Hypomnemata Glossopoetica

Wm S. Annis

January 22, 2022

1. Phonology

Normally illegal clusters may occur in particular grammatical contexts, and thus look common (*cf.* Latin -**st** in 3SG copula **est**).

Hierarchy of codas:[1] n < m, ɳ, ŋ < ɳ « l, ɾ < r, ɻ, ʁ < ɭ, ɽ « t < k, p < s, z, c, q, ʃ < b, d, g, x h « w, j. There is a slight place hierarchy: alveolar < velar < retroflex or tap. Classes percolate, such that in complex codas, if Nasals, Resonants and Stops are permitted you usually expect N, R, S, NR, NS and RS as coda sequences. Other orders are possible, but the above rule is common-ish.

Hierarchy of clusters (S = sonorant, O = obstruent), word initial: OS < OO < SS < SO; word final: SO < OO < SS < OS. Onset clusters tend to avoid identical places of articulation, which leads to avoidance of things like **tl, dl, bw,* etc., in a good number of languages. /j/ is lightly disfavored as C2 after dentals, alveolars and palatals; /j/ and palatals are in general disfavored before front vowels.

Languages with *sC-* clusters often have codas. S+STOP < S+FRIC / S+NASAL < S+LAT < S+RHOT (the fricative and nasal are trickier to order).

Even if a particular C is a permitted coda, its allowed environment may be quite restricted. Potential constraints: forbidden before homorganic stop; or homorganic nasal; geminates forbidden. Solutions: delete with compensatory vowel lengthening; debuccalize (become fricative, glottal stop, delete without compensation); nasal deletion with nasalized vowel remaining; tone wackiness.

Lower vowels are preferred as syllabic nuclei; high vowels are more prone to syncope (either midword or finally). Content words less likely to elide.

Sonority hierarchy:

> low vowels > mid vowels (except /ə/) > high vowels (except /ɨ/) >
> /ə/ > /ɨ/ > glides > laterals >
> flaps > trills > nasals > /h/ >
> voiced fricatives >
> voiced stops and affricates, voiceless fricatives >
> voiceless fricatives, voiced stops and affricates >
> voiceless stops and affricates

Vowel devoicing starts with lowest sonority (/i/, /u/).

Hiatus resolution: 1) elide, contract, add glide /j/ or /w/, with the glide sharing the same frontness or roundness of V1; 2) insert /h/ or /ʔ/; 3) insert coronal consonant, usually /t/ or some /r/.

[1] Some attested single-C coda inventories: {n ŋ}, {n ŋ t k}, {n ŋ m p t k}, {n m l r}, {n m w j}, {n ŋ m l r j}, {n m r d}, {d l s x}, {m b k l z r}, {n l w j t k}, {n ŋ m l p t k}, {w n m r k t v ʃ ʒ}.

Metathesis (where P = /p, b, m/, etc.). Obstruent + resonant can switch. PK > KP occasionally, but not the other way. T{PK} > {PK}T reasonably common.

Many things can happen to a consonant following a nasal: NT > ND; NS, NZ > NTS, NDZ; NT > NT^{h}; ND > NN; but also: ND > NT; NT^{h} > NT; NN > ND.

Consonant strength hierarchy (in some African languages, generalized): continuants, semivowels and ʔ > voiced stops > nasals > voiced and nasal geminates, voiceless stops > voiceless geminates > nasal clusters.[2]

Some suffixes may occur only in pausa, or are different there.

Syllable weight: closed syllables may be heavy or light depending on the coda consonant, with sonorants more often making weight and stops not (in general CVV > CVC > CV). Also, low vowels are generally heavier than high, and central vowels less heavy than non-central. You can get word-length dependent affix allomorphs of the same number of syllables, but with heavier vowels in one and lighter in the other.

Consider palatalization, labialiazation, nasalization, etc., as phenomena which apply to a *syllable* rather than a particular segment. This can generate substantially differing outcomes in historical processes (*v.* Chadic family), especially as differing branches apply the effects more closely or more distantly.

For determining stress accent, Malagasy has some light final syllables. These are the remnants of final consonants that got vowels tacked on after a switch to a dominant CV structure (and the other final consonants deleted).

1.1. ATR Harmony. Some systems are dominated by the feature of the word root, causing harmony both to the left and the right. Other systems are based on syllable position, *e.g.,* a system where -ATR vowels can follow, but not precede, +ATR vowels. An originally larger vowel inventory may reduce, giving a situation where one apparent vowel can trigger +ATR in some words and -ATR in others. Different morphological systems may take different rules within one language.

May apply to all vowels, or just to -ATR /ɪ ʊ/ (or /ɛ ɔ/).

/a/ may be transparent, or may block. Or may be -ATR, with /ə/ the +ATR version.

Common systems: 9V /iɪeɛaɔoʊu/; 7V(M) /ieɛa(ə)ɔou/ or 7V(H) /iɪɛa(ə)ɔʊu/. -ATR usually dominant in 7V(M), +ATR usually dominant in 7V(H).

1.2. Tone. In many Dogon languages, postnominal adjectives and demonstratives (but not numbers or other quantifiers other than one) cause all tones to lower in the preceding NP. If multiple adjectives follow, all are lowered except the last. The head noun of an internally headed RC is also so lowered.

1.3. Reduplication. The original word-initial consonant of a reduplicated root may undergo lenition, or fortition. Malagasy **fantatra** > **fanta̲pantatra**.

The **inflectional** uses include: number (for nouns; for verbs marking actor number or event plurality; in noun compounds either or both elements may reduplicate, sometimes in free variation); very rarely to encode possession (either of 3rd person, or 1st and 2nd, usually with partial reduplication); frequent in verbs for frequentative, habitual or progressive aspect, but also imperfective, inchoative and perfect(ive); Margi participles are formed by partial or full RED (seen in a few other languages).

RED may be used to repair a word shape that has taken some affix.

The **derivational** uses include: ordinal numbers from cardinal; distribution ("3 each"); valency reduction in verbs (including antipassive), as well as reciprocals (and mutuality in other word classes, "face to face"); diminutive (and both endearment and contempt); associatives ("someone with X"); similarity; sort/kind of N; disposition ("someone prone to X, likely to X"); common in indefinites ("who" > "whoever"); in verbs, a lack of control, disorder, carelessness, pretense, attempt; incrementality ("gradually, little by little, one by one"); spread-out or scattered ("here and there, looking around"); non-uniformity ("zig-zag, now and then, in several colors, hodge-podge").

[2]Bilin has consonant "ablaut" in the plural, where the penultimate or ultimate consonant (occasionally both) altnernate in the plural: *b-f, d-t, d-s, d-ʃ, r-t, l-t, r-l, dʒ-ʃ, g-k, x-k, gʷ-kʷ, xʷ-kʷ, x-kʼ, xʷ-kʼʷ, w-kʷ.*

Not infrequently used to name insects and birds, and the word for "baby."

Noun to verb ("to wear an X"), adjective or adverb; verb to noun (agent noun, action noun, instrument), adjective or adverb; adjective to adverb.

For adjectives and other vocabulary of quality: diminutive attentuative, augmentative, intensification.

The sense of "collectivity" might be used where we would expect a single term ("broom, flight of stairs," etc).

Some affixes may simply require the root take RED, with Makah having several types of RED and affixes selecting one.

2. Word Classes

	Reference	**Modification**	**Predication**
Object	referring, argument phrase (**noun**)	nominal attributive phrase	predicate nominal
Property	deadjectival nominal	attributive phrase (**adjective**)	predicate adjectival
Action	complement clause	relative clause	clause (**verb**)

Universal One. A lexical class used in a nonprototypical propositional act function[3] will be coded with at least as many morphemes as in its prototypical function, as in *bright* > *bright-**ness*** (prototypical functions are marked in bold in the chart).

Universal Two. A lexical class used in a nonprototypical propositional act function will also have no more grammatical behavioral potential than in its prototypical function. For example, a predicate nominal will not have more verbal inflectional possibilities than a full verb.

Individual languages will break up that 9x9 grid differently in terms of constructions, and where individual words end up.

3. The Noun

Nouns are the most frequently borrowed word class.

3.1. Possession.

"Possession" can cover a wide range of relationships:

- Ownership (or temporary possession)
- Whole-part relationship (body part, part of an object)
- Kinship
- Attribute of a person, animal or thing ("Bob's temper")
- Orientation or location ("the front of the house"; useful when body part terms are used for location to use different marking)
- Association ("my teacher," but also dwellings, house to homeland, and personal clothing and goods)
- Nominalization

Most languages do not have the wide range of possessive uses found in English or Greek in the same construction. The first three in the list above are most central.

Marking for inalienable possession is generally smaller (fewer syllables, no classifier, simpler construction) than the marking for alienable.

Genitive marked on R, pertensive marked on D. Affixed possession markers may also induce pertensive marking. Genitive marking frequently has other functions, pertensive rarely.

[3]Reference, modification or predication.

Different systems of marking may occur in two or even three groups depending on possessor: pronoun, proper noun, kin term, common human noun, common animate noun, common inanimate noun (with contiguous constructions between groups).

Splits across whole-part D: external body parts, internal body parts, genitalia, body fluids, parts of animals, parts of plants, parts of artifacts.

3.2. Classifiers. Number classifiers are most common with "one" and "two," possibly obligatory only with these; rarely with different classifier forms for higher numbers. More classifier types may be available with lower numbers.

With stative verbs, the classic Dixon set are most likely to take classifiers; or postural verbs. With transitive verbs, more likely on high-agency, high-patientness verbs (handling, cooking, killing and the like). With any verb class, optional classifiers may mark salience or "completeness" of the classified feature.

When not used with their expected word class, classifiers may be derivational.

3.2.1. Gender / Noun Class. Variable gender assignment may code size and/or shape and/or posture (upright vs. horizontal). Phonological gender assignment (initial, final, or both, are possible determinates) typically restricted to nonanimates. Function of referent may change gender (water as drink vs. water as part of the landscape).

Agreement: clefting may interfere or inhibit agreement; mixed class resolution may default to the least marked form, but the behavior for animate vs. inanimate may pick different forms, and coordination of mixed animacy may simply be avoided altogether (a different construction, or repeating the phrase with different subjects, etc.); phonetics may inhibit some kinds of agreement. Semantic agreement override: attributive < predicate < relative < personal pronoun.

Different word classes may pick different genders for the least marked, catch-all form.

It is rare but possible to have two noun class systems operating in a single language, with different agreement rules operating on different word classes. In particular, "pronominal" agreement for pronouns and verb agreement[4] and "nominal" agreement for adjectives (and sometimes numbers). Pronominal agreement in general is focused on pronouns, is a smaller system, aligns with animacy, sex or humanness, and has a fairly transparent semantic basis. Nominal agreement systems tend to be larger, focus on animacy, sex, shape and size, and the semantics may be much less clear. Demonstratives align with either pattern, from language to language. Where the two systems have semantic overlap, the markers may be quite different, or cover different spaces.

Hierarchy (tendency) for site of marking: pronouns > verbs > demonstratives > adjectives > numerals. Also possible: verbs > adjectives > pronouns.

3.2.2. Noun Classifiers. Similar to gender, and historically may lead to it, but classifiers are different. The semantics are generally clearer; some languages may allow cooccurance of classifiers (SPECIFIC GENERAL, as in "HUMAN MAN boy"); single N may take different classifiers to determine precise meaning; may be used anaphorically (esp. across clauses of a single sentence), possibly with other syntactic functions (relative clauses). Attend to coreferentiality.

May be small (2–3) or larger (20, or more).

"Social status" may be encoded in human classifiers ("initiate, known," etc.). Some classifiers encode inherent nature (person, bug, tree), some function (edible, drinkable, movable, etc.). Multiple classifiers will usually encode one of each.

Classifier on N may be omitted once class established, or obligatory always.

May be separate Q for "unknown class" vs. "known class but no more."

Noun classifier systems often related to number classifier systems, but they may be different. Number classifiers are much less likely to be optional. Both systems may occur in a single NP.

[4]Paumarí has m./f. for demonstratives, but a separate *ka*/non-*ka* agreement system on the verbs, showing up as a prefix.

3.2.3. Number Classifiers. Sometimes there is a catch-all classifier, sometimes many nouns don't take any classifier.

May not go beyond ten, or are dropped with 10s units.

Classifiers for humans/animates may have different forms with different numbers; with low numbers the N + CL form can be suppletive.

Affixed classifiers more likely to code animacy, and to be required; independent classifiers code shape, consistency, etc., and may be more or less optional, with different speakers at different competencies. Rarely, affixed and independent classifiers may appear together.

A single noun may take different classifiers to focus salient characteristic.

3.2.4. Possessive Classifiers. Can code: shape, consistence, animacy of possessum; relationship of possessor to possessum (as in Oceanic indirect possession); or, very rarely, class of possessor.

There may just be a bunch of words for "of" which match class (possibly highly suppletive).

3.2.5. Verbal Classifiers. Many originate from noun incorporation, possibly competing with it. However, some verbal classifiers may be clearly related to numeral or other classifiers in the language. Another origin is from verbs, via grammaticalized serial verb constructions.

Existential verbs may distinguish animate from inanimate; a few languages elaborate existentials quite a lot (in a container, movable vs. immobile, non-human animal, etc.).

3.3. Case. SOV languages more likely to have case marking than SVO.

3.3.1. Formal Marking.

A	O	E	peripheral	
w	x	y	z	Latin (very common)
w	x	y	y	Jarawara
w	x	x	y	Kinyarwanda
w	x	x	x	Creek

In Creek there is **-t** on subjects and **-n** on nonsubjects.

3.3.2. Accusative. This may be sensitive to the animacy of the object, with high animacy objects more likely to take an overt accusative (so-called "differential object marking"). DOM may be sensitive to the relative animacy of A and O. Or it may vary with definiteness. Even in languages with DOM, O marking may be required in all ditransitive constructions.

In many languages, ACC may be used on adverbs of distance and duration. It is also regular on nominalized complement clauses ("I know he-is-here-ACC").

Sometimes motion verbs take a goal in the ACC.

Historically, ACC may evolve from DAT, more rearely from INST.

3.3.3. Allative. A huge range of possibilities here. Occurring more often: true allative ("to, towards, reaching for"), purpose ("use it for that, in order to"), conceptual ("think about, occur to"), recipient (dative), timepoint ("at TIME"), addressee ("talk to me"), perceptual ("look at, listen to"), reason ("because of, ran from fear"). Some less frequent possibilities include: temporal boundary ("by/until TIME"), benefactive, possessive, porportion or rate ("3 out of 4, 3 at a time"), equivalence ("as, in exchange for"), subordinator ("although, when, while"), emotional target ("hard for, be angry at").

3.3.4. Ablative. An ablative may mark not only the source but also the path of motion (perlative).

3.3.5. Locative. Very rarely (Páez) a language may have several locatives which mark posture (standing, lying, hanging, leaning, etc.).

3.3.6. DPM. Differental Place Marking is possible, with the hierarchy: human > common inanimate > place. Shorter locative marking becomes more likely as you go down the hierarchy (such as Latin using bare cases without prepositions for the names of towns and small islands, and a few things like *domus*). Names of cities seem to come in for this a bit more than other things. Further, there might be special additional morphology when animates take locative marking, such as the Basque *-ga(n)-* extension for LOC, ABL, and ALL. Different, often more complex, adpositional constructions might be required for animate place targets.

3.3.7. Vocative. Vocative expressions often have characteristic vowel lengthening, stress shifts, or tone changes. In support of such pitch alterations, overt vocative morphology will usually have fewer consonants than other sorts of case marking. Mid vowels are weakly preferred to high.

Swedish has both long and short forms of "father" and "mother," *fader:far, moder:mor*. The long forms cannot be used as vocatives (except *fader* in some archaic relgious forms).

4. The Pronoun

There may be separate forms for "PRN alone" and/or "PRN also" unrelated to similar expressions for nouns.

Languages with 1/2 systems are not common, but not rare either.

Sg/du/paucal/pl is far more common that sg/du/trial/pl.

It's not especially common for sg and pl forms to be related (as in Chinese).

Many languages:

	Singular	Plural
1ST	1	12, 13 (and 11)
2ND	2	22, 23
3RD	3	33

However, one might get something like:

1	13 (and 11)
2	12, 22, 23

The basic 1/2/3 sg/pl system may be extended to include a special 12 "me and you" form. This may involve the innovation of a dual throughout the pronoun system. Or it may be just a normal system with an inclusive/exclusive distinction. Or a minimal/augmented system,

Minimal	Non-minimal
1	1 + others
1+2	1+2 + others
2	2 + others
3	3 + others

The non-minimal may include "unit augmented" (one other person, producing a dual in many forms) and just "augmented" (the "plural"). Minimal/augmented are common in Australia, Austronesian, South America, and a few in North America, though the unit augmented/augmented distinction is nearly restricted to Australia.

3sg = 3pl is a common neutralization. 2pl = 3pl (and 2du = 3du) can happen.

Using 2du, 2paucal or 2pl as a mark of respect occurs in Europe, Oceania, Australia.

There is a relationship between the indefinite "someone" and 1non-sg forms (French; Caddo, etc). Usually it's 1pl.inc that does this.

Question-based indefinites may be marked with some other morpheme, often related to words for "be," "want," "perhaps," "or" or "also."

4.1. Demonstratives and Deixis.

- two way
- three way: proximal, distal, remote; proximal, medium, further away or imprecise; proximal, medial (not far, known to 1 and 2), distal
- four way: proximal (to 1), proximal (to 2), medial, distant
- five way: priximal visible, proximal audible, medial, distant, imperceptible
- six way (Godoberi): close to 1, close to 2, that at some distance from 1, that at some distance from 2, that down there, that (aforementioned)

Rarely, pronominal uses may require nominalizer.

Some systems have forms specifically used for anaphora, or only permit particular forms to be used anaphorically (slight preference for distal forms for this?).

In a very few languages, there may be special forms always accompanied by physical gestures (English "yei high").

May have overtones of familiarity, endearment, pejorative, empathy. Like diminutives, only context may make clear a positive vs. negative interpretation. Even proximal deixis can be pejorative.

Demonstratives can be recruited as filler, both within a sentence (a particular form will be preferred), or as "um, ah."

Splits: more distinctions for modifier, with fewer for independent, pronoun-like forms.

Adverbial locatives ("here, there") often match demonstrative distinctions (not infrequently by derivation or clear relation), but may have fewer distinctions. Rarely, particular clause positions may have fewer options (Jarawara, one way distinction clause initial, "here/there," two way distnction clause final).

Manner adverbials ("thus, like this, like that"), again often related to pronouns, may make fewer distinctions.

Verbal demonstratives: "do/be (like) this/that." Can be both anaphoric and cataphoric.

ℜ: Demonstrative agreement with addressee (in addition to usual), with traces of such agreement for ADJ.

ℜ: Ik has demonstratives that locate objects in time, with a five way (non-past, recent past, removed past, remote past, remotest past) distinction that matches the temporal adverbs.

4.2. Indefinites. Cross-linguistically, indefinites derived from interrogatives are about twice as common as those derived from nouns. If a derivational element is used on interrogative indefinites, often related to "be, want, perhaps, or," or "also." Kannada uses "or" clitic to form "some" indefinites, "also" clitic for "any" indefinites. A few languages split function, with interrogatives for "any" indefinites and nouns for "some" indefinites.

If a language has Q-derived indefinites, there may still be some additional marking of a question when used interrogatively, often the same Q particle used in polar questions. Or an interrogative mood. Also Q-word focus constructions (fronting, particle, clause intonation) may distinguish.

Indefinites off the the left of the Haspelmath map are more likely to be appreciative, and those to the right depreciative.

5. The Adjective

Comparatives neither rare nor universal. Can be marked with (1) affix and adpositional phrase; (2) adposition only; (3) coordination ("X is big but Y is small" = "X is bigger than Y").

Breaks along word class distinctions follow the adjective class: full NAV (all of noun, adjective and verb in the lexicon); N[AV] (i.e., property-concept words go with verbs); [NA]V; and [NAV] (all three classes conflated).

Resultative adjectives are most likely in languages in which there are complex verbal predicates (serial verbs and particle verbs). "I hammered it flat" slots into complex verbal constructions naturally. French, for example, simply doesn't have these.

6. Numerals

Bases found in human languages: 2, 3, 4, 5 (hand), 6, 8 (like 4, counting the spaces between digits), 12, 20, 60. Hybrids: 5+20(+80), 2+5, 10+20. Base 10 is by far the most common. Lower bases are less common, and will generally only go up to a few powers, with a base 4 system, for example, only reaching 8, 16, or 32.

The higher power units may not be usable alone. That is, unlike English "ten," some languages may require it be "one ten" in the same way 20 is "two tens."

The words for 2 and 2nd often convey sense of "another" ("secondary").

Overcounting: some systems overshoot, then work backwards, so that "seven thirty" is 27 (old Turkic). Subtraction: 8 = two from ten; 31 = one plus 10 from 2*20.

A very few languages (PNG) uses different bases for counting different things (as in Bukiyip, which uses 3 and 4).

7. The Verb

"Labile transitivity" is very common. Be clear on S = A *vs.* S = O for intransitive constructions. Some languages may skew in favor of one or the other, but others do not. English "cook," with transitive and both intransitives ("the chef cooks; the chef cooks stew; the stew is cooking") is unusual.

Lability is a dodge! If the language has rich mechanisms for changing transitivity, it's less frequent.

Labile verbs tend to cluster semantically, 1) destruction and strong property change, "break, boil, freeze, dry, go out, melt, dissolve, burn, destroy, break, split, kill/die;" 2) motion and spatial configuration, "rock, roll, sink, spread, close, open, connect, rise/raise, stop, fill, turn;" 3) phase, "begin, finish;" 4) non-physical effect, "change, improve, develop;" and 5) verbs with an animate patient, "wake up, learn, gather." A single language may pick several clusters, or only a few items from one.

A transitive verb may be used intransitively in an extended intransitive sense, too.

Verb pairs ("die/kill," "eat/feed") for which the S/O role are primarily animate are more likely to pattern with one basic transitivity for the simplest stem (i.e., "kill" is primary with a derivation for "die"), while those with primarily inanimate S/O role verbs ("boil/boil," "burn/burn", "fall/drop") will also have a primary transitivity (in English, these are often labile or suppletive). In general, though, the basic role for animate verbs will be intransitive, with augmentation for the transitive form. On the other hand, in the large, some few languages are fairly intransitive, some quite strongly transitive.

Stative verbs often mark fewer tense or aspect distinctions (such as Eng. -ing, "*I am knowing"). Or, they may take some additional marking obligatorily (Turkish).

Among other categories, some languages have verbs that mark ease or difficulty. The marking for "ease" on a transitive verb may signal a small O.

In richly marked verbs, these may not be marked: 3RD inanimate sobjects, all 3RD objects, 3RD topical subjects, 3RD absolutives, all 3RD of any kind.

7.1. Affix Order. The fewer person markers there are, the more likely they are to be prefixing. But: object marking on the verb is prefixing with more than chance frequency. Prefixing in general is less common with OV than with VO, except for object marking, with a distinct preference for object prefixing with OV.

If T and A are on the same side of the verb, A is closer; if different, the order is T VERB A.

7.2. Person Marking. Ways to express the subject of a verb, with overt subject vs. no overt subject:

- Marta came-she / came-she (70%)
- Marta came-she / she came-she (very rare outside Germanic sphere)

- Marta came / she came (7%ish)
- Marta came / came-she (14%)
- Marta came / came (7%ish)

Marking of both A and P is somewhat more common than all other options: no agreement, marking A only, marking P only (rare), marking of either (very rare). Agreement may be either by affixes, or by clitics that will tend to attach themselves prosodically to the VP, or in Wackernagel position.

With objects especially, but subjects also sometimes, the verbal marking may only occur as a pro-form replacing a missing overt object; or according to some feature of the arguments, such as animacy, definiteness, high referential prominence, word class (NP vs. PRO); or by tense, aspect, polarity, mood, or even clause type.

Grammaticalization path of O marking is probably different from that of S/A.

7.3. Tense Aspect Mood.

7.3.1. Aspect. Perfective marking has a tense hierarchy: past > future > present (and the present may be reinterpreted, for example as habitual or future).

7.3.2. Telicity. Tense and default aspect interpretations of verbs may be very dependent on telicity.

(PFV) Achievement > Accomplishment > Activity > State (IPFV)

The most natural location for perfective marking is on achievement verbs; the most natural for imperfective is state (though state verbs themselves can be odd). Cross-marking (IPFV on achievement verb) may have other interpretations (such as iterativity in English "I was sneezing").

Im/perfectives may have a default tense interpretation (present/past), which can split anywhere along the hierarchy above. In Even, aorist activities are present, aorist achievements and accomlishments are recent past.

7.3.3. Mood. It's not unusual in a single language for there to be more kinds of constructions for expressing mood and the future than for tense and aspect.

"Irrealis" often associated with future, or becomes future historically.

7.3.3.1. Frustrative. A frustrative may be a separate word, a clitic, or (most common) an inflection. As usual, a range of meanings possible, with no expectation that a single language will use a frustrative for all senses (mostly derived from Amazonian data):

- action was accomplished, but without expected result (often unspecified, but inferrable from context)
- action was accomplished but the speaker finds it irritating or inconvenient (in a few languages, this sense only in negative clauses)
- speaker's unwilling participation
- incompletive, the action was started but not accomplished ("nearly")
- incompletive, the expected action didn't even start
- counterfactual

If cliticized to a noun, often depreciative rather than simply scoping over clause.

May be restrictions on combining with: tense, polarity, mood (such as, never with imperative, or never in future).

7.4. Evidentiality, Mirativity, Egophoricity. All of evidentiality, mirativity, and egophoricity potentially have polysemous marking, so that certain uses of, say, evidential marking might be used for a mirative. May be merged with rest of TAM system, or not.

Evidentials may distinguish: visual evidence, non-visual or sensory, inference, assumption (including reasoning, assumption, or general knowledge), hearsay, and quotative for overt reference to the source. Evidentials may be limited to particular tenses or aspects; least common in the future. Most common two-way distinction is reported vs. everything else. Common three-way: direct, hear-say, inference.

Egophoric marking[5] prototypically marks EGO for first person and NONEGO for everything else in statements, and marks EGO for second person and NONEGO for everything else in questions. Evidential distinctions more likely in NONEGO. Again, past is most likely to mark the most distinctions. Might be separate marking for agent vs. experiencer "subject." Egophoric marking might only occur with volitional predicates ("cook," but not "be sick", Tibetan). Direct egophoric marking might be used in a clause where the speaker has some relevant and intentional involvement ("he's at the airport-EGO," e.g., because I took him). Egophore marking likely to be more distant from verb stem than TAM or person.

Egophoric marking centrally codes epistemic source or authority. It might also code for control and agency, such that NONEGO marking with first person can indicate unintential acts ("I ate a bug-NONEGO," i.e., by accident).

7.5. Valency. Possible that applicatives can only be used for animate arguments. Less often, may only be used with intransitive verbs.

It may be that applicatives in general are less frequent in languages with rich case systems (less necessary).

Antipassive marker may code for humanness (Rgyalrong).

In languages that restrict causatives in transitives to verbs of congnition and perception (know > inform), verbs of ingestion and consumption (eat > feed) may pattern with them.

7.5.1. Passive. Most often synthetic, with the marking usually, but not always, closer to the verb stem than TAM marking.

If auxiliary, most common intransitives are *be, become,* and verbs of motion; most common transitives are *get, receive, suffer, touch;* even *eat* is attested.

Accessibility in promotion of passive: DIRECT OBJECT > INDIRECT OBJECT > ADJUNCT.

Actor most often represented by an oblique: instrumental, locative (preposition, *by*, ὑπο, etc.), genitive. Rarely, special marking is employed.

7.6. Semantic Types and Their Frames.

AFFECT (*hit, cut*)	Agent	Target	Instrument	
GIVING (*give, lend*)	Donor	Gift	Recipient	
SPEAKING (*speak, tell*)	Speaker	Addressee	Message	Medium
THINKING (*consider*)	Cogitator	Thought		
ATTENTION (*see, hear*)	Perceiver	Impression		
LIKING (*like, love, hate*)	Experiencer	Stimulus		

In some languages, some *affect* verbs may require an inanimate agent, usually specific, such as food making a person sick, *etc*.

In very few languages are LIKING verbs like ANNOYING verbs, where the experiencer and stimulus are in O and A roles. However, extended intransitives, with an oblique experiencer, are a bit more common.

ATTENTION and LIKING may have extended intransitive frames. Or dative subjects. Or oblique objects.

[5]Some scholars consider egophore a kind of person marking. This seems less likely as more data comes in.

WANT verbs may have extended intransitive frames.

The types with more than two arguments may have lexical or construction splits to determine which of the arguments is in the primary O slot ("tell" vs. "speak (French)" vs. "say").

For GIVE there may be i) single verb with different constructions, ii) gift role is always in O function, iii) recipient role always in O function (rare; amusingly, with DAT for the gift role). Or, ii and iii may have different lexical items.

7.7. Associated Motion. In addition to the basic trans- and cislocative, which link a motion to the main event of the verb, they may code relative time of motion (prior: DO.ARRIVING, "when, after going, X"; simultaneous: DO.GOING, "while going, X"; prior: DO.AND.LEAVE, "before going, X"). They may distinguish simple deictics ("come in order to") vs. associated motion ("come while doing"). There may be a distinction between "go" and "come" vs. "go back" and "come back." Arrernte distinguishes a "quickly" or hurried series, as well as "do coming through." No deictic center may be present for a small subset (e.g., "do here and there"). In some systems, the elaboration for S/A is reduced, but there is a series for O on transitive verbs,[6] "I saw-GO(O) a man" for "I saw a man going away from me." Finally, the target of motion may be coded for stability (temporary or permanent, "enter.COME.TEMPORARY" "come in," vs. "sit.GO.PERM" "go and sit and stay there").

Very rarely the two core associated motion affixes may be used together for "away and back" sense.

Expansion of the basic system often includes "up" and "down" senses (and these may be the only extensions); and then "in" and "out." Erromangan has a misdirective, "away from expected direction, off to the side," etc., and Northern Paiute includes random motion in the set.

Rarely these affixes may be confined to imperatives. Affixes might be different depending on transitivity of verb, number of subject, or, very rarely, tense or aspect (probably from SVCs).

These locate an event in space much like tense locates an event in time. Though using these may add locative arguments to clause, the main verb is always foreground. Further, they may link discourse, with things like "he came, and he argued-CIS with me," with "come" and CIS marking the same path information.

Directional affixes may be grammaticalized in various ways (inward = perfective, downward = progressive, upward = imperative; cis = inceptive, change of state, trans = endpoint of activity); or more metaphorical senses (down = not fully satisfactory, up = better, back = return to health or satisfactory state).

Like tense and aspect, these locatives may be restricted or forbidden in subordinate or nominalized clauses (these are inflectional, not derivational).

Associated motion appears to occur in about 1/3rd of languages, with certain hotspots (Australia, the Andes and Amazonia). It is less likely in languages with serial verbs. The forms appear to be diachronically unstable, with potentially wide variation between otherwise closely related languages.

7.8. Participles. Participles may have an *inherent orientation* or a *contextual orientation.* European languages have inherent orientation, where the role participant is determined by the participle form (active participle for agent, passive for patient). In contextually oriented participles, several different participants (agent, patient, location) can all be encoded, and usage determines the interpretation. The relative accessibility hierarchy will be in play here, too. Contextual oriented participles are generally generous in what they will accept on the hierarchy.

ℜ: Separate inherent orientation participles distinguishing S from A role.

Locative orientated participles ("live.in-PCPL city") and instrumental orientied participles ("book write-PCPL pen") are possible, but rare-ish. Instrumental may be interpreted as reason, attaching to a generic noun meaning "way" or "time."

It is possible to have passive participles (i.e, oriented on P) in languages which do not have finite passive constructions at all. It might be obligatory to state the agent in such constructions. Passive participle forms are likely to be the form used to relativize further down the accessibility hierarchy.

[6]With deixis oriented to the verb subject. Not sure if that's universal.

S+P absolutive participles may only be permitted for low-agency intransitives ("fallen leaf" but not "danced woman"). Telicity sometimes a core consideration.

Participle forms often have several functions in addition to adjective-like attribution: adverbial ("watching the children, ..."), clausal argument ("I saw he was not running"), action nominals ("running is good"). Participles and nominalizations may be hard to distinguish formally.

Sometimes participle markers include TAM information, sometimes not. Markers that do not include TAM information are sometimes restricted in what verbal TAM marking they may occur with. Or, separate TAM markers are used with particples. When the participles code TAM, it is normal for fewer distinctions to be made. Finally, participles may make no TAM contrasts at all.

There may be separate negative participle forms. Also, particular participle forms may not be permitted with negation at all, requiring a clause to be recast into a particular participle form to be negated (such as the absolutive participle in Georgian). Nominal negating constructions may be preferred. Rarely, it might not be permitted to negate any participle types.

Arguments to participles may be as for verbs, or as for verb nominalizations (genitives frequent).

Even languages without an adjective class can have things quite like participles, which are used in relative and attributive constructions.

Participles in some languages are the only way to manage a relative clause. Or, some marker plus the participle forms a separate relative construction.

7.9. Converbs. Moderately more likely in SOV languages. Most things called converbs require same subject, but an argument can be made for inflected converb-like functions being part of the same phemomenon (cf. Coptic). Forms not inflected for person may or may not require same subject, though same subject is most common. Argument marking may be identical to a main clause, or different. There may be restrictions on negation. Rarely, converb clause may occur at the other margin of the clause, often with iconic restrictions, as in a purpose or intention converb allowed to follow the clause, while imperfective may not (in an SOV language).

Simplest functions: aspectual, imperfective (simultaneous) and perfective (anterior, usually). Some languages have many forms (all from Northern Akhvakh): locative (may take location case marking), inceptive ("from the moment X-ing began"), immediate ("as soon as"), anterior ("before X-ing"), imminent ("just before X-ing"), nonposterior (so fast that the converb event may not have time to occur at all, "come down here-CVB before something bad happens"), conditional, concessive ("although"), similative ("in the same way as"), gradual ("the more..., the more...," with the main clause just normally marked), cause ("because"), purposive ("in order to").

In languages with rich converb inventories, some may be highly restricted with respect to the main clause verb: Akhvakh progressive converb only occurs with "be, remain, see, find."

In Turkic languages (and some others), converb forms are the LEX in auxiliary constructions.

8. The Adverb

In languages otherwise without a well-defined adverb class, or a regular way of forming adverbs, there may be root adverbs for the domains SPEED ("quickly, slowly") and VALUE ("well, poorly, bad"). In addition to root adverbs, these can show up as derivations from adjectives which may differ from other adverb-like constructions somehow (such as being zero-marked). These two semantic cores are rarely the only one in any particular language.

9. Number

The distinction between "paucal" and "plural" is context dependent.

If there is a trial, its use may signify salience of the number, with the plural being used for three much of the time.

Number suppletion in verbs: not common, aligns ergatively. Usually only 1–4 verbs take it, but may reach a couple dozen. If there is a sg/du/pl distinction, at least some verbs will just have sg/pl (or non-pl/pl). Among intransitives, the posture verbs are most commonly suppletive, *sit, stand, lie,* and most likely to show numerous number distinctions (du). Next: *enter, go, be big, die/be dead, hang, arrive, run, come, fall, cry, be little*. Transitives: *kill, put/place, throw, give, break, take, bring, carry*.

9.1. Countability Classes. In languages with a singulative-collective distinction the hierarchy is: substances < aggregates < individuals; with individual most likely to be singular by default, with a marked plural. In languages with singulatives, aggregates most likely to take those.

Singulatives are rather common with small fruits and vegetables, small animals, sometimes paired body parts, and groups of people (Welsh: moch *pigs* vs. mochyn *pig;* mwyar *blackberries* vs. mwyaren *blackberry*).

10. Constructions

Any schematic construction may, like lexical constructions, be polysemous. For example, the English ditransitive:

- X CAUSES Y TO RECEIVE Z (central sense), "Joe gave Sally the ball."
- Conditions of satisfaction imply X CAUSES Y TO RECEIVE Z, "Joe promised Bob a car."
- X CAUSES Y NOT TO RECEIVE Z, "Joe refused Bob a cookie."
- X ACTS TO CAUSE Y TO RECEIVE Z at some future time, "Joe bequeathed Bob a fortune."
- X ENABLES Y TO RECEIVE Z, "Joe permitted Chris an apple."
- X INTENDS TO CAUSE Y TO RECEIVE Z, "Joe baked Bob a cake."

Similarly, a particular language may have one construction for a particular job, such as a ditransitive construction, but others have several with different pragmatic or pivot significance ("I gave him the book" *vs.* "I gave the book to him").

10.1. Grammatical Relations & Alignment. Grammatical relations are sets of arguments that are treated the same way by some contruction, such as being assigned the same case or causing the same kind of agreement.

Nouns and verbs may have separate alignments, such that nouns may align ergatively but the verbs nominatively, for example. Grammatical relations are construction-specific.

S = intransitive subject, A = (di)transitive agent, O = transitive object (often this is P), T = ditransitive theme, G = ditransitive goal (or ground).

Rarely, transitive A is distinguished from ditransitive A (and not by a separate case, but by a separate construction).

- {S} intransitive subject, nominative
- {S, A} subject, nominative; accusative alignment
- {A} transitive subject, ergative
- {O, T} direct object, accusative; indirective alignment
- {O, G} primary object, dative; secundative alignment
- {T} secondary object
- {G} indirect object, dative

- {S, O, T} absolutive; nominative; ergative alignment
- {S, O, G} absolutive; nominative; ergative alignment

O arguments may map to different relations based on: animacy, humanness, definiteness, specificity or saliency. In Nepali, animate O are marked as G. In Swahili, O is marked on verb only if animate and/or known. Or, non-salient O may be incorporated into verb.

A arguments are the inverse of O: the lower the salience or animacy the more likely they are to get overt marking, so that a human may get nominative, a stone ergative. In some languages inanimate {S, A} is basically impossible, with voice changing or incorporation to handle things like "the wind broke it."

- speech act participant > kin/name > human > animate > inanimate > mass
- specific > nonspecific referential > generic/nonreferential
- known/topical/thematic/definite > new/focal/rhematic/indefinite
- singular > plural

Role marking may be sensitive to other roles within the clause. Yurok only marks O when there is a 3P subject. Sahaptin only marks A if the O is a SAP, Tauya only if O is human.

It is unusual for languages to have obligatorily filled grammatical role positions (very much unlike English).

In Eastern Kiranti, a specific O, G argument can be an NP (adjective attribution, etc.), while nonspecific must be a bare N.

Floating quantifiers may prefer a particular grammatical relation. In Tagalog, *lahat* "all," floats to Wackernagel, but goes with the proximative relation. In Yélî Dnye, quantifiers float to the preverbal position, but go with {S, O}

Focus constructions may be different for different relations.

The non-ergative clauses in so-called split ergative systems are rarely nom-acc (A = NOM, P = ACC, S = NOM). Instead one can get A = ABS, P = ABS, S = ABS (common), A = ABS, P = OBL, S = ABS (common) or even A = ERG, P = ABS, S = ERG (common in Mayan). The aspect splits often reflect (historically) intransitive AUX constructions. Thus, in rare circumstances, one may get splits in future clauses, or with negation, if they reflected an original AUX construction.

10.2. Complement Clauses. These come in three types.

1. Fact type, indicate that something did take place. Similarly marked to main clause. If subject is the same across clauses, it isn't likely to be omitted. Usually marked with some complementizer element which will have other functions in the language (often "say" or "be like"). Complementizer may code reliability (sure fact vs. possible fact).

2. Activity type, indicating extension in time. Often similar to a noun phrase, but will still have a subject. If the subject is the same, it may be omitted. Or the verb may have a special form. Generally less specified in TAM and negation than a main clause; may not include same bound pronominal elements.

3. Potential type, typically less like a main clause than the Fact type and less similarity to a NP than the Activity type. In some languages, the subjects must be the same. Reduced TAM and pronominal marking. Implicit type reference to same or posterior time. Generally a special verb form ("infinitive") or may take marking similar to dative or some other case.

Languages may range from one to 5–7 complement types, with subtypes.

Attention verbs ("see, hear, show") typically take Activity complement. May take Fact, for completed actions or of state. "Find, discover" are expected to take Fact.

Thinking verbs ("think (of, about, over), consider, imagine") take Fact, or sometimes Activity ("think about"). "Assume, suppose" take Fact. "Remember, forget" take Fact, with English unusual in taking potential ("I remembered to shut off the stove"). "Know, understand" take Fact or Potential.

Deciding ("decide, resolve, choose") take Fact or Potential.

Liking ("love, prefer, regret, fear") take Activity or sometimes Fact. "Enjoy" takes Activity.

Speaking has several subtypes. "Say, inform, tell" usually take Fact. "Report" takes Fact or Activity. "Describe, refer to" takes Activity. "Promise, threaten" takes Potential, which may be in the indirect object slot. "Order, persuade" generally take Potential.

10.2.1. Epistemic Stance. In addition to things like the subjunctive, the epistemic stance towards an embedded clause can be affixes, clitics, particles, adverbs, and auxiliaries. Jacaltec has different two complementation conjunctions for verbs of speaking, one indicating certainty, one indicating reservation or hearsay.

10.2.2. Expletive Negation. In fear clauses, especially if not indicative, negation may be a requirement of the construction, though it is not negating the fear clause.[7] If a language allows different complement constructions for fear-verbs, one may take expletive negation while the other does not (cf. Russian *čto* vs. *kak...by ne*).

These may evolve from wish constructions, or from negative purpose clauses.

When fear clauses look like embedded questions, even if indicative, expletive negation is more likely.

Fear-verbs likely to take this: *fear, prohibit, hinder, prevent, avoid, deny, refuse*. In a few languages positive predicates (*hope*) might have some constructions with expletive negtation.

Expletive negation may also occur in: exclamatives, emphatic questions, concessive conditionals, before clauses, until clauses, polite requests, comparatives.

10.3. Desententialization. The most desententialized is purpose clauses: purpose $<$ before $<$ after, when $<$ reason, reality condition. Also: phrasal, modal $<$ desiderative, manipulation[8] $<$ perception $<$ propositional attitude,[9] knowledge $<$ utterance.

10.4. Negation. Negative affix only a bit less common than negative particle. Very rarely, might be circumfix.

Asymmetric negation (different structure under negation beyond mere presence of marker) reasonably common, with a mix of symmetric and asymmetric within a single language a bit more common than either alone. Possible asymmetries: lexical verb loses finiteness (sometimes with AUX negator, sometimes other AUX with nominalized, negated LEX); negatives use irrealis of some sort; negatives have some sort marking used for asseveration in non-negated; difference in grammatical categories (TAM, argument structure; Harar Oromo removes all person marking distinctions in negative). Negation is stative, so negated clauses may partake of some stative construction behavior. The discourse context of negatives *contra* affirmatives necessarily means the details matter less.

Nonstandard negative constructions more likely in: imperatives, existentials, nonverbal clauses.

Negative imperatives might be (with WALS counts): usual imperative + usual negation (23%); usual imperative + different negation (37%); different imperative + usual negation (11%); different imperative + different negation (29%).

Circumfix all over planet, just not common.

Possible distinctions, beyond simple "not:" noun vs. non-noun, verb vs. non-verb (where adj. may pattern with either); different negator for the future; prohibitive; participle negation; negative particle for non-existence (which may be like "no" or plain negation for nouns). Non-verb negator can be grammaticallized from "other, different; refuse, not want."

10.5. Questions. Focus questions ("do *you* want it?") usually look like polar questions, but might have a special affix or structure; could have affinity with alternative question construction.

Rarely, negative polar questions might be distinct from positive polar questions.

[7] French has it optionally: *je crains qu'il* (*ne*) *vienne* "I'm afraid he might come." vs. *je crains qu'il ne vienne pas* "I'm afraid he won't come."

[8] *urge, suggest,* etc.

[9] *believe* and the like

Having unmarked polar questions is extremely rare, whether lacking an affix declaratives do have (Sheko) or being identical in form and intonation to a statement (Yeli Dnye). Most common is a marked polar question and unmarked declarative, though marked declaratives can happen (Crow, Sabanê).

Intonation is among most important ways to mark questions, though it usually is accompanied by other markers. Less than 20% of one survey (of 955 languages) use intonation alone. The rising pattern, as in English, is in no way universal. In West Central Africa, an areal feature of questions involves falling pitch, vowel lengthening, and a breathy end to the intonation unit.

Word order change to mark questions is extremely rare, basically confined to Western Europe.

ℜ: Koasati uses a glottal stop infix as question marker (usually morpheme boundary).

ℜ: Halkomelem uses auxiliary verb for polar questions.

10.6. Conditionals. The protasis of a condition and formal marking of the topic may be identically or similarly marked (including "if X ADP" or the like for topic marking).

11. Discourse

11.1. Definiteness. In languages with (in)definite articles, where indefiniteness is marked by artices within the NP, indefinite nouns tend to be unrestricted in core argument position.

In languages without (in)definite articles, non-specific arguments are more likely to have restricted core case functions, and definiteness distinctions tend to be encoded in the verb phrase, generally with various valency operations (DOM, incorporation, antipassives, etc.). In such languages, existential expressions are preferred to introduce novel discourse topics. Even in a language with articles, such as French, disprefers indefinite subjects, using an existential + rel. expression in colloquial speech. If semi- or de-transitivizing verbs forms are available in an article language, they might be used.

Subjecthood cline: definite > generic > specific indefinite > indefinite. In some languages, subjects *must* be definite with event verbs.

Objecthood cline: definite > specific indefinite > generic > indefinite. Again, indefinite objects may trigger valency reducing forms, making objects oblique (or incorporated).

Individuation may be part of the clines above, with, for example, "the stars" as a mass treated more like an indefinite than "the moon."

11.2. Topicality. Generic topicality hierarchies: speaker > hearer > 3RD; human > animate > inanimate; agent > dative > patient; large > small (and adult > child); possessor > possessed; definite > indefinite; pronoun > full NP.

Referential hierarchies: Speech-act participants > Kinship/Name > Human > Animate > Inanimate; Specific > Non-specific referential > Generic; Known/Topical/Thematic/Definite > New. Different languages take different approaches for SAPs, with both 1 > 2 and 2 > 1 found (the latter a politeness matter, apparently) and number may play a role, too. Found in the wild: 1PL/2PL > 1SG > 2SG, 1PL > 2, 2PL > 1 > 2SG.

11.3. Focus. Focus constructions may differ by argument type, clause type, polarity, etc.

Focus markedness (from least to most): indefinite NP, definite NP, pronoun, bound pronoun, zero.

WH-questions generally require a focused answer. Further, if the language has *ex situ* WH words (initial, preverbal, etc.), the focus phrase will usually end up in the same position.

Clauses with focused elements tend not to have overt topics.

Focused subjects may move into object position. In an SOV language, the focused subject may move immediately before the verb. In a very few languages, subjects may be incorporated into an intransitive verb to focus them.

In some (SOV) Omotic languages, what look like the subject agreement affixes (clitics?) can move around and attach to a focused element. With content questions, the often fronted question word always gets the subject. In Huallaga Quechua, the evidential marker goes on the focus (or the verb, in the most neutral, predicate focus situation).

11.4. Functions. Discourse particles, expressions, and adverbs serve a set of core discourse functions.

Structure. What are usually called PUSH and POP introduce and recover from minor digressions, "by the way, but, anyway." The CHECK function asks the listener to verify, "isn't it?" There are two editing functions, REPAIR which fixes information, "no, rather...," and EXEMPLIFY which gives an example, "for example, for instance." The HESITATE function is filler to let the speaker figure out what they're going to say next. UPTAKE indicates that the speaker has heard and understood the listener, and is ready to move on, "so."

Coherence. These functions address knowledge assumptions and attitudes, with KNOWN indicating assumed shared knowledge, "as you know, indeed," and REVISED indicating an updated assumption, "after all." A speaker might restate something "in other words."

Attitude. POSITIVE and NEGATIVE "unfortunately." INDIFFERENT "as far as I'm concerned." SURPRISE "did you actually?" and PREFER "I would rather." The SMOOTH function can generate a vast stream of words which tone down, lighten, and moderate face threats. And EMPHASIZE marks something as high on a scale, "really."

Discourse particles in practice might indicate only one of the major functions, or might, depending on use, cover several. Alternately, several functions might be joined into a single discourse construction that expresses all those functions simultaneously. A discourse construction may have a particular function interpretation only questions.

Different function constructions may be possible in different places in a discourse: initiation of a turn, reply (or reply-initiate) to a turn, midclause, or uttered alone, with different nuances in meaning possible.

Discourse construction frequency and selection may vary by register, genre, politeness, etc.

12. Lexicon

Consider a few phonesthemes.

Diminutive/medial/augmentative may code gender of the speaker (Weining Ahmao, in the classifier system).

Special formal/elevated vocabulary word shapes may have several patterns of systematic relationship between it and base level words (Javanese).

VSO languages are more likely to have even noun-to-verb ratios (lots of verbs derived from nouns), while SOV languages are more likely to have N+V idioms taking up the slack, resulting in more nouns. This includes SVO languages that lean OV. Regardless, there are almost always more nouns than verbs (though perhaps not by many).

Respect forms may contain morph for "lord" or "sky."

Some Austronesian languages have an "anger" vocabulary, not swearing, for verbs and nouns. Some common deformation patters.

Hunting, fishing or territorial languages: to conceal your intent from animals or spirits, or with tabu considerations; circumlocution and non-systematic deformation common practices for this.

Body parts become special concepts in stages: 1) a region of the human body, 2) a region of an inanimate object, 3) a region in contact with an object and 4) a region detached from the object. The landmark path ("extremity, peak" $>$ "head") goes in reverse.

Over time words move from external and objective to subjective and grammaticalized ("boor" from farmer to oaf, "feel" from touch to experience emotion, "insist" from perservere to demand to believe strongly).

Once a conceptual metaphor has taken root (SEEING is UNDERSTANDING) vocabulary merely related to the concept of seeing may also be dragged into the metaphor later. You don't expect "brilliant = intelligent" to happen

until the first metaphor is well established.

12.1. Movement. Verbs of rotation may distinguish features: internal vs. external axis; elevation over landmark; control; single vs. repeated turn.

12.2. The Body. Many languages have a "psychic being" that experiences emotions, "my psyche became frightened, is happy," etc., instead of the entire person experiencing them. This psychic being may be an internal organ: liver, heart, stomach, diaphragm, etc.

Some body parts may be conceived of as parts of other body parts, where "my lips" must be "the lips of my mouth."

Some languages with obligatory inalienable possessives on body part nouns may have a special "derelationalization" construction. Tzutujil has a suffix (while possessors are a prefix).

Inessential body parts ("beard, scar") may be treated differently than vital ones. Predicate possession of a vital body part is likely to have additional interpretations ("I sit with a belly" = satisfied with food; "he really has ears" = he has good ears).

Body part properties may be expressed: "on/to W there is an A P," "W has an A P," "W is provided with an A P," "W's P is A." A-P (or P-A) compounds are not uncommon, as both properties and exocentric nouns.

Instrumental function of body part terms may be marked in unexpected ways ("go on/by foot; carry in my hand, on my shoulder").

A body part may be etymologically related to its most obvious or common function (such as "fist" < "pound").

A person experiences and interacts with the world via their body. Expressions describing that may involve the person or the body part in many ways across languages.

- The person may be possessor ("wash his face," "he died by the hand of the executioner"). In some languages the person is excluded strongly except as possessor ("he hung his neck" rather than "he hung himself").
- The person may be indicated with an oblique (often dative), ("to me the belly aches, I wash to you the hands, I burnt to me the hand").
- In topic-prominent languages, the person may be the topic ("elephant-TOP nose is long").
- The body part may be a locative ("I have a pain in my ear"), ablative ("I have a pain from my ear"), or some other location construction.
- Or, the person may be the locative ("at the child the head aches").
- Body part terms may be assumed to be part of the agent or subject of a clause ("I opened the eyes" = my eyes, "she snapped with the finger" = her finger).
- In a few languages, both person and part may take identical role marking ("dog-ACC you-grabbed tail-ACC").

12.3. Location and Posture. Some languages are more location precise, some are more posture precise. In either situation, there is enough information for a listner to figure out the answer to "where is X?" by insepction.

Unmarked location construction for stereotypical, canonical situations have several types: Type 0: no verb at all in BLC (Basic Locative Construction); Type 1: single locative verb (1a copula, as English; 1b locative or existential verb, as Japanese); type 2 which has a small, 3–7 set of locative verbs; type 3, which has a large set of dispositional verbs, 9-100, such as Mayan. A few languages are mixed, such as Goemai which uses both type 2 and 3 together in SVCs.

BLC hierarchy: animate ground > figure pierced > ground pierced > adhesion > core scenes (ring on finger > apple on skewer > arrow in apple > stamp > cup on table / fruit in bowl). The core scenes are most common BLC use, with the full hierarcy quite rare (of course, English).

Non-canonical location will use more complex constructions.

Locative copula may make animacy distinctions (Japanese). Or, it may distinguish from known and unknown (used in questions, or when the situation is unclear).

Type 2: *sit, stand, lie, squat, hang, be near, be located as a mass, be supported in a medium* (*like water);* also dispersion, *spread, cover,* and attachment *be fixed, be stuck to, be bound*. Which verb is used for a particular item may have complex decision trees invoking not only orientation, but prominence, animacy, respect, containment ("the dog sits in the living room" even if it is standing, due to the room), contact, duration of state of affairs, size relation between figure and ground. Default may need to be memorized like gender, but non-canonical orientation means a different verb can always be used. "Sit" seems often the default for uncertain situations. "Hang" may get quite surprising usage (implies permanence, due to attachment, or customary behavior, such as terrestrial animals going about their normal business). May be separate "stand" verb for inanimates.

Type 2 verbs tend to grammaticalize: imperfective, progressive (> present), or become articles. The default posture verb for an item may pervade the language across all constructions in subtle ways: Yêlí Dnye has not only posture verbs, but internal action (*sit down, stand up*) verbs of putting (*put standing, put sitting*), verbs of taking (*take sitting X, take hanging X*).

12.4. Perception. Expressions of cognition frequently resort to verbs of seeing, hearing, smelling, or feeling. "Seeing = knowing" is not as universal as European languages make it seem, but is still widespread across various features (evidentials often make the visual evidential the default, least marked).

A single verb meaning both "see" and "hear" is possible, with instrumental (ear) or object nouns (voice, sound) indicating a "hear" sense is intended.

Polysemous senses of perception verb *roots* often disambiguated by different case structure or complementation constructions. Perception verbs occasionally align with cognition verbs in using different case relations.

Perception verbs might reject derivation (causative, passive, etc.).

Imperatives of several perception verbs (and "know") used as attention getting discourse markers.

"See" might: warning, "look out, beware;" "discover;" suggest special knowledge; "know" in perfective or stative forms; "look after;" "have an opinion, judge." Or: "desire." Expressions of socializing might use sight verbs. They eye may connote (sexual) desire or aggression in general (extended eye contact expected in some cultures, but highly aggressive and threatening in others — communicants might not even look at each other). Seeing certain things is a common tabu, but hearing rarely comes in for such tabus. Sight might be supernatural.

For sight, may be distinct lexical item depending on thing seen, breaking along: surprising object, new object, old object, fact, future, experiential perfect ("I had seen this man"), present (more than one break possible, though not more than three lexical items seen in paper data). Some languages may make one end of this continuum the most common default visual verb, others the other end.

"Hear" might: knowing, understanding, remembering. Often, "obey." May be colexicalized with sense of smell. The ear can be the organ of understanding, memory, but also of emotions, intention, and obligation.

Highly elaborated vocabulary for tastes and smells happen, but are not common. In the languages that do have richer aroma terms, food smell vocabulary is common, possibly including cooking method; the smell of rotting and feces seems to be common; animals, as well as blood and/or fat in meat preparation; strongly medicinal or herbal aromas; mold and mustiness.

Experience	Activity	Phenomenon
see	look (at)	look, seem
hear	listen (to)	sound
feel	feel/touch	feel (like)
smell	smell	smell (of/like)
taste	taste	taste (like)

There may be different or identical lexical items of any of the types; or differing by construction (as English for the bottom three). S=O lability (as in English "taste") applies even to sight verbs in some languages.

Positive and negative judgement of perceptions are likely to split for sight ("beautiful, ugly") with just "good" and "bad" for the other senses much of the time, with occasionally specialized terms following no particular pattern.

Viberg's lexicalization and markedness hierarchy: sight > hearing > touch, taste, smell (i.e., more lexical distinctions for sight than for smell).

If there's a lexical difference between experience and activity, experience verbs will often be more transitive. In ergative languages this holds less often, and might even have intransitive constructions for both.

May be internal, "proprioceptive" sense constructions, for things felt happening to oneself ("feel" vs. "feel internally" is attested).

Distinction between objective ("look") and subjective ("seem") (or non-mediated vs. mediated). May be a single mediated verb across all senses.

A [-control] root such as "see" may become [+control] in the imperative ("look at!")

12.5. Pain. Crosslinguistically few root words to describe pain. Rather, idioms from four domains are commonly used: burning (including smolder, shine, boil), destruction or deformation (sharp or pointed instrument; break, tear; press, pull; burst), sounds (usually for sensations in the ears or head), and motion (circular movement, twist, jump). Atelic verbs may be nominalized to produce telic/stative constructions.

The body part experiencing pain may be locative ("in my hand"), a subject ("my hand hurts"), or an agent ("my hand hurts me").

Loss of functionality is often expressed by lack of motion (lock up, stiffen, be like a stick).

Derivations locating pain or illness (*-osis, -itis, -algia,* etc), as in modern medical terminology, are quite rare.

12.6. Temperature. Core words for temperature may make only two distictions ("hot/warm" vs. "cold/cool"), three ("cold/cool" vs. "warm" vs. "hot"), or four ("cold, cool, warm, hot"). Some languages may have unique intensifiers for extreme temperatures, usually related to words for processes or things that exemplify the temperature (such as ice or burning).

Temperature may refer to three domains: tactile ("the plate is hot"), ambient ("it's hot here"), and personal-feeling ("I am hot"). Languages like English or Italian may use a single word ("cold," "freddo") for all three; there may a unique word for each domain, as in East Armenian; or the domains may be split up among two words. Expressions for the personal-feeling domain are most likely to be different in some way. The ambient domain generally makes the most distinctions, the personal-feeling the fewest. Personal-feeling terms may be restricted to "uncomfortably hot" and "uncomfortably cold."

Terms in the ambient domain may mark source of the heat, humidity, the effect on breathing, windiness, etc., such as one for "hot" and another for "hot and humid." Some terms may be restricted to season, such as Palula "cool" *šidaloó*, which can refer to pleasant coolness in summer, but is not used to describe not-too-cold in the winter.

There may be separate tactile domain temperature terms for food and especially water.

12.7. Clause Connection. Historical sources for *and:* adv. and prep. of linear succession, "in front, after, before, then" (as English *and* from PIE **hanti* 'in front'); focal additive particles, "also, too;" paragraph linkers ("besides, moreover, and then"); comitative markers ("with" > "and"); verbs meaning "go, bring" in narrative contexts (Hdi *là* "to go" > "and"); pronominal roots (PIE **tó* > Hittite *ta* "and").

Historical sources for *or:* distal "that, other" (PIE **au-* > Lat. "aut"); interrogative particle; free choice verbs (Fr. "soit... soit..."); dubitative particles "perhaps," etc.; denied conditional clause "if not, if it is not so."

Historical sources for *but:* spatial meaning of distance (separation), closeness or opposition (OE "be utan" > "but," "in stead," German "sondern" separate > "but rather"); temporal meaning of overlap, "while;" temporal mean-

ing of continuity "always" (Eng "still" constantly > "nonetheless"); causal and resultative ("therefore"); comparative meaning "more, bigger" (Lat. *magis* > It. "ma").

Semantic map of contrast: oppositive ("I bought X, but he bought Y") > corrective ("I'm not Xing, I'm Ying") > counterexpectative ("I'm tall but bad at basketball").

12.8. Compounding. There are quite a few possible relationships between the head H and modifier M in N+N compounds. Compounds might be formed by 1) simply cramming nouns together, as in English; 2) root + some linking element + N (for euphonic reasons, or some other ligature required in the language in compounds); 3) N PREP N (as in French *de* and Japanese *no*); 4) genitive; 5) N + head-marked N (Turkish); 6) relational adjective + N (such as "iron-ADJ + road" for "train" in Russian).

- Similarity
  - Taxonomy: an M is a kind of H: *oak tree;* an H is a kind of M: *bear cub*
  - Coordination: an H that is also an M: *boy king*
  - Similarity: an H that is similar to M *kidney bean*
- Containment
  - Containment: an H that is contained in M: *orange seed;* an H that contains M: *seed orange*
  - Possession: an H that is possessed by M: *family estate;* an H that possesses M: *career girl*
  - Part: an H that is part of M: *car motor;* an H that M is part of: *motor car*
  - Location: an H located at/near/in M: *house music;* an H that M is located at/near/in: *music hall*
  - Time: an H that occurs at/during M: *summer job;* an H at/during which M occurs: *hunting season*
  - Composition: an H that M is composed of: *wheat flour;* an H that is composed of M: *sugar cube*
  - Topic: an H that is about M: *history book*
- Causation (source, goal)
  - Direction: an H whose goal is M: *sun worship;* an H that is the goal of M: *sales target*
  - Source: an H that is a source of M: *sugar cane;* an H whose source is M: *cane sugar*
  - Causation: an H that causes M: *tear gas;* an H that M causes: *sunburn*
  - Production: an H that produces M: *song bird;* an H that M produces: *birdsong*
  - Usage: an H that uses M: *oil lamp;* an H that M uses: *lamp oil*
  - Function: an H that serves as M: *buffer state*
  - Purpose: an H intended for M: *animal doctor*

H and M elements in compounds may be metonymic ("cottage cheese, spaghetti western, godfather, roadkill") or metaphorical ("milk tooth" and the "-father" element in "godfather").

12.9. Metaphors and Idiom. Common source: BODILY MANIFESTATION FOR EMOTION. The face often (actions, colors), but also eyes. Other organs get up to shenanigans, including blood.

Colors seem to often carry emotional connotations. What a particular color signifies will vary from culture to culture: in Thai, a green eye or face signify anger.

Metonymy: result for whole, salient feature for whole (where whole might be thing or event), instrument for action, effect for cause, producer for product, both whole for part and part for whole, perception for thing perceived (sound for event causing it).

Metonymies often chain (eye > vision > attention > desire, as in "I have an eye on that new computer" vs. "keep an eye on it"), with idiom and metaphor applicable at all steps. Back, buttocks > *back part* > *behind* > *after;* or *behind* > *follow, support*. Belly > *inside part* > *inside* > *inclusive, during;* or > *pregnancy* > *offspring*. Ear > *hearing* > *attention,* (*dis*)*regard, obedience, hearsay*. Eye > *vision* > *attention, beauty*. Head > *top part* > *over, beginning, end*. Mouth, tongue > *speech* > *word, speech act*.

In Kolyma Yukaghir there is a separate ACC just for use on 1/2 SG/PL when the subject is 1/2 SG/PL, the "pronominal accusative."

Form dependency: each element's forms may depend on choices made in the feature above them (such as fewer tense forms in the negative, for example). The final three are potentially interdependent: polarity $<$ tense, aspect, evidentiality $<$ person, reference classification (gender, etc.) $\Leftrightarrow$ number $\Leftrightarrow$ case.

Hierarchy of diminutives and augmentatives: noun $<$ adj., verb $<$ adv., numeral, pronoun, interjection $<$ determiner. These markers may be distributed over several words in a sentence.

If the noun takes prefixing morphology, so will verbs.

Serial verb constructions more likely in SVO languages than in SOV. SOV languages are more likely to use converbs, though a very few have both converbs and SVCs. Having a generalized "and" (*i.e.,* the same for NP and VP) is weakly correlated with not having SVCs.

Pseudo-coordination ("try and see") is typically restricted to a very small set of verbs, and in Germanic languages at least is often used in aspectual constructions ("sit and ..." in Norwegian for progressive, "return and..." in Arabic for repetition). Straight-up aspect verbs ("begin") can also use such a construction. "Go" and "come" with purpose may fall into this construction. Also used for: causative ("make/cause and"), intention ("plan and"), reason, conditional.

Writing systems: vertical and horizontal lines, easiest to identify, most frequent component of most symbols in a set. Obliques and diagonals do not mix with vertical and horizontals too much (K, A, Z, less frequent than E, H, F or W and X). Vertical symmetry (M, A, W) more common that horizontal (K, D, E).

13.1. Oblique Strategies. Don't ask what it does, but how it moves.

Don't ask what it does, but what it looks like.

Don't ask what it does, but who it is important to.

Don't ask what it does, but who it hangs out with.

Add something which has no history.

Verb: action vs. means vs. result.

Pick an animal and make it a central metaphor.

Pick an ancient technology and make it a central metaphor source (stonework, ceramics, weaving, herding, sailing and knots, etc.).

Meaning comes from familiarity.

What if the arguments or referent switch animacy?

14. Lexical Functions

These functions are a condensed notation to represent relationships in lexical collocation, though some will correspond to valency changes, derivations, etc.

14.1. Structure. These are used for narrative and pragmatic purposes, textual cohesion, etc., and are not simply empty. They allow one to shift focus, saliency, topic. Some of this will be handled by valency tricks in some languages.

FUNC$_0$() = LVC meaning *happen to take place* which has the keyword lexicalized as subject: *the possibility exists, time flies, the day passes by, the rain falls*.

FUNC$_1$() = LVC meaning *originate from,* connects keyword as subject with agent as DO: *responsibility lies* (*with sb*), *the blow comes* (*from sb*), *support comes* (*from sb*). **FUNC**$_1$(*blow*) = *comes from sb*.

FUNC$_2$() = LVC meaning *concern, apply to,* connects

keyword as subject with the object: **Func**$_2$(*blow*) = *falls upon sb.*

Oper$_1$() = "carry out, perform, act, do" LVC which connects the subject and the action as DO: *take a bath, vacation; have a look, bath, shower; give sb/sth a smile, laugh, shout*. Very common. **Oper**$_1$(*attention*) = *pay*.

Oper$_2$() = "undergo, meet" LVC which connects patient, recipient, experiencer as subject to action as DO: *get a benefit, have an attack* (*of a disease*)*, take advice, undergo inspection*. **Oper**$_2$(*attention*) = *draw*.

Labor$_{ij}$(=) LVC which connects *i*th element as subject to *j*th element as DO, with keyword as secondary object. **Labor**$_{12}$(*interrogation*) = *to subject sb. to an interrogation;* similar, *treat someone with respect*.

14.2. Fulfillment. All these produce verbs or LVCs that satisfy "to fulfill the requirement of, to do with X what you are required to, X fulfills its requirement, designed to." The fulfillment may not be seen as such by someone undergoing it: the fulfullment of some disease is death. May be different terms for different types of fulfillment (psychological, physical, etc.).

Fact$_n$() = syntactic actant *n* filfulls it's own requirement: **Fact**$_0$(*doubt*) = *be corroborated,* **Fact**$_0$(*knife*) = *cut,* **Fact**$_1$(*turn*) = *be someone's turn,* **Fact**$_2$(*ship*) = *transport people or cargo*. **AntiFact**$_0$(*accusation*) = *is fabricated,* **AntiFact**$_1$(*accusation*) = *withdraw,* **AntiFact**$_2$(*accusation*) = *denies charges.*

Real$_1$() = LVC for "act accordingly to the situation, use as forseen:" *exercise authority, use a telephone, speak a language, keep a promise*.

Real$_2$() = LVC for "react according to the situation," *respond to an objection, satisfy a requirement, give in to persuasion, get a joke, confirm a hypothesis.* **AntiReal**(=) *fail an exam, reject advice, turn down application*.

Labreal$_{ij}$() = LVC corresponding to **Labor**() above. **Labreal**$_{12}$(*gallows*) = *string someone up*. Others: *cut something with a saw, hold something in reserve.* **Labreal**$_{13}$() = *burn with shame, waste one's health*.

14.3. Verbal. Usually from nouns.

Copul() = copula: *work as a teacher, serve as an example,* **InCepCopul**(*ill*) = *fall ill.*

Involv() = involve, affect a non-participant: *light floods the room, snowstorm catches/hits, smell filled the room*.

Manif() = manifest, become apparent in someone/something: *joy bursts, scorn drips*.

Prox() = be about to, be on the verge of: *on the edge of despair, on the brink of disaster, verge of tears, thunderstorm brews*. Usually as **ProxOper**$_1$(*despair*).

Prepar() = prepare X for, get X ready for normal use. Usually as **PerparFact**$_0$(*car*) = *fill up the car*.

Obstr() = function with difficulty: *eyes blur, economy stagnates, short of breath*. **CausObstr**() common: *gun jams, rope tangles, traffic snarls*.

Son() = emit characteristic sound: *whip cracks, bell chimes, cane swish, leaf rustles*.

Stop() = stop functioning: *lose one's breath, voice breaks, heart stops* (*or breaks*).

Excess() = function in an abnormally excessive way: *heart has palpitations, engine races, sweat rolls down, teeth grind*.

Sympt() = represents bodily reaction to X. Joined in complex relationships with the rest: **Obstr**(*speech*) + **Sympt**(*anger*) = *sputters with anger;* **Obstr**(*breath*) + **Sympt**(*anger*) = *chokes with anger*.

14.4. Nominal.

S_n() = the *n*th participant (agent noun, object noun, etc.). S_1(*teach*) = *teacher,* S_2(*teach*) = *student,* S_3(*teach*) = *subject matter.* More interesting when modified, e.g., **(AntiBon)**S_1().

In addition are S_{loc}(), S_{instr}(), S_{mod}() (manner, *way of life*), S_{res}() (result, S_{res}(*split*) = *crack*), S_{med}() (means).

Cap() = the head of: *pope, captain, emperor,* **Cap**(*university*) = *president*.

Equip() = staff, crew of: *crew, company, personnel.*

MULT() = collection of: *bouquet, group*.

SING() = unit of entity: *rain drop, snowflake, act of violence*.

PEL() = covering: **PEL**(*bean*) = *pod,* **PEL**(*tree*) = *bark,* **PEL**(*book*) = *cover; binding.* **LIQUPEL**(*bean*) = *to shell,* **LIQUPEL**(*rabbit*) = *to skin*.

FAS() = "face," front: *front of house, bow or prow of ship, nose of plane.* **ANTIFAS**() = *ship stern, tail of plane*.

14.5. Adjectival.

$\mathbf{A}_n()$ = determining property of *n*th participant from the viewpoint of its role in the situation; quite like participles with verbs. $\mathbf{A}_0$(*brother*) = *fraternal,* $\mathbf{A}_0$(*city*) = *urban,* $\mathbf{A}_1$(*delight*) = *delightful,* $\mathbf{A}_1$(*anger*) = *in anger, angry,* $\mathbf{A}_2$(*shoot*) = *under fire*. $\mathbf{A}_2$(*analyze*) = *under analysis*. Often more useful modified.

ABLE$_n$() = can easily, prone to: **ABLE**$_1$(*cry*) = *tearful,* **ABLE**$_2$(*trust*) = *trustworthy*.

QUAL$_i$() = predisposed, of *i*th probable argument: **QUAL**$_1$(*cry*) = *sad,* **QUAL**$_1$(*laugh*) = *cheerful,* **QUAL**$_2$(*laugh*) = *awkward, absurd*.

14.6. Evaluation. May be combined with **ANTI**: **MAGN**(*temperature*) = *high*, **ANTIMAGN**(*temperature*) = *low*. May be quite different depending on word class, and take multiple forms: **MAGN**(*smoker*) = *heavy, chain-smoker;* **MAGN**(*to smoke*) = *like a chimney*.

BON() = good, generally held praise: *neatly cut, heroic struggle, fruitful analysis*.

CENTR() = center, culmination: *height of the crisis, summit of glory, prime of life*.

DEGRAD() = degraded, lowered: *discipline decays, house becomes dilapidated, patience wears thin, temper frays, teeth decay*.

MAGN() = immensely, very: *shave close/clean, condemn strongly, infinite patience*. Might be quantitative or temporal (speed).

VER() = real, genuine, as it should be, meeting intended requirements: *genuine surprise, walk steadily, loyal citizen, legitimate demand, precise instrument, well-deserved punishment, restful sleep*.

ANTIVER() might have "too much" or "too little" options.

14.7. Other.

RESULT() = the expected result of; **RESULT**(*buy*) = *own,* **RESULT**(*to have learnt*) =*know*.

CONV$_{ijk}$() = converse, reorders arguments. **CONV**$_{21}$(*include*) = *belong,* **CONV**$_{21}$(*precede*) = *follow*.

FIGUR() = figurative, standard received metaphors: *curtain of rain, pangs of remorse, flames of passion*.

14.8. Modification. These are only used in combination with others. **INCEPFUNC**$_1$(*anger*) = *anger rises*. **CONTOPER**$_1$(*power*) = *retain one's power*. **PERMFUNC**$_0$(*aggression*) = *condone aggression*

INCEP(begin, start) (= **ANTIFIN**())
CONT(continue, maintain, retain)
FIN(cease, stop)
CAUS(causative)
LIQU(liquidate, stop, divert) (= **ANTICAUS**())
PERM(permit, allow, condone)
PLUS(more)
MINUS(less)
ANTI(negates)

Apparently common blends:

ANTIMAGN

CONTFACT$_0$() = *luck holds*.

CAUSFUNC$_0$() = *find an answer, conduct a campaign, produce an effect*. Fairly common.

CausFunc$_1$() = *open the way, cause damage, give an answer*. Fairly common. Can be modified: **CausPlusFunc$_1$**(*risk*) = *increase, raise,* **CausMinusFunc$_1$**(*consumption*) = *reduce*

IncepOper$_1$() = *take an attitude, start a session, obtain a position.*

ContOper$_1$() = *keep silence, follow an example, keep one's balance, lead a busy life.*

A fuller example:
IncepOper$_1$(*habit*) = *acquire, form, take to.*
FinOper$_1$(*habit*) = *drop, get out/rid of.*
LiquOper$_1$(*habit*) = *break, wean from.*
Liqu$_1$Oper$_1$(*habit*) = *kick, shake off.*
CausFunc$_1$(*habit*) = *instill into, inculcate.*

These may also produce simultaneous functions, [**Magn** + **Oper$_1$**](doubt) = *be plagued by doubt.* [**Ver** + **Oper$_1$**](health) = *have a clean bill of health.*

15. Sound Changes

p

/p/ → [k] / _C
/p/ → [k] / u_#
/p/ → [kʷ]
/p/ → [ʔ] / _#
/p/ → [ʔ]
/p/ → [f]
/p/ → [f] / V_
/p/ → [ɸ]
/p/ → [h]
/p/ → [v] / V_V
/p/ → [f] / #_
/p/ → [h] / _{u,o}
/p/ → [w] / V_V
/p/ → ∅ / _st
/p/ → ∅ / V_tV

pr

/pr, kr/ → [bl, gl]
/pr/ → [t]

pt

/pt/ → [tː]

p'

/p'/ → [b]
/p'/ → [p] / V_
/p'/ → [p]

t

/t/ → [d] / #_
/t/ → [ʔ] / _#
/t/ → [n] / _#
/t/ → [s] / i_
/t/ → [s] / #_
/t/ → [h] / #_
/t/ → [k] / _l, _r
/t/ → [k] / _j
/t/ → [tʃ] / _#
/t/ → [ʃ] / i_i
/t/ → [h] / _t
/t/ → [tʃ] / _ʔ
/t/ → [r] / X_X
/t/ → [l] / X_X
/t/ → [tɬ] / _{a,u}
/t/ → [j] / #_i
/t/ → ∅ / n_
/t/ → ∅ / n_V
/t/ → ∅ / _st
/t/ → ∅ / s_
/t/ → ∅ / #_VrC
/t/ → ∅ / k_#

t'

/tʼ/ → [t]

tʷ

/tʷ/ → [p]
/tʷ/ → [tʃʷ]
/tʷ/ → [d]
/tʷ/ → [gʷ ~ g]
/tw/ → [tew]

tk

/tk/ → [g] / X_X
/tk/ → [k]

ts

/ts/ → [s] / V_
/ts/ → [s]
/ts/ → [sʼ]
/ts/ → [z] / V_V
/ts/ → [ʃ] / _r
/ts/ → [ʃ] / V_
/ts/ → [tʃ] / _i, V_
/ts/ → [ʃ]
/ts/ → [h]

tsʼ

/tsʼ/ → [s] / V_
/tsʼ/ → [s] / V_
/tsʼ/ → [tʃʼ] / V_
/tsʼ/ → [ʒ] / V_
/tsʼ/ → [sʼ]

tɬ

/tɬ/ → [k]

tɬʼ

/tɬʼ/ → [ɬ] / _C[+sibilant]

tʃ

/tʃ/ → [ts]
/tʃ/ → [θ]
/tʃ/ → [t] / #_
/tʃ/ → [dʒ] / #_i
/tʃ/ → [d] / _#
/tʃ/ → [ʃ] / _q

tʃʼ

/tʃʼ/ → [tʃ]

c

/c/ → [s]
/c/ → [t] / _V{s,c}V
/c/ → [t] / _#
/c/ → [s] / #_{a,u}

k

/k/ → [kʷ] / _a
/k/ → [g] / _{w,j}
/k/ → [g] / _l
/k/ → [g] / #_
/k/ → [g] / _(V)r
/k/ → [ʔ] / _#
/k/ → [ʔ]
/k/ → [g] / a_a
/k/ → [ts] / _i
/k, g, x/ → [ts, dz, s] / _{i, æ}
/k/ → [tʃ] / #_Vt
/k/ → [q] / _V[+low]
/k/ → [x] / a_
/k/ → [j] / V_tV
/k/ → [i] / #_n
/k/ → ∅ / _n
/k/ → ∅ / _st
/k/ → ∅ / t_
/k/ → ∅ / _w
/k/ → ∅ / X_X
/k/ → ∅ / V_wV

kl

/kl/ → [ð]

km

/km/ → [kV¹pV¹]

kr

/kr/ → [gj] / medial

/pr, kr/ → [bl, gl]

ks

/ks/ → [ts] / _V

kx

/kx/ → [s]

/kx/ → [xː]

k’

/k’/ → [k]

/k’/ → [q]

/k’/ → [χʷ] / #_

/k’/ → [ɣ]

kː

/kː/ → [q]

kʷ

/kʷ/ → [t] / _{i,e}
/kʷ/ → [b]

/kʷ/ → [p]
/kʷ/ → [k] / #_

/kʷ/ → [q] / _a
/kʷ/ → [w] / X_X

/kʷ/ → [k] / _i
/kʷ/ → [xʷ] / V_V

q

/q/ → [k] / _i
/q/ → [g]
/q/ → [ʔ]

/q/ → ∅ / _i
/q/ → [ʔ] / h_
/q/ → [x]

/q/ → [h]
/q/ → [χ]
/q/ → [ʕ]

q’

/q’/ → [ʔ]

/q’/ → [q]

b

/b/ → [p] / N_
/b/ → [p] / #_
/b/ → [p’] / #_
/b/ → [m] / _j

/b/ → [m] / V_V
/b/ → [m] / #_Vn
/b/ → [m] / #_a
/b/ → [m] / _VN

/b, d, g/ → [m, n, ŋ] / _#
/b/ → [m]
/b/ → [v] / V_
/b/ → [w] / _#

/b/ → [w] / V_V

d

/d/ → [r] / X_X
/d/ → [r]
/d/ → [t] / _w
/d/ → [t] / N_

/d/ → [ð] / r_
/d/ → [k] / _w
/d/ → [ʧ] / _{i,u}
/d/ → [t’] / #_

/d/ → [l] / #_
/d/ → [l] / _V[-nas]
/d/ → [l]
/d/ → [n] / _V[+nas]

/b, d, g/ → [m, n, ŋ] / _#
/d/ → ([ð] →) [j/h] / V_

dʷ

/dʷ/ → [b]

/dʷ/ → [dʒʷ]

dz

/dz/ → [z]
/dz/ → [d] / #_

/dz/ → [c] / N_
/dz/ → [ts’] / #_

/dz/ → [j]
/dz/ → [dʒ]

dɮ

/dɮ/ → [ɣ]

dʒ

/dʒ/ → [g]
/dʒ/ → [ʒ]
/dʒ/ → [d] / #_

/dʒ/ → [j]
/dʒ/ → [j] / #_
/dʒ/ → [dz]

/dʒ/ → [ð] / i_
/dʒ/ → [ð] / u_a
/dʒ/ → ∅ / i_i

g

/g/ → [k] / #d,wV_
/g/ → [k’] / #_
/k, g, x/ → [ts, dz, s] / _{i, æ}
/g/ → [dʒ] / #_Vd

/g/ → [dʒ]
/g/ → [k] / #_
/g/ → [k] / _w
/g/ → [k] / N_

/g/ → [m] / _n
/g/ → [w] / V_
/g/ → [w]
/g/ → [ŋ]

/b, d, g/ → [m, n, ŋ] / _#
/g/ → ∅ / X_X

gʷ

/gʷ/ → [w]

/gʷ/ → [β] / _V

/gʷ/ → ∅ / _o

ɢ

/ɢ/ → [ɣ]

/ɢ/ → [q]

ʔ

/ʔ/ → ∅ / _{p,t,k}

/ʔb ʔd ʔg/ → [p t k]

/ʔr ʔv ʔh/ → [t m {f,s}]

m

/m, n, ŋ/ → [p, t, k] / _#
/m/ → [ʔ] / _p

/m/ → [b] / #_
/m/ → [v] / _V

/m/ → [w] / #_
/m/ → [w] / _B

mb

/mb/ → [mː]

mn

/mn/ → [mV¹nV¹]

/mn/ → [bl]

n

/m, n, ŋ/ → [p, t, k] / _#
/n/ → [l] / CV_m
/n/ → [x] / _t
/n/ → [m] / #_

/n/ → [d] / #_
/n/ → [l] / #_
/n/ → [t] / #_, _k, _kʷ, _h
/n/ → [r]

/n/ → [m] / u_V
/n/ → [h] / V_V
/n/ → ∅ / _j
/n/ → ∅ / _{s,ts}

/n/ → ∅ / _ti
/n/ → ∅ / _s

nb

/nb/ → [m]

nd

/nd/ → [nː]

nr nl

/nl, ml/ → [nd]
/nr/ → [d]

/nr nw/ → [nd mb]
/nr/ → [rː]

/nr/ → [s]

ɲ

/ɲ/ → [n]

ŋ

/m, n, ŋ/ → [p, t, k] / _#
/ŋ/ → [g] / #_u
/ŋ/ → [g] / _$

/ŋ/ → [n] / #_{a,o,e}
/ŋ/ → [j] / #_
/ŋ/ → [h]

/ŋ/ → [x] / _#
/ŋ/ → [x]
/ŋ/ → [n]

/ŋ/ → [w] / C_
/ŋ/ → ∅

ŋg

/ŋg/ → [ŋː]

f

/f/ → [p]
/f/ → [w] / _V

/f/ → [s]
/f/ → [x] / _O

/f/ → [xʷ]

ɸ

/ɸ/ → [w] / V_V

/ɸ/ → [f] / #_

/ɸ/ → [h]

s

/s/ → [x] / _t
/s/ → ∅ / _{p,t,k}
/s/ → ∅ / #_
/s/ → ∅ / V_V
/s/ → [t] / _V{s,c}V
/s/ → [h] / _#

/s/ → [h] / #_
/s/ → ∅ / C_V
/s/ → ∅ / V_C
/s/ → [dʒ] / {i,u}_V
/s/ → [h] / V_V
/s/ → [x] / {r,u,k,i}_

/s/ → [f] / #_r
/s/ → [b] / _r
/s/ → [ʃ] / V_V
/s/ → [ʃ] / _xʷ
/s/ → [ʃ] / _r
/s/ → [θ]

/s/ → [ʃ] / V_k
/s/ → [ɬ]
/s/ → [x] / _l

sk

/sk/ → [f]

/sk/ → [ʃ]

/sk/ → [j]

sn

/sn/ → [tsV¹rV¹]

/sn/ → [sV¹nV¹]

/sn/ → [sː]

sr

/sr/ → [ft]

/sr/ → [sn]

st

/st/ → [ʃ]

/st/ → [ts]

ʂ

/ʂ/ → [h]

/ʂ/ → [s]

/ʂ/ → [ʃ] / _a

θ

/θ/ → [t]
/θ/ → [s]

/θ/ → [ʃ] / C_
/θ/ → [x] / _{p,k}

/θ/ → ∅ / _{p,t,k}
/θ/ → [r]

/θ/ → [f] / #_

ʃ

/ʃ/ → [x] / _B
/ʃ/ → [x] / _k

/ʃ/ → [s] / _{p,k}
/ʃ/ → [tʃ] / #_a

/ʃ/ → [s]
/ʃ/ → [h]

ɕ

/ɕ/ → [s]

x

/x/ → [h] / _{p,k}
/x/ → [ʃ] / _{i,e}

/x/ → [h]
/x/ → [s]

/k, g, x/ → [ts, dz, s] / _{i, æ}
/x/ → [j] / _V

/x/ → [w] / a_o
/x/ → ∅ / _k

xt

/xt/ → [iθ]

xk

/xk/ → [g]

/xk/ → [kː]

xʷ

/xʷ/ → [f ~ () ɸ]
/xʷ/ → [kʷ]

/xʷ/ → [h] / V_V
/xʷ/ → [ħ]

/xʷ/ → [w]

xʲ

/xʲ/ → [ɕ]

/xʲ/ → [ʃ]

/xʲ/ → [s]

χ

/χ/ → [k] / #_

v

/v/ → [p] / _s
/v/ → [p] / _C

/v/ → [b] / _r
/v/ → [f]

/v/ → [ft] / _s
/v/ → [w]

/v/ → [m] / #_V
/v/ → ∅ / u_V

z

/z/ → [ts] / Vj_
/z/ → [d] / Vj_
/z/ → [d] / V_

/z/ → [d]
/z/ → [g]
/z/ → [s]

/z/ → [ð] / X_X
/z/ → [ð]
/z/ → [h]

/z/ → [r]

/z/ → [j]

ʒ

/ʒ/ → [z]

ð

/ð/ → [d]
/ð/ → [z]

/ð/ → [j]
/ð/ → [n] / _#

/ð/ → [t]
/ð/ → [s] / _i

/ð/ → [h]

ʒ

/ʒ/ → [j] / #_

/ʒ/ → [ʃ] / #_

ɣ

/ɣ/ → [k] / #_
/ɣ/ → [k] / _#
/ɣ/ → [k] / _C
/ɣ/ → [g] / #_Vx

/ɣ/ → [q]
/ɣ/ → [x] / #_
/ɣ/ → [ʁ]
/ɣ/ → [j]

/ɣ/ → [w] / V_
/ɣ/ → [i] / C_V
/ɣ/ → ∅ / V_V
/ɣ/ → ∅ / V_u

/ɣ/ → ∅ / u_V

ɣʲ

/ɣʲ/ → [ʒ]

ʁ

/ʁ/ → [q] / #_

/ʁ/ → [q] / _#

/ʁ/ → [q] / _C

h

/h/ → [ɦ] / #_Vs
/h/ → ∅ / _#
/h/ → [x] / _{p,k}
/h/ → ∅ / _{s,x}

/h/ → ∅ / C_
/h/ → ∅ / _{p,t,k}
/h/ → [ʔ] / _#
/h/ → [ʔ] / V_V

/h/ → [j]
/h/ → [j] / _{i,e}
/h/ → [w] / _{o,u}
/h/ → [j] / _#

/h/ → [w] / #_

l

/l/ → [d] / #_VC
/l/ → [n] / #_VN
/l/ → [n] / V_#
/l/ → [n] / _#
/l/ → [n]
/l/ → ∅ / {i,e}_#

/l/ → [j] / {a,o,u}_#
/l/ → [j] / #_iC
/l/ → [j]
/l/ → [d] / _{C,#}
/l/ → [r]
/l/ → [r] / _#

/l/ → [ɬ] / _t
/l/ → [ɮ]
/l/ → [l̥] / V_#
/l/ → [d] / n_
/l/ → [ɾ] / V_V
/l/ → [r] / #nV_C

/l/ → [t]
/l/ → [h] / _m
/l/ → [l̥]

ln

/ln/ → [lː]

lw

/lw/ → [lgʷ ~ lg]

ɣ

/ɣ/ → [ʒ] / _{a,u}

ɭ

/ɭ/ → [ɖ] / a_

ɬ

/ɬ/ → [l]
/ɬ/ → [s] / #_VC

/ɬ/ → [ʃ]
/ɬ/ → [xʲ]

/ɬ/ → [s]

ɬʷ

/ɬʷ/ → [ʃ]

r

/r/ → [t] / _{i,u}
/r/ → [t] / _#
/r/ → [t]
/r/ → [d] / s_
/r/ → [d] / x_
/r/ → [d] / k_

/r/ → [n] / #_
/r/ → [h] / C_
/r/ → [s] / _k
/r/ → [n]
/r/ → [ʃ] / _O
/r/ → [l]

/r/ → [h] / _#
/r/ → [ɣ]
/r/ → [n] / _C
/r/ → [d] / #_
/r/ → [j]
/r/ → [ʐ] / #_

/r/ → [ts] / _i
/r/ → [n] / _j
/r/ → [w] / o_{a,e,i}
/r/ → [j] / e,i_{a,o}
/r/ → [h] / ʃ_
/r/ → [h] / s_

/r/ → [l] / x_
/r/ → ∅ / _j
/r/ → ∅ / w_
/r/ → ∅ / Vh_
/r/ → ∅ / k_
/r/ → ∅ / n_
/r/ → ∅ / Xk_X

ɾ

/ɾ/ → [g]
/ɾ/ → [n] / _#
/ɾ/ → ∅ / V_V

rn

/rn/ → [ɳ]
/rn/ → [rː]

rd

/rd/ → [r]
/rd/ → [l]

w

/w/ → [p] / _t
/w/ → [g] / _V
/w/ → [kʷ] / _V
/w/ → [g]
/w/ → [b] / i_#
/w/ → [gʷ]
/w/ → [v]
/w/ → [u] / t_
/w/ → [f] / #_
/w/ → [β]
/w/ → [m]
/w/ → [m] / _u
/w/ → [m] / u_
/w/ → [m] / _Vn
/w/ → [j]
/w/ → [j] / {r,k}_i#
/w/ → [j] / V_{o,u}
/w/ → ∅ / C_
/w/ → ∅ / {a,o}_i
/w/ → ∅ / {t,r}_i
/w/ → ∅ / t_i
/w/ → ∅ / _o
/w/ → ∅ / k_a
/w/ → ∅ / #_
/w/ → ∅ / _C
/w/ → ∅ / #_i

j

/j/ → [s]
/j/ → [dʒ] / #_
/j/ → [d] / #_
/j/ → [t] / #_
/j/ → [dz]
/j/ → [t] / _#
/j/ → [d]
/j/ → [r]
/j/ → [ð] / V_V
/j/ → [ð] / #_
/j/ → [ʒ] / {#, V, ʔ}_V
/j/ → [w] / _o
/j/ → [dʒ] / _V
/j/ → [tʃ] / _V
/j/ → [dʒ] / _V[+high]
/j/ → [d] / _V[+low]
/j/ → [h] / #_
/j/ → ∅ / C_
/j/ → ∅ / _ts
/j/ → ∅ / {tʃ,ʃ}_e
/j/ → ∅ / {ts,ʃ,w}_a
/j/ → ∅ / _e

a

/a/ → [o] / _w
/a/ → [ə] / _Ce
/a/ → [i] / _Ci
/a/ → [u] / _Cu
/a/ → [i] / _Ce
/a/ → [e] / _Ci
/a/ → [e] / _Co
/a/ → [e] / _re
/a/ → [o] / _Cu
/a/ → [ə]
/a/ → [aː] / _C#
/a/ → [o] / _#
/a/ → [ɛ] / {t,n}_#
/a/ → [e] / _{t,s}#
/a/ → [æ] / _Ce
/a/ → [æ] / iC_
/a/ → [æ] / _ra
/a/ → [æ] / r_

aː

/aː/ → [ɔa]
/aː/ → [ua]

æ

/æ/ → [i] / {s,ʃ,x}_Ca, _Ci
/æ/ → [e]
/æ/ → [e] / _j
/æ/ → [a]
/æ/ → [ɑ] / unstressed
/æ/ → [ia] / _Ca

ai

/ai/ → [ɛː]

/ai/ → [ɛ]

au

/au/ → [ɔː]

/au/ → [ɔ]

aw

/aw/ → [o] / _#

wa

/wa/ → [u] / #_

/wa/ → [o] / #_

wə

/wə/ → [u] / #_

e

/e/ → [i] / #N_C
/e/ → [i] / #b_
/e/ → [i] / #_
/e/ → [i] / k_
/e/ → [ja] / #_C{e,i}

/e/ → [i] / _g
/e/ → [i] / #C_C
/e/ → [i] / _Ci,u
/e/ → [æ] / {P,K}_
/e/ → [æ] / _P

/e/ → [a] / _{r,l}
/e/ → [a] / _s
/e/ → [i] / _ŋ
/e/ → [o] / _w
/e/ → [o] / w_

/e/ → [i] / r_ʔ
/e/ → [je] / _#

eː

/eː/ → [æ]

/eː/ → [i]

/eː/ → [ie]

/eː, oː/ → [iːe, uːo]

eʊ

/eʊ/ → [egu]

ew

/ew/ → [ju]

ej

/ej/ → [i]

/ej/ → [e] / _#

we

/we/ → [o]

ɛ

/ɛ/ → [i] / _N

/ɛ/ → [ie] /

i

/i/ → [e] / {r,l}_e
/i/ → [e] / _C{C,#}
/i/ → [əj] / _V
/i/ → [e] / _k#
/i/ → [iə] / _{k,ŋ,h,l,r}#
/i/ → [e] / _Ca
/i/ → ∅ / #_{z,dʒ}
/i/ → [e] / tʃ_

in

/in/ → [ŋ] / #_C

iw

/iw/ → [i] / _#
/iw/ → [u]

ij

/ij/ → [ej] / _C

ə

/ə/ → [i]
/ə/ → [o] / _h{p,w,m}
/ə/ → [i] / _hC
/ə/ → [a] / _x
/ə/ → [u]
/ə/ → [a] / _ʔ#
/ə/ → [a] / _C#
/ə/ → [i] / {u,a}_
/ə/ → [o]

o

/o/ → [e]
/o/ → [wo]
/o/ → [u] / _C{i,u}
/o/ → [u] / _N
/o/ → [u] / _ŋ
/o/ → [e] / w_{r,s,t}
/o/ → [u] / _Ci
/o/ → [wo] / _#

ow

/ow/ → [u]
/ow/ → [o] / _#

oː

/oː/ → [aː]
/oː/ → [uː]
/oː/ → [u]
/oː/ → [ua]
/oː/ → [wo]
/eː, oː/ → [iːe, uːo]

ø

/ø y/ → [e i]
/ø y/ → [we wi]
/ø/ → [i] / {s,ʃ,x}_Cu
/øː/ → [wi] / {k,g}_
/ø y/ → [{o,u} i] / _Ci

u

/u/ → [i]
/u/ → [o] / _C{C,#}
/u/ → [əw] / _V
/u/ → [aw]
/u/ → [o] / _k#
/u/ → [uj] / _{t,s}#
/u/ → [uə] / _{k,ŋ,h,l,r}#
/u/ → [o] / _Ca
/u/ → [o] / _{p,b}i
/u/ → [wə] / #_
/u/ → ∅ / k_w

uw

/uw/ → [ow] / _C

y

/ø y/ → [e i]
/ø y/ → [we wi]

/y(ː)/ → [i(ː)]
/y(ː)/ → [e(ː)]

/y/ → [ə] / _r
/ø y/ → [{o,u} i] / _Ci

uwa

/uwa/ → [oː]

∅

∅ → t / #_s
∅ → t / s_r
∅ → k / s_w
∅ → ʔ / #_V

∅ → h / #_V
∅ → s / k_i
∅ → j / {i, e}_V
∅ → j / #_a

∅ → w / o_V
∅ → e / {ts, s, n}_r, n_k
∅ → e / {t, k}_{r, n}, s_w
∅ → e / {t, s}_w

∅ → i / r_w
∅ → i / C_R

16. Sound Systems

16.1. Chipaya. Onset clusters: /s ʃ/ + /p/ + (/x/); /s ʃ/ + /k q/ + (/x xʷ χ χʷ/); /t/ + /x xʷ χ χʷ/; /tʃ l/ + /x/. Possible codas: /x χ/ + /p t k q l r/ + (/t/); /xʷ χʷ/ + /k q/ + (/t/); C + /t/.

16.2. Assiniboine. Onset clusters: /p/ + /t s ʃ tʃ/; /tk/; /k/ + /t s ʃ tʃ m n/; /s ʃ/ + /p t k tʃ m n/; /x/ + /p t tʃ m n/; /mn/. No codas.

16.3. Oksampmin. (C)(C)V(C). Onset choice: any C except /w j/ or labialized stop. Onset cluster: C_2 may be /j w l x/. Onset /sk/ also seen. All C may be final except prenasalized stops.

16.4. Pech. (C)(C)V(C)(C). Onset cluster C_2 must be /r/. Final single: all C execpt /p t kʷ b/. Final cluster must have /r/ as first element.

16.4.1. Ingessana. (C)V(C)(C). All C in simple codas except /ʔ kʼ dʒ/. Coda cluster word finally only, C_1 is /r l n m/ and C_2 is obstruent.

16.4.2. Aguacatenango Tzeltal. (C)(C)V(C)(C). Onset cluster C_1 must be /s ʃ h/, with any consonant following. Coda cluster limited to /h/ + voiceless stop or affricate.

16.4.3. Mamaindê. (C)(C)V(C)(C). Onset cluster: /kʰ tʰ k h/ + /w/, or /h ʔ/ + /l n j w s/, or /ʔm/. Stops and nasals only for simple coda, with stop or nasal + /ʔ/ for coda cluster.

16.4.4. Qawasqar. (C)(C)(C)(C)V(C)(C)(C). Biconsonantal onset: /f q qʼ s t/ + /tʃ s t tʼ j w q/. Triconsonantal onset have same restrictions for C_1 (/qsq qst sqw/). All simple codas allowed. Complex coda: /f j l m n p q r t w s/ + /s q/. Triconsonantal codas include /lqs rqs qsq/.

16.4.5. Bardi. (C)V(C)(C). Single coda all except /p/. Coda cluster is /l r ɻ/ + nasal homorganic with following stop.

16.4.6. Burushaski. (C)(C)V(C)(C). Complex onset: /p b pʰ t d tʰ g/ + /ɾ j/. All C except /w j/ in simple coda. Coda clusters: voiceless fricative + /k/; or sonorant + /t k ʂ ɕ ts tɕ ʈʂ/.