The Permutation Dimension: A Mathematics of Metathesis

Modelling reordering as an operator on the consonantal skeleton (protoform-agnostic)

Alejandro Toledo Martínez — Independent researcher (ORCID: 0009-0000-1277-9697)

Abstract ‹stub›

The pilot's operator $\Delta(a, b)$ lives in feature space $\mathbb{F}_2^n$ *at a fixed slot*; by construction it cannot represent a reordering of segments (the monotonic aligner mis-scores metathesis as substitutions and gaps). This paper adds the orthogonal dimension: metathesis as a permutation of the consonantal skeleton's positions. It (i) fixes a mathematical object for the metathesis event, (ii) builds an alignment-free detector, (iii) measures the intra-language algebra of a system's reorderings and (iv) compares it across systems. State the headline once results exist. ‹fill›

1. Introduction — the axis the feature operator cannot see

- Recap: the pilot operator is a feature-difference *at aligned slots*; alignment is monotonic (order-preserving), so reordering is invisible and leaks as noise. ‹stub›
- Metathesis is common and famous (Spanish *murciélago* ← *murciégalo*; *cocodrilo* ← *crocodilo*; *palabra* ← *parabola*; *milagro*, *peligro*): from adjacent swaps to long-distance moves. ‹stub›
- Thesis: metathesis is a second, orthogonal dimension of the change operator — order, not features — and it has its own measurable algebra, within a language and across languages. Discover-then-contrast, as always. ‹stub›

2. The object — what mathematical event models a metathesis? *(the central design question)*

A menu of candidate formalisms, from simplest to richest; the paper selects and justifies one (or a layered set).

- (A) Permutation $\pi \in S_k$ of the skeleton's positions — the object; its action on the order is the *event*. Two coderivatives with the same consonant multiset but different order differ by π. Brings for free: composition (group), cycle type, parity/sign, inversion count, and whether a language is restricted to adjacent transpositions (Coxeter generators of S_k). ‹stub›
- (B) The full operator $(\pi, \Delta) \in S_k \times \mathbb{F}_2^n$ — reorder *and* change features together. The pilot's operator is the $\pi = \mathrm{id}$ face; metathesis is $\pi \neq \mathrm{id}$. This is the honest event, since real metathesis usually co-occurs with feature change. ‹stub›
- (C) Rearrangement distance (genome-rearrangement mathematics; Bafna–Pevzner): the event is a sorting operation (transposition / reversal / block-move); the distance between two skeletons is the minimum number of such operations. Handles adjacent swaps (*murciélago*, 1 transposition) and long moves (*cocodrilo*, 1 block-move) on the same scale, and resonates with the genetics analogy the programme already uses. ‹stub›

- (D) Braid / crossing view — if we track *how* segments cross rather than only the endpoint permutation. Likely more than needed at first; noted for completeness. ‹stub›
- Recommendation to argue: core object $\pi \in S_k$; honest operator (π, Δ); magnitude via rearrangement distance. ‹settle›

3. Detecting metathesis without reconstruction (and without a monotonic aligner)

- Alignment-free detector: for each coderivative pair, compare the consonantal *skeletons*; a same multiset, different order pair is a metathesis candidate. No NW, no protoform. ‹stub›
- Measure the reordering: the permutation π relating the two orders (and its rearrangement distance); co-occurring Δ recorded separately. ‹stub›
- Nulls/controls: chance rate of same-multiset/different-order under a segment-shuffling null; the pilot's discipline carried over (concept-permuted D_R, language-level resampling). ‹stub›

3a. Pilot on IE-CoR (iecor) gold cognacy — first numbers

From `paper3/metathesis_iecor.py` over expert IE-CoR cognate sets; consonant skeletons from raw segments (no OAS, no reconstruction); one reflex per language; persisted to `data/db/paper3.db` (`metathesis`, `metathesis_perm`, `metathesis_pair`).

In one line. On clean Indo-European data, metathesis is *rare but real*, and when it happens it is *local* — a swap of neighbours, not an arbitrary shuffle.

Object settled (§2): the pilot uses (A) $\pi \in$ S_k, with the inversion count as the size of the rearrangement. It is the cheapest honest object; the fuller operator (π, Δ) waits for §3b.

How many, and how sure. We examined 2,424 gold cognate sets — each with at least two languages and a skeleton of at least two consonants — and the 43,283 cross-language reflex-pairs they contain.

Of those pairs, 6,361 (14.7 %) keep the *same* consonant multiset. Only 29 of them reorder it. Those 29 are the metathesis candidates, spread over 17 sets (0.7 % of the total).

Two cautions keep this honest. First, metathesis really is rare: 99.5 % of the skeleton-preserving pairs keep the same order. Second, the detector is conservative — it demands an exact multiset match on raw segments, so a single transcription or diacritic mismatch throws a pair out. So 29 is a floor, not a ceiling.

Is 29 more than chance? Yes. Among length-matched *non-cognate* pairs, the same-multiset-different-order rate is only 0.030 %; our candidates run about 2× above it. They are cognate-driven, not coincidence.

The headline: the permutation dimension is LOCAL. We classified each candidate's π against a uniform-S_k baseline — the rate you would see if reordering were random (a single adjacent swap then has probability $(k-1)/(k!-1)$):

skeleton length k	candidates	adjacent transpositions	observed adj %	uniform-S_k adj %
2	8	8	100 %	100 %
3	12	7	58 %	40 %
4	8	2	25 %	13 %
5	1	0	0 %	3.4 %

At k = 3 and k = 4 adjacent (neighbour) swaps are over-represented 1.5–2× vs uniform — metathesis concentrates on *local* transpositions rather than arbitrary reorderings, the first quantitative statement of the permutation dimension. (Numbers are small — this is directional evidence, to be firmed up at scale and on further families.)

The specular TR↔RT case. Among length-2 candidates, 5/8 (62 %) are obstruent↔liquid flips (e.g. `s l ↔ l s`) — the exact consonantal-skeleton pattern this axis was built to catch.

Cross-language recurrence (§5, first glimpse). The *same* skeleton swap recurs across different language pairs — `t r v ↔ t v r`, `s l ↔ l s`, `m r ↔ r m`, `s k n ↔ k n s`, `s p r s ↔ s r p s`. *Caveat:* pair-level counts can inflate a single set observed across several languages; the write-up must recount recurrence at the independent-innovation (set/branch) level.

What stays open (§7). Direction (which order is older) is undecided reconstruction-free; the detector is a lower bound; k-per-length counts are small. Next: relax the detector (allow one co-occurring Δ, i.e. the full (π, Δ)), scale n and add a second family, and split recurrence by branch/area.

3b. The full operator (π, Δ): tested, and a method boundary

The pilot detects *pure* reordering ($\Delta = 0$). But real metathesis often comes with a sound change too, so we relaxed the detector to the honest operator (π, Δ) — reorder *and* allow a substitution.

The rule is simple: two same-length skeletons are a candidate if some permutation π, followed by at most SUBCAP substitutions, makes them match — with π not the identity, and with the reorder *paying for itself* (it must fix at least two mismatches, so a plain substitution cannot pose as a metathesis). At SUBCAP = 0 this reproduces the pilot exactly (29).

The real question is whether relaxing adds signal or just noise. We tested it against the same anagram-by-chance null (`paper3/metathesis_v2.py`; `data/db/paper3.db`, tables `metathesis_v2` …):

operator	candidates	signal ratio (observed / chance)
$\Delta = 0$, any π (the pilot)	29	1.94×
$\Delta = 0$, local π only	17	1.23×
$\Delta \leq 1$, any π	173	1.16×
$\Delta \leq 1$, local π	44	0.96×

The result: a clean negative. Every relaxation *lowers* the signal-to-noise.

Allowing even one substitution collapses the ratio, from 1.94× down to 1.16×. The reason is simple: a permitted substitution is exactly what lets *random* pairs match too, so the null rises as fast as the raw count. Restricting the reorder to local swaps does not save it either (0.96×) — and the 96–100 % "adjacent" figure there is *tautological*, since we imposed locality, so it is not evidence.

The extra candidates are plausible one by one. The swap $s\,r\,\chi \leftrightarrow \theta\,\chi\,r$ with $s \to \theta$ (a Grimm-plus-metathesis) even recurs across three language pairs. But taken together they are indistinguishable from chance for a bag-of-skeletons method.

What it means. The real, above-chance signal lives in the pure reorder ($\Delta = 0$) — and it is strongest when long-distance reorders are included (1.94× vs 1.23× for local-only), which vindicates the pilot's choice.

The full (π, Δ) operator is the right *object*, but it cannot be detected cleanly without alignment. To separate a reordering from a co-occurring substitution you have to line the two up, and that needs joint cognate-set alignment (MSA), not pairwise skeleton comparison. That is the honest next tool — and a boundary result in its own right: *reconstruction-free metathesis detection is reliable for pure transposition, and degrades to chance the moment substitution is admitted.*

3c. Why does metathesis happen? A mathematical–cryptological reading

Historical linguistics documents *when* metathesis spreads but says little about *why* it arises; the causal accounts live in other disciplines — neural parallel planning (Fromkin, Garrett, Dell), phonotactic optimisation (Prince & Smolensky), developmental phonology (Stoel-Gammon), aphasia (Blumstein, Levelt), rhythm (Goswami), self-organisation (Bybee, Croft, Kirby), and, as a distinct register, psychoanalytic displacement (Freud, Lacan). Rather than choose one, we read the whole ladder through our object, the transposition operator $\pi \in S_k$, which turns several of these "why"s into statements about the *same* mathematics and yields testable predictions. (Disciplines enter as support; the theory is ours.)

1. Why *local*? — minimal word length in the Coxeter generators. Speech is planned *in parallel*: co-active, nearby segments exchange (Dell). In S_k that is exactly an adjacent transposition — a single Coxeter generator. "The brain swaps co-active neighbours" and "metathesis is a length-1 word in the adjacent-transposition generators" are the *same statement*, and §3a's measured over-representation of adjacent swaps (low inversion count) is the algebraic fingerprint of that neural fact: metathesis lives near the identity on the Cayley graph of S_k.

2. Why *tolerated at all*? — it is the information-preserving sound change. A transposition cipher preserves the multiset of symbols (which is why frequency analysis alone cannot break it — you must anagram). Phonologically the consonant *multiset* is the root's identity (the Semitic extreme, framing §3.6). So metathesis re-encodes order while conserving content: its equivocation on the root is zero, $H(\text{multiset} \mid \text{reflex}) = 0$ — unlike a merger, which destroys information irreversibly. A language can *afford* metathesis precisely because it does not reduce recognisability. That is a crisp reason, in our own information-theoretic taxonomy, why an otherwise "irregular" change is stable enough to enter the key.

3. Why *toward* a particular order? — sorting to sonority. OT says rising-sonority onsets (CR/PL/TR) are preferred; in our terms that is a target permutation, and metathesis is a *sorting* move toward it — the inversion count (§3a) is the number of sonority-sequencing violations a swap repairs, and our TR↔RT result (62 % of length-2 candidates) is exactly the sonority-swap axis. Tested (`paper3/metathesis_sonority.py`; `data/db/paper3.db`, `sonority_test`): with a rising-sonority proxy over the (vowelless) skeleton, metathesis orders are not more sonority-harmonic than random permutations of the same multiset (ΔC ≈ -0.02, sign p ≈ 0.6) nor than a non-metathesis attested control ($\Delta\Delta \approx 0$) — no evidence of sonority *optimisation* in aggregate. What *is* visible: metathesis toggles the sonority axis — 76 % of pairs differ in fall-count, 7/8 length-2 pairs straddle rise↔fall (TR↔RT) — so the operation lives *on* the sonority dimension without demonstrably descending it. Two honest limits: the *directional* "restoration" claim (marked $\rightarrow$ unmarked) is untestable reconstruction-free (both orders are attested; we lack the arrow of time), and a proper SSP test needs syllable onsets — the vowels the skeleton drops (the recurring corpus wall). A vowel-aware, direction-controlled test is the honest next step.

4. Why *rhythmic*? — a keystream-conditioned transposition. The rhythm accounts (Goswami; our endorhythm) make the temporal template an oscillatory signal; a misaligned skeleton is re-ordered to fit. Cryptologically this is a transposition whose π is conditioned on a prosodic keystream — the transposition-layer analogue of the polyalphabetic/keystream conditioning of framing §3.2 (Verner). And like Verner's accent, the rhythmic conditioning may be erased in the reflexes, which is why a corpus without prosody cannot recover it. Prediction: metathesis rate should correlate with a language's prosodic/rhythmic properties — the empirical handle the endorhythm hypothesis needs.

5. Why does a slip become a law? — micro-noise renormalised into the key. Individual speech-exchange errors are a stochastic channel applying a small π with low probability — transient

noise. A *historical* metathesis is that noise entering the shared key: population dynamics (replicator–mutator, framing pillar 5) fix the variant, and it earns its place only when it pays its description-length cost (MDL) in the community's code. This is a *renormalisation* — micro-scale articulatory noise coarse-grains into a macro-scale rule — and it explains §3a's rarity (0.7 % of sets) and §3b's negative: most transposition noise never pays for itself, and admitting a substitution makes the operation indistinguishable from that noise.

A distinct register (declared, not asserted). Psychoanalysis (Freud's *Verschiebung*, Lacan's sliding signifier) reads reordering as *displacement* on the chain of signifiers — metathesis as displacement at the phoneme level. We keep this as the programme's third register (constructed meaning, not found fact): a hermeneutic reading of the same π, honest about being an interpretation rather than a mechanism.

The reframed question. "Why does metathesis exist?" becomes: *what makes a content-preserving transposition a recurrent, low-cost, sometimes evolutionarily-stable move on the Cayley graph of the skeleton?* — answered, in our terms, by (1) nearness to the identity, (2) zero equivocation on the root, (3) descent toward a sonority target, (4) an optional rhythmic keystream, and (5) the MDL/population threshold it must cross to enter the key. Each clause is a measurement, not a metaphor.

4. Results I — the intra-language algebra of metathesis

- For each language/system: the set of π's it uses — is it a restricted sub-structure of S_k? Only adjacent transpositions? Particular cycle types? Do they compose (a sub-monoid)? ‹fill›
- Distribution of rearrangement distance; which positions/segments move most. ‹fill›

5. Results II — metathesis across languages

- Do the same π patterns recur across related/contact systems? Is there specular inversion (the TR↔RT case the consonantal-skeleton view was built to catch)? ‹fill›
- Does the *permutation* signal carry system/areal information the feature-change signal does not (contrast with the pilot's distributions)? ‹fill›

6. Discussion ‹stub›

The order dimension beside the feature dimension: the operator algebra as $\mathbb{F}_2^n \times S_k$; connection to genome rearrangement; what metathesis adds to "where genealogy lives" (pilot Annex A).

7. Limitations and scope ‹stub›

Skeleton extraction and multiset equality are approximations; partial metathesis co-occurring with deletion is hard; direction of the reordering (which order is earlier) needs the same reconstruction-free caution as Paper 2.

8. Reproducibility ‹stub›

New `make` targets (metathesis detector, permutation/rearrangement measures); durable results in `data/results/`.

References ‹stub›

Reuse the pilot's; add: symmetric-group / Coxeter basics; genome-rearrangement (Bafna & Pevzner; Hannenhalli & Pevzner); metathesis typology (e.g. Blevins & Garrett) for the *contrast* set.