaa tulle
Pirriya tharli	Pitta-Pitta tharli	Wangkayutyuru tharli	Yarluyandi tharli

m·n|FLESH OR MEAT · binary · 11 languages · enrichment 10×

Dharawal murni	Dyirbal miña	Gudjal minya	Guugu-Yimidhirr minha
Kaanju miña	Kugu Nganhcara minha	Kukatj miny	Kuku Yalanji miña
Kuuku-Ya’u minya	Umpila minya	Yidiny miña	

m·n|ear · binary · 11 languages · enrichment 9×

Badjiri maŋa	Dyirbal munga	Gunggari manga	Gunya manga
Guwa maŋa	Jaru manga	Kungkari manga	Margany manga
Pirriya manga	Yiningay munga	Yirandali munga	

b·n|ear · binary · 11 languages · enrichment 19× · emergent 3·n

Biri bina	Dharumbal bina	Djabugay bina	Gamilaraay binna
Gangulu bina	Mayi-Thakurti binna	Wargamay bina	Wulguru bina
Yagara binne	Yidiny binə	Yuwaalaraay Binna	

k·r|eye · binary · 11 languages · enrichment 14× · emergent 2·r

Jiwarli kuru	Kartujarra kuru	Martu Wangka kuru	Ngaanyatjarra kuru
Payungu kuru	Pintupi-Luritja kuru	Pitjantjatjara kuɽu	Thalanyji kuru
Tharrgari kuru	Wangkatja kuru	Warriyangga kuru	

m·n|THROAT · binary · 10 languages · enrichment 10×

Badjiri manu	Djabugay manu	Gunya manu	Guugu-Yimidhirr manu
Guwa manu	Kaanju manu	Kuku Yalanji manu	Kurtjar ma:n
Margany manu	Umpila manu		

w·l|UPPER LEG (THIGH) · binary · 10 languages · enrichment 19×

```
Jiwarli wulu     Kurrama wulu       Mudburra wurlu  Panyjima wulu
Payungu wulu     Thalanyji wulu     Tharrgari wulu  Warriyangga wulu
Warumungu warli  Yindjibarndi wulu
```

n·l|belly · binary · 10 languages · enrichment 13×

```
Jiwarli ngarlu    Karajarri ngarlu     Kariyarra ngarlu     Kurrama ngarlu
Ngarla ngarlu     Nyamal ngarlo        Nyangumarta ngarlu_  Panyjima ngarlu
Tharrgari ngarlu  Yindjibarndi ngarlu
```

m·k|bone · binary · 10 languages · enrichment 29× · emergent m·2

```
Diyari muku      Guyani muku         Kukatj muk       Kurtjar mu:k
Ngamini muku     Nhirrpi muku        Wangkumara muku  Yandruwandha muku
Yarluyandi muku  Yawarrawarrka muku
```

Kinship and person

The pronominal system appears with its duals, which is what distinguishes this family from the three preceding: n·l|WE TWO (INCLUSIVE) with 46 languages and 109× enrichment, p·l|THEY TWO, alongside n·j|I, n·n|WE (EXCLUSIVE) and n·r|THOU.

n·l|WE TWO deserves a look for its enrichment: 109× with 46 languages is high for a binary of nasal plus lateral, both frequent consonants. It comes from the fact that almost every language attesting that concept says it with the same skeleton: coverage drives the figure, not rarity.

The kinship vocabulary is dense — mother's brother, father's sister, wife — and reflects the lists the corpus comes from, which record kinship systems for these languages in detail.

n·l|WE TWO (INCLUSIVE) · binary · 46 languages · enrichment 109×

```
Badjiri ngulli  Dharuk ŋali       Djapu ŋali          Gupapuyngu ngali
Jaru ngaali     Karajarri ŋali    Kuku Yalanji ngali  Mathi-Mathi ngali
Ngarla ngali    Ngiyambaa ngali:  Panyjima ngali      Thalanyji ngali
Wakaya ngal_    Wik Mungkan ngal  Wulguru ngali       Yidiny ngali
```

n·n|WHO · binary · 38 languages · enrichment 33×

```
Badimaya ngana  Bilinarra ngana  Darkinyung ngaan  Durubul ngan
Gurindji ngana  Jiwarli ngana    Kukatj ngaan      Kurrama ngana
Mudburra ngana  Ngarluma ngana   Panyjima ngana    Thalanyji ngana
Wajarri ngana   Warnman ngana    Yagara ngan       Yindjibarndi ngana
```

n·j|I · binary · 30 languages · enrichment 109× · emergent n·1

```
Adnyamathanha ngayi  Bilinarra ngayi          Gumbaynggir ngaya  Gurindji ngayu
Guyani ngayi         Kariyarra ngayi          Kukatja ngayu      Kuuku-Ya'u ngayu
Mayi-Kulan ngayi     Mayi-Kutuna ngayu        Mudburra ngayi     Ngarla ngaya
Ngawun ngayu         Northern Mangarla ngayu  Warlmanpa ngayu    Yindjibarndi ngayi
```

n·n|WHAT · binary · 29 languages · enrichment 23×

Bunganditj nan	Dharuk ngang	Gunya ngani	Karajarri ngana
Kukatj nganya	Kurrama ngani	Kuuku-Ya'u ngaani	Mathi-Mathi nhangi
Mayi-Thakurti ngani	Ngawun ngani	Nyangumarta ngani	Pakanh nganhnhi
Wardandi nang	Warnman ngana	Wik Mungkan nge:n	Yalarnnga nhangu

n·n|WE (EXCLUSIVE) · binary · 25 languages · enrichment 27×

Dyirbal ngana	Gunya ngana	Kaanju ngana	Kuuku-Ya'u ngana
Mayi-Kulan ngana	Mayi-Kutuna ngana	Muruwari ngana	Ngarrindjeri nga:ne
Ngiyambaa Ne#ene	Nhirrpi nganu-	Northern Mangarla -ngani	Pitta-Pitta ngana
Umpila ngana	Warungu ngana	Wik Mungkan ŋan	Yawarrawarrka ngani

m·n|WHAT · binary · 23 languages · enrichment 22×

Arabana minha	Badjiri minna	Batyala minya	Diyari minha
Duungidjawu miña	Gamilaraay Minya	Gumbaynggir minya	Malyangapa minha
Mithaka mina	Ngamini minha	Ngunawal minya	Nhirrpi minha
Pitta-Pitta minha	Wangkumara minha	Yandruwandha minha	Yarluyandi minha

t·n|THEY · binary · 22 languages · enrichment 65× · emergent 1·n

Diyari thana	Jiwarli thana	Kalkatungu thina	Mabuiag tana
Mayi-Kulan thana	Mayi-Yapi thana	Muruwari thana	Ngamini thana
Payungu thana	Pitta-Pitta thana	Thalanyji thana	Wangkumara thana
Warriyangga thana	Western Arrarnta itne	Yandruwandha thana	Yarluyandi thana

n·n|THAT · binary · 20 languages · enrichment 18×

Adnyamathanha anhanha	Dhuwal ngunhi	Djambarrpuyngu ŋunhi	Djapu ŋunhi
Gudjal nguna	Gumatj ŋunhi	Gupapuyngu ŋunhi	Gurindji nginya
Mathi-Mathi nhunhi	Ngunawal niny	Nhirrpi nhunu	Nyamal ngunya
Payungu ngunha	Thalanyji ngunha	Tharrgari ngunha	Warriyangga ngunha

n·r|THOU · binary · 20 languages · enrichment 23×

Adnyamathanha nhurra_	Awabakal nú-ra	Jiwarli nhurra	Karajarri -nyurra
Malgana nhura	Martu Wangka -nyurra	Muruwari nhura	Ngarla ñurra
Nyangumarta nyurra	Payungu nhurra	Pitjantjatjara ñurra	Pitta-Pitta nhura
Tharrgari nhurra	Wajarri ñurra	Wargamay ñurra	Warlpiri =nyarra

j·n|name · binary · 20 languages · enrichment 53× · emergent 1·n

Gurindji yini	Kariyarra yini	Kartujarra yini	Kukatja yini
Martu Wangka yini	Mudburra yini	Ngarinyman yini	Ngarla yini
Northern Mangarla yini	Nyamal yini	Nyangumarta yini	Panyjima yini
Tharrgari yini	Wajarri yini	Warluwarra yini	Warriyangga yini

m·m|FATHER · binary · 18 languages · enrichment 66×

Badimaya mama	Bunganditj mam	Dhudhuroa mama	Kariyarra mama
Kartujarra mama	Kurrama mama	Malgana mamma	Mathi-Mathi maami
Ngadjumaya mama	Nyamal mama	Pallanganmiddang mama	Panyjima mama
Payungu mama	Pintupi-Luritja mama	Pitjantjatjara mama	Wajarri mama

n·n|I · binary · 18 languages · enrichment 13×

Badjiri ngunyi	Dhudhuroa ninna	Diyari nganhi	Gudjal nganya
Gunya nganha	Guugu-Yimidhirr nganhi	Kuku Yalanji ngaña	Kurtjar nga:nh
Margany nganha	Mithaka nganyi	Ngaiawang nganna	Nhirrpi nganyi
Wangkumara nganyi	Wirangu ngunna	Yandruwandha nganyi	Yanyuwa ngarna

j·r|THOU · binary · 18 languages · enrichment 46× · emergent 1·r

Diyari yura	Gunggari yurra	Gunya yura	Guugu-Yimidhirr yurra
Guwa yurra	Kukatj yir	Kuku Yalanji yurra	Mayi-Kulan yarra
Mithaka yura	Ngawun yarra	Nhirrpi yurra	Wangkumara yurra
Warungu yurra	Yandruwandha yura	Yanyuwa yirru	Yarluyandi yurra

n·n|YOU · binary · 17 languages · enrichment 13×

Adnyamathanha nhinha	Dhangu nhonu	Djinang ñuni	Durubul ngin
Duungidjawu ŋin	Gooreng Gooreng ngin	Guyani nina	Kaanju nguna
Kalkatungu nyini	Kuuku-Ya'u ngunu	Mathi-Mathi nginna	Narrungga nhinni
Nhanta nyini	Parnkala ngina	Umpila ngunu	Yan-nhangu nhunu

n·n·t|YOU · ternary · 17 languages · enrichment 54× · emergent n·n·1

Gurindji nyuntu	Karajarri nyuntu	Kariyarra ñinta	Kartujarra nyuntu
Kurrama ñinta	Martu Wangka nyuntu	Mirniny ngurntu	Muruwari ñintu
Ngarinyman nyuntu	Northern Mangarla nyuntu	Pakanh nhinta	Panyjima ñinta
Pintupi-Luritja nyuntu	Pitjantjatjara nyuntu	Warlmanpa nyuntu_	Wik Mungkan nint

n·t|I · binary · 16 languages · enrichment 42× · emergent n·1

Arabana antha	Bunganditj ngathu	Dharumbal ngatha	Garlali ŋathu
Jiwarli ngatha	Mabuiag ngath	Malgana natha	Muruwari ngathu
Ngiyambaa Na#ttu	Nyamal ngatha	Payungu ngatha	Thalanyji ngatha
Wajarri ngatha	Wangkangurru anta	Warriyangga ngatha	Yabula Yabula ngatha

n·n|THIS · binary · 16 languages · enrichment 13×

Alyawarr nhenh	Antekerrepenhe nhenh	Badimaya nhinha	Gundungurra Nin
Gunya inany	Mithaka nuniyi	Ngarigu niny	Ngunawal nguna
Parnkala nanni	Pintupi-Luritja ngaanya	Pitjantjatjara nyanga	Wajarri ñaña
Warnman nyarni	Western Arrarnta nhanhe	Yalarnnga ninyi	Yan-nhangu nhaŋu

n·n·d|YOU · ternary · 16 languages · enrichment 59× · emergent n·n·1

Bilinarra nyundu	Darkinyung ngindi	Dharumbal nginda	Dyirbal nginda
Gumbaynggir ngi:nda	Guugu-Yimidhirr nhundu	Jaru nyundu	Malgana nanda
Mudburra nyundu	Ngiyambaa Nindu	Payungu nyinda	Thalanyji nyinda
Wajarri nyinda	Watjuk nunda	Yidiny ñundu	Yindjibarndi nyinda

n·l|HE OR SHE · binary · 15 languages · enrichment 25×

Dharumbal nhula	Gunggari ngula	Gunya nhula	Guugu-Yimidhirr ñulu
Kaanju nhila	Kuku Yalanji ñulu_	Kuuku-Ya'u ngulu	Margany nhula
Mayi-Kulan nhulu	Mayi-Kutuna nulu	Mayi-Yapi nhulu	Umpila nhulu
Warungu nyula	Wik Mungkan nil	Yugambeh neule	

p·l|THEY TWO · binary · 14 languages · enrichment 47× · emergent 3·l

Guwa pula	Jiwarli pula	Mabuiag ipal	Mayi-Kulan pala
Mithaka pula	Muruwari pula	Ngawun pala	Nhirrpi pula
Payungu pula	Pitta-Pitta pula	Thalanyji pula	Wangkatja pula
Warlpiri pala	Wik Mungkan pul		

n·p|WIFE · binary · 14 languages · enrichment 56× · emergent n·3

Arabana nupa	Karajarri nyupa	Kariyarra nyupa	Kartujarra nyupa
Muruwari nupa	Ngarla ngoopa	Ngarrindjeri napa	Nhirrpi nhipa
Nyamal ngoopa	Panyjima nyupa	Wangkayutyuru nhupu	Warnman nyupa
Yandruwandha nhipa	Yawarrawarrka nipa		

n·m·n|MATERNAL UNCLE (MOTHER'S BROTHER) · ternary · 13 languages · enrichment 55×

Adnyamathanha ngamarna	Bilinarra ngamirni	Gurindji ngamirni	Guyani ngamarna
Jaru ngamirni	Mayi-Kulan ngamirn	Mayi-Thakurti ngamirn	Mayi-Yapi ngamirn
Mudburra ngamirni	Parnkala ngamarna	Warlmanpa ngamirni	Warlpiri ngamirni
Warumungu ngamirni			

k·k|MATERNAL UNCLE (MOTHER'S BROTHER) · binary · 12 languages · enrichment 55× · emergent 2·2

Aghu-Tharrnggala kekə	Arabana kaka	Diyari kaka	Kariyarra kaka
Kartujarra kaka	Martu Wangka kaka	Ngarinyman kaka	Ngarluma kaka
Northern Mangarla kaka	Wangkangurru kaka	Warluwarra kaka	Yarluyandi kaka

w·n·d|WHERE · ternary · 12 languages · enrichment 61× · emergent w·n·1

Badimaya wandi	Dharumbal wunda	Gundungurra Wunda	Gunya wanhdha
Guugu-Yimidhirr wanhdha	Malgana warnda	Mathi-Mathi winhdha	Mayi-Kulan wandha
Mayi-Thakurti wandhu	Ngawun wandhi	Ngunawal wanDa	Pallanganmiddang wanda

n·m|MOTHER · binary · 11 languages · enrichment 20×

Adnyamathanha ngummie	Djabugay ngama	Gurindji ngamayi	Guugu-Yimidhirr ngamu
Guyani ngami	Jaru ŋamai	Kukatja ngama	Kuku Yalanji ngamu
Parnkala ngami	Wangkayutyuru ngama	Warluwarra ngama	

m·k·l|PATERNAL AUNT (FATHER'S SISTER) · ternary · 11 languages · enrichment 162× · emergent m·2·l

Bilinarra mukurla	Gumatj mukul	Gupapuyngu mukul	Gurindji mukul	
Jaru mukul	Jiwarli mukul	Kariyarra mukul	Mudburra mukula	
Ngarinyman mukurl	Ritharrngu mukul	Thalanyji mukul		

w·r|WHO · binary · 11 languages · enrichment 17×

```
Arabana waRa     Djinang wari     Gumbaynggir waarru  Gupapuyngu wäri
Ngamini waRa     Nhirrpi waRa     Pitta-Pitta wara    Ritharrngu wara
Wangkumara wara  Yan-nhangu wara  Yarluyandi wara
```

j·n·d|YOU · ternary · 11 languages · enrichment 88× · emergent 1·n·1

```
Biri yinda            Garlali yundu   Mayi-Kulan yundu   Mayi-Kutuna yundu
Mayi-Thakurti yoondo  Ngarla yinda    Ngawun yundu       Warungu yinda
Yanyuwa yinda         Yiningay yinda  Yirandali yoondoo
```

w·n·c|WHERE · ternary · 10 languages · enrichment 79× · emergent w·n·1

```
Durubul winja        Duungidjawu wanja  Dyirbal wañja  Kartujarra wanyja
Martu Wangka wanyja  Ngarla wanyja      Nyungar wiñca  Warlmanpa wanji
Warungu wanyja       Yidiny wañja:
```

w·n·t|WHERE · ternary · 10 languages · enrichment 53× · emergent w·n·1

```
Jiwarli wantha       Kaanju wantu      Malyangapa wanta  Mayi-Yapi wanhtha
Narrungga wanti      Nhanta wantha     Pakanh wantu      Payungu wantha
Pitta-Pitta winhtha  Thalanyji wantha
```

n·r|YOU · binary · 10 languages · enrichment 10×

```
Bunganditj ngooro  Jaru nyurra               Jiwarli nhurra      Karajarri -nyurra
Ngaiawang ngurra   Northern Mangarla nyurra  Nyangumarta nyurra  Pitta-Pitta nhura
Tharrgari nhurra   Warriyangga nhurra
```

n·r|name · binary · 10 languages · enrichment 12×

```
Aghu-Tharrnggala ŋor  Biri nharri      Bunganditj nuri  Colac nere
Gunggari ngari        Gunya nhari      Kaurna narri     Margany nhari
Wathawurrung narri    Yugambeh ngarri
```

Nature and sky

n·r|CAMP with 42 languages is the block's second widest-reaching code, and "camp" is a concept the Australian lists record and those of the other three families do not. Each family's corpus brings the questions of its own surveys, and that shows here more than anywhere else.

k·l appears with "fire", "louse" and "boomerang": three unrelated concepts under one skeleton, which is the cost §2.3 declares.

n·r|CAMP · binary · 42 languages · enrichment 42×

```
Arabana nura        Darkinyung ŋarra  Gumbaynggir ngu:ra   Jaru ngurra
Kartujarra ngurra   Kungkari ngurra   Martu Wangka ngurra  Muruwari ngurra
Ngarluma ngurra     Nhanta nurro      Pirriya ngurra       Thalanyji ngurra
Wangkangurru ngura  Warlpiri ngurra   Warriyangga ngurra   Yandruwandha ngurra
```

n·p|water · binary · 16 languages · enrichment 60× · emergent n·3

Adnyamathanha ngapa	Diyari ngapa	Garlali ngap:a	Jiwarliny ngapa
Karajarri ngapa	Karuwali ngapa	Mithaka ngapa	Muruwari ngapa
Nhirrpi ngapa	Pitta-Pitta ngapu	Warlmanpa ngappa	Warlpiri ngapa
Warumungu ngappa	Yandruwandha ngapa	Yarluyandi ngapa	Yawarrawarrka ngapa

m·n|night · binary · 14 languages · enrichment 12×

Dhangu munha	Djapu munha	Gumatj munha	Gupapuyngu munha
Kartujarra munga	Kukatja munga	Kuuku-Ya'u maana	Martu Wangka munga
Ngaanyatjarra munga	Pintupi-Luritja munga	Pitjantjatjara muŋa	Warlmanpa munga
Warlpiri munga	Yan-nhangu munha		

k·l|fire · binary · 13 languages · enrichment 25× · emergent 2·l

Jiwarli karla	Kurrama karla	Malgana kalu	Malyangapa kal:a__kadla
Mirniny karla	Ngadjumaya kalla	Panyjima karla	Payungu karla
Thalanyji karla	Wardandi karla	Warriyangga karla	Watjuk kalla
Wirangu kulla			

k·k·r|EAST · ternary · 12 languages · enrichment 128× · emergent 2·2·r

Jiwarliny kakarra	Karajarri kakarra	Kartujarra kakarra	Martu Wangka kakarra
Mirniny kakarra	Northern Mangarla kakarra	Nyangumarta kakarra	Nyungar kakarr
Panyjima kakarra	Warlmanpa kakarra	Warnman kakarra	Warumungu kakurru

w·l·r|moon · ternary · 12 languages · enrichment 95×

Jiwarli wirlarra	Kariyarra wilarra	Kurrama wilarra	Ngarla wirlarra
Ngarluma wilarra	Nyamal Weellara	Panyjima wirlarra	Payungu wirlarra
Thalanyji wirlarra	Tharrgari wirlarra	Wajarri wilarra	Yindjibarndi wilarra

m·n·t|earth · ternary · 11 languages · enrichment 31× · emergent m·n·1

Dhangu munatha	Dhuwal munatha	Djambarrpuyngu munatha	Djapu munatha
Gumatj munatha	Gupapuyngu munatha	Ngaanyatjarra manta	Pintupi-Luritja manta
Pitjantjatjara manta	Wangkatja manta	Yan-nhangu munatha	

Animals

k·l|louse, w·r·n|CROW. A small block, and its size is set by the hundred-widest cut, not by the animal lexicon.

g·j|fish · binary · 20 languages · enrichment 112× · emergent 2·1

Biri guyu	Dhangu guya	Dharumbal guya	Djabugay guyu
Djapu guya	Djinang guyi	Dyirbal guya	Gamilaraay guya
Gunggari guyu	Gunya guyu	Gupapuyngu guya	Margany guyu
Ngaiawang gooeya	Ritharrngu guya	Wiradjuri guuya	Wirangu guya

w·k·r|CROW · ternary · 13 languages · enrichment 88× · emergent w·2·r

```
Jiwarli wakurra       Kariyarra Wakura   Karuwali wokeri      Kurrama wa:kurra
Mithaka wakiRi        Ngarla wa:kurra    Ngarluma wakurra     Nyangumarta wakurra
Panyjima wakurra      Thalanyji wakurra  Warriyangga wakurra  Yawarrawarrka wakiri
Yindjibarndi wakurra
```

w·g·n|CROW · ternary · 11 languages · enrichment 138× · emergent w·2·n

```
Awabakal wágun    Dharuk wugan       Durubul wagan      Gumbaynggir wa:gan
Ngaiawang waugan  Ngunawal wagun     Wiradjuri wa:ga:n  Wulguru wagan
Yagara waa-gun    Yirandali wagunna  Yugambeh waga:n
```

k·l|louse · binary · 11 languages · enrichment 27× · emergent 2·l

```
Jiwarli kulu    Kukatj kuul       Kurrama kulu   Ngarla kulu
Ngarluma kulu   Nyungar kuul      Panyjima kulu  Payungu kulu
Thalanyji kulu  Warriyangga kulu  Watjuk kolo
```

w·k·n|CROW · ternary · 10 languages · enrichment 93× · emergent w·2·n

```
Badjiri wakan     Gunya wakarn     Guwa wakarna        Kalkatungu wakan
Kungkari wakarn   Margany wakarn   Mayi-Kutuna wakarn  Muruwari wakan
Ngiyambaa wa:kan  Yiningay wakarn
```

m·n|ant · binary · 10 languages · enrichment 10×

```
Badimaya minga  Jiwarli minga        Kurrama minga    Ngaanyatjarra minga
Nhanta minga    Pitjantjatjara miŋa  Thalanyji minga  Tharrgari minga
Wajarri minga   Wangkatja minga
```

Plants and food

m·j appears with ”vegetables” (22 languages, 81×), ”fruit” (17) and ”house” (15). The first two are plausibly the same field; the third is something else falling in the same skeleton.

m·j|VEGETABLES · binary · 22 languages · enrichment 81× · emergent m·1

```
Adnyamathanha mayi  Biri mayi               Gangulu mayi     Kaanju mayi
Kartujarra mayi     Kuku Yalanji mayi       Kuuku-Ya’u mayi  Margany mayi
Ngarla mayi         Northern Mangarla mayi  Nyamal mayi      Pakanh mayi
Umpila mayi         Warlpiri miyi           Warumungu mayi   Wik Mungkan may
```

m·j|fruit · binary · 17 languages · enrichment 56× · emergent m·1

```
Jiwarliny miyi        Kaanju mayi      Kartujarra mayi         Kugu Nganhcara mayi
Kurtjar maay          Kuuku-Ya’u mayi  Margany mayi            Martu Wangka mayi
Ngaanyatjarra mayi    Ngarla mayi      Northern Mangarla mayi  Pakanh mayi
Pintupi-Luritja mayi  Umpila mayi      Warlpiri miyi*1*        Warumungu mayi_
```

Verbs of action and state

w·n·k covers ”be alive” (16 languages), ”speak” (14) and ”say” (10) — three neighbouring concepts with the same skeleton and enrichments between 33× and 52×.

And here something should be said about the segmentation §5.3 measures. Pama-Nyungan languages have verbal morphology with tense and aspect suffixes, and the corpus cites them in forms that sometimes carry them. The published catalogue does not apply automatic segmentation — §5.3 explains why — so part of what appears in these blocks as skeleton variation is suffix, not root.

n·n|sit · binary · 24 languages · enrichment 17×

```
Dhayyi nhina      Dhuwal nhina      Garlali nyina     Kartujarra nyina
Kukatja nyina     Malyangapa ninga  Mirniny ngina-    Ngadjumaya nyina-
Nhanta nyina      Nhirrpi nhina     Pirriya yinaangu  Ritharrngu nhiina
Warluwarra nyina  Warungu nyina     Yan-nhangu nyena  Yandruwandha nhina
```

w·n·k|BE ALIVE · ternary · 16 languages · enrichment 52× · emergent w·n·2

```
Karajarri wankayi  Kartujarra wanka  Kurrama wanka          Martu Wangka wanka
Mudburra wanka     Ngarla wanka      Ngarluma wanka         Nyangumarta wanka
Panyjima wanka     Payungu wanka     Pintupi-Luritja wanka  Pitjantjatjara wanka
Thalanyji wanka    Wakaya wanku      Warnman wanka          Warumungu wankka
```

w·n·k|SPEAK · ternary · 14 languages · enrichment 41× · emergent w·n·2

```
Guyani wangka   Jiwarli wangka       Kariyarra wangka  Kurrama wangka
Ngarla wangka   Nhanta wangka        Nyamal wangka     Nyungar wangka
Payungu wangka  Thalanyji wangka     Warlpiri wangka   Warriyangga wangka
Wirangu wangka  Yindjibarndi wangka
```

j·n|WALK · binary · 11 languages · enrichment 26× · emergent 1·n

```
Batyala yenna      Bunganditj yan  Dhurga yana-   Gamilaraay yenni
Gundungurra yánna  Guwa yana       Mirniny yango  Pallanganmiddang yan-
Wathawurrung yan-  Wulguru yana    Yugambeh yana
```

p·t|bite · binary · 11 languages · enrichment 41× · emergent 3·1

```
Kaanju patha  Kugu Nganhcara patha  Mayi-Kulan patha  Mirniny patha-
Ngawun patha  Pakanh patha          Parnkala paiata   Pitjantjatjara paṯa
Umpila patha  Wik Mungkan path      Wirangu patha
```

m·n|fly_V · binary · 11 languages · enrichment 9×

```
Alyawarr ameng  Antekerrepenhe ameng  Awabakal meini          Kartujarra muungu
Karuwali mongi  Kukatja munga         Ngaanyatjarra muungu    Pintupi-Luritja muungu
Wakaya imungu   Wangkatja muungu      Western Arrarnta manga
```

n·n|see · binary · 11 languages · enrichment 8×

```
Bilinarra nya-ng     Kalkatungu nhanyi  Martu Wangka nyangu  Mathi-Mathi nhanga
Ngadjumaya nangu-    Ngaiawang n@a@an   Ngarinyman nyanya    Nyungar nhungaa
Pitjantjatjara ṉaŋa  Warlpiri nyanyi    Yagara nan
```

n·m|see · binary · 11 languages · enrichment 20×

```
Dhangu nha:ma  Dhayyi nhäm        Dhuwal nhäma       Djambarrpuyngu nha:ma
Djapu nhäma    Mayi-Kulan nhama   Mayi-Kutuna nhama  Mayi-Yapi nhama
Ngawun nhama   Ritharrngu nhaama  Yan-nhangu nhäma
```

p·c|bite · binary · 10 languages · enrichment 60× · emergent 3·1

```
Dyirbal paca      Kartujarra paca  Kungkari paca     Northern Mangarla paca
Nyangumarta paci  Pirriya paca:    Pitta-Pitta paca  Wangkayutyuru paja-
Wangkumara paca   Yidiny paca
```

n·r|hear · binary · 10 languages · enrichment 10×

```
Darkinyung ngara    Dharuk ngára        Diyari ngara-      Gumbaynggir ngarra-
Malyangapa ngaRa-   Nhirrpi ngaRa       Wangkumara ngara-  Wargamay ŋa:ra
Yandruwandha ngara  Yorta Yorta ngarr-
```

w·n·k|say · ternary · 10 languages · enrichment 33× · emergent w·n·2

```
Kariyarra wangka   Kartujarra wangka  Ngarla wangka          Ngarluma wangka
Nhanta wangka      Panyjima wangka    Pitjantjatjara wangka  Wajarri wangka
Warlmanpa wa-ngka  Wirangu wangka
```

Objects and culture

m·r|SPEARTHROWER, w·n·l|BOOMERANG, w·n|DIGGING STICK. Three pieces of material culture with reaches of eleven to fifteen languages.

m·r appears here and also with ”hand” (69 languages) and ”black” (12). It is the catalogue’s most productive code, and its recurrence across three apparently unrelated concepts is exactly what the reverse index asks one to watch.

w·n|DIGGING STICK · binary · 19 languages · enrichment 31×

```
Arabana wana       Diyari wana        Jiwarli wana       Kariyarra wana
Kukatja wana       Martu Wangka wana  Nhanta wana        Nyamal warnu
Nyangumarta warnu  Nyungar wan        Panyjima wanu      Pitjantjatjara wana
Tharrgari wana     Wajarri wana       Wangkangurru wana  Wangkatja wana
```

m·r|SPEARTHROWER · binary · 15 languages · enrichment 13×

```
Alyawarr amerr     Arabana amira    Jiwarli mirru           Kurrama mirru
Ngarluma meero     Nyungar meero    Pintupi-Luritja mirru   Pitjantjatjara miru
Thalanyji mirru    Tharrgari mirru  Wajarri mirru           Wardandi miera
Warriyangga mirru  Watjuk meero     Western Arrarnta merra
```

m·j|house · binary · 15 languages · enrichment 49× · emergent m·1

```
Jaru maya          Jiwarli maya      Karajarri maya     Kariyarra maya
Martu Wangka maya  Ngarla maya       Ngarluma maja      Nhanta maya
Nyangumarta maya   Panyjima maya     Payungu maya       Thalanyji maya
Tharrgari maya     Warriyangga maya  Yindjibarndi maya
```

w·n·l|BOOMERANG · ternary · 14 languages · enrichment 93×

Badjiri wangul	Biri wangal	Dharumbal wangal	Djabugay wangal
Gangulu wangal	Gudjal wangal	Gunggari wangal	Gunya wangal
Kuku Yalanji wangal	Margany wangal	Wargamay waŋal	Warungu wangal
Wulguru waangal	Yiningay wangala		

w·m·r|SPEARTHROWER · ternary · 14 languages · enrichment 174×

Adnyamathanha wamara	Biri wumara	Bularnu wamara	Darkinyung wamar
Dharuk womra	Gangulu wumara	Gumbaynggir wamaarr	Pitta-Pitta wamirra
Wangkumara wamara	Warumungu wamiri	Wiradjuri womar	Yalarnnga wumara
Yorta Yorta womerra	Yuwaalaraay wamara		

k·l|BOOMERANG · binary · 11 languages · enrichment 22× · emergent 2·l

Gurindji karli	Jiwarliny karli	Kartujarra karli	Kukatja karli
Martu Wangka karli	Ngaanyatjarra karli	Nyamal karli	Pitjantjatjara kali
Warlmanpa karli	Warlpiri karli	Watjuk kaile	

Qualities

j·r|sharp shares a skeleton with j·r|tooth (21 languages) and j·r|THOU (18). Three concepts, the first two with a semantic neighbourhood the reader may judge; the third without.

This is the fourth time this warning appears in the appendix. In a family where the catalogue produces 2,092 codes over 360 concepts, a skeleton recurring across several concepts is constant, and one must carry that into every block.

m·r|black · binary · 12 languages · enrichment 10×

Adnyamathanha maRu	Diyari maru	Garlali murru	Jiwarliny maru
Ngaanyatjarra maru	Ngamini maRu	Northern Mangarla maru	Pintupi-Luritja maru
Pitjantjatjara maɽu	Wangkatja maru	Yalarnnga maru	Yarluyandi maru

g·l|ANGRY · binary · 12 languages · enrichment 48× · emergent 2·l

Biri guli	Bunganditj guli	Dharumbal gului	Djabugay guli
Durubul gula	Gangulu gul	Guugu-Yimidhirr guli	Jaru goli
Mathi-Mathi guli	Ngarinyman guli	Warungu guli	Yugambeh gula:

p·k|rotten · binary · 11 languages · enrichment 70× · emergent 3·2

Gurindji puka	Mirniny puka	Muruwari puka	Ngarluma puka
Nyamal puka	Nyangumarta puka	Paakantyi puka	Panyjima puka
Warlmanpa pukka	Warlpiri puka	Warumungu pukka	

j·r|sharp · binary · 11 languages · enrichment 32× · emergent 1·r

Karajarri yoṛa	Kariyarra yiri	Ngaanyatjarra yiri	Nyamal yiri
Nyangumarta yiri_	Panyjima yiri	Payungu yiri	Southern Walmajar jiri
Tharrgari yiri	Warlpiri yiri	Warriyangga yiri	

m·n|good · binary · 10 languages · enrichment 8×

Diyari manyu	Guwa manyo	Kaanju mini	Kaurna mane
Kugu Nganhcara mini	Kuuku-Ya’u mini	Mabuiag mina	Pakanh minni
Umpila mini	Yindjilandji 'mʊngʊ		