Makhtl

Consonants—the skeleton of language. We're somewhat pampered, in this regard; we're used to having our hands held when it comes to reading, with our vowels and our spaces, but if we look at a few, more osseous languages (and many still living),[1] we can find systems which operate without the guide rails we're used to: languages which only write consonants or, better yet, only write some consonants, using semantic rather than phonetic components to indicate meaning in a way that is more akin to 漢字 (Hànzì) than anything we're probably familiar with.[2] This essay will be broken down into a few sections: a description of the various stages of the Egyptian language and the scripts used to write them, an explanation of triconsonantal root systems and how they complement consonant-centric writing systems, and finally (the real meat and potatoes) an exploration of both these topics with an example conlang.

Having been advised to [shoot high, aim low](#) in explanations, we'll be building up our model of these systems from a relatively low level—a measure intended to ensure we don't leave any gaps and reach as broad an audience as possible. With that said, let's begin!

---

[1] The most frequent examples being the writing systems used for or derived from those used for the Semitic languages such as Arabic, Hebrew, Aramaic, and Pahlavi. I eventually want to explore this last example in greater depth, perhaps in some other post, as it is a fascinating example of a writing system developing logograms with a totally different underlying mechanism than the other major logographies: Chinese, Egyptian, Mayan, etc.

[2] I will hopefully create another post soon exploring the logosyllabic writing system—the type that includes the Chinese and Mayan scripts—as these operate a bit differently than those we'll be talking about today.

1 | AN BRIEF OVERVIEW

Most people will probably associate the Egyptian language with hieroglyphs, and while we will be discussing them to some degree, we are especially interested in the later systems that evolved out of them: Hieratic and Demotic. These began as cursive forms of hieroglyphs, but they developed a great many quirks of their own as time went on. In descending order of formality (with plenty of exceptions) we have Hieroglyphs, Hieratic, and finally Demotic. Coptic, the modern form of the Egyptian language, uses an alphabet largely adopted from Greek, a few modifications made here and there; we'll mostly leave it for another time, given that it is (mostly) not directly descended from hieroglyphs, but we will touch on some of the phonological and grammatical features of the Coptic language, so it is not entirely beyond the scope of this essay.

Ancient Egyptian is an Afro-Asiatic language, related—albeit somewhat distantly—to the Semitic, Berber, and Cushitic languages (among others) of North Africa and the Middle East. One might understand its relationship to these as being similar to Greek's relationship to the Romance and Germanic languages: it bears certain similarities due to having descended from the same source but is, in many ways, quite different. The situation is only made more complex by the changes the language underwent throughout history, becoming in some ways less akin to those Semitic languages you're likely to be familiar with. To explain those differences, we'll have to touch on three (of the) classifications languages can fall into: analytic, fusional, and agglutinative.

Analytic languages are those like Chinese, Indonesian, or Vietnamese which largely don't inflect their words (meaning nouns don't change their form or take endings for case or possession, verbs don't do so for tense, etc). English is somewhat analytic, but it holds onto some older endings (we still conjugate our verbs and inflect our pronouns, for example), so it doesn't fit neatly into this category. Really, it straddles the line with our next classification: fusional languages.

These are languages like Spanish or Russian which have endings that contain a lot of meaning. Take the verb "hablar" in Spanish; this means "to speak." It has a form "hablo," where the ending "-o" indicates that the subject is the first-person singular "I," that the verb is in the present tense, and that the mood is the indicative (what is being said is true or believed by the speaker). In contrast, the form "habláremos" indicates a first-person, plural subject, the future tense, and the subjunctive mood (what is being said is not true, not believed, or uncertain). As you can tell, both these forms have a lot of meaning packed into relatively small endings—in other words, these meanings are "fused" into one ending.

Both English and a great many of the Semitic languages fit somewhere between the fusional and analytic categories, and for much of its history Egyptian was in a similar situation. But Coptic, the last stage of the language (that is still used by the Coptic Church) underwent changes that moved it into our last category of languages: those which separate the complex affixes of fusional languages into numerous ones that are compounded onto one another.

Agglutinative languages include such languages as Japanese, Turkish, Finnish, and most importantly for our purposes, Coptic. These are characterized by their use of numerous affixes—prefixes or suffixes, less often infixes or circumfixes—that compound to form more specific meanings. For example, we might look at the Japanese word, 思う, /o.mo.u/, "to think." To make it polite, we use a different base and add the polite ending "-ます" /-masu/, and to making it both polite and negative, we add on the negative ending, yielding: 思いません /omo.i.ma.sen/. Each suffix contributes a small bit of meaning, and they're added on, one after the other, to stack meaning onto a verb.

Some languages, such as Turkish, have similar complexity in their nouns. For example, the Turkish word "evimde" is a combination of the word "ev," or "house," with the first-person, singular possessive suffix (indicating "my") and the locative case (indicating "at"). As a whole, it means "at my house," with the underlying structure "ev-im-de," or "house-my-at." The order of these components might be a bit confusing, but it will suffice for now to say that both Japanese and Turkish are examples of languages that are considered "head-final," a concept we'll explain shortly.

2 | CHARACTERISTICS OF AFROASIATIC LANGUAGES

Now that we've covered the various ways languages encode information, it will hopefully make more sense now when I say that Middle Egyptian was fusional but Coptic is agglutinative.

While I imagine this theory does not hold up when compared to actual historical changes, as language creators it may serve us well to imagine such changes as cyclical: agglutinative languages undergo changes which fuses their many affixes together; fusional languages have these affixes worn away entirely (or speakers simplify inflections because, as humans, we're quite lazy); and analytic languages begin affixing their particles, auxiliary verbs, and other periphrastic constructions to the words they modify. Egyptian underwent something like this, going from fusional to analytic to agglutinative, and this will inform our attempts to emulate its historical development in our own constructed language. It is often the case that old structures in language linger long after new methods have arisen to mark the same information they encoded—or when this information ceases to be overtly marked. For example, Old English had an alternative method for marking plurality in "weak" nouns that has been lost in all but a few modern words such as "brethren" (the plural of "brother"). The term has shifted in meaning since the default method for marking plurality became "-s," but that in itself gives us good insight into how a language might grow and change, lexicalizing its old methods for marking information when new methods arise.

To stick with kinship terms, we might imagine a hypothetical word ⟨ma⟩ /ma/, or "mother," which is "pluralized" via a noun class system as in Zulu. Noun classes are a complicated subject, but for now it will suffice to say that each noun class has a prefix and plurality is indicated not via the method you're probably used to but instead by changing the noun's class: that is, if a noun is a member of

the class for "person," it is made plural by removing the class prefix and adding in its place the class prefix for "people." Let's say, for now, that the class prefix for "person" is ⟨a-⟩ /a/, so a singular "mother" would be ⟨ama⟩ /ama/, and if I wanted to make that into "mothers" I would need to change that prefix to the class prefix for "people," let's say ⟨a'a-⟩ /aʔa/, which would yield ⟨a'ama⟩ /aʔama/. Thus, "mother" is ⟨ama⟩ /ama/ while "mothers" is ⟨a'ama⟩ /aʔama/.

Some years pass and the language loses its noun class system: either speakers begin to see them as part of the root or they simply fall away entirely (some mixture of the two appears to have occurred in Akan).[3] In addition, speakers began using a regular suffix to indicate plurality, let's say ⟨-n⟩ /n/. But speakers keep using ⟨a'ama⟩ /aʔama/ despite the fact that the new word ⟨aman⟩ also means "mothers." There are simply two, equally legitimate ways to pluralize ⟨ama⟩. This continues for a time, until ⟨aman⟩ wins out, but speakers don't just drop ⟨a'ama⟩ entirely. Instead, it shifts meaning; it begins to be used to talk about one's metaphorical mothers, one's maternal ancestors. Many more years pass, and now speakers use ⟨aman⟩ as the plural form of ⟨ama⟩, but they also refer to their maternal ancestors as ⟨a'ama⟩, even though that word used to have the same meaning that ⟨aman⟩ holds now. This sort of thing is a perfect way to derive new vocabulary while weaving complex etymology and grammar into your langauge, developing a sense that the language is living and growing.

Something we also often see, if we're sticking with plurality, is the development of parallel methods of pluralization. The Afroasiatic languages,

3 Emmanuel Kweku Osam (1993): The loss of the noun class system in Akan, *Acta Linguistica Hafniensia: International Journal of Linguistics*, 26:1, 81-106.

particularly the Semitic ones, are known for their "broken plurals," or nouns which don't take affixes to mark plurality but instead change the vowels inside their roots. I've heard a few explanations as to their origins, and not having any particular expertise on the topic I'll simply say that we'll derive ours from an old noun class system of the sort we just talked about, even if that's not how they arose in the Semitic languages.

These broken plurals are (often) part of a broader system known as triconsonantal roots. You might've heard of them—and if you are a conlanger, you almost certainly have, and you're probably tired of it—but in any case, most people are aware of them due to Hebrew or Arabic. In these systems, most words are derived from three-consonant roots such as the classic, ك-ت-ب, K-T-B. Both nouns and verbs are formed by adding affixes to this root and changing the vowels that get placed between its three consonants. The root, K-T-B, has to do with writing, so if we add three a's after each consonant, it yields كتب /kataba/, which means "he wrote." We can say "I write" by changing those vowels, yielding أكتب /aktubu/.

Middle Egyptian worked similarly, but it seems to have lacked the broken plurals that we find in Arabic. In his book on Middle Egyptian, James P. Allen tells us that the standard pluralization rule—suffixing "-w"—is "absolutely consistent in Egyptian: all nouns form their plurals by it, without exception."[4] During the bulk of this essay, we'll be making a little example language that'll feature some of the systems we've explored, specifically those from Coptic and

---

[4] Allen, Japes P. *Middle Egyptian: An Introduction to the Language and Culture of Hieroglyphs*. 2nd ed. (New York: Cambridge University Press, 2010), 38.

Middle Egyptian, but I simply enjoy broken plurals too much not to include them as well.

But before I get ahead of myself, we should talk about logoconsontal writing systems, the real meat and potatoes that I originally began this essay intending to explore.

3 | HIEROGLYPHS, HIERATIC, AND DEMOTIC

For now, we'll set Hieratic and Demotic aside and focus on Hieroglyphs, the thing most people imagine when you say "Ancient Egyptian." Hieroglyphs function on a hybrid system that is in some ways akin to Arabic or Hebrew and in others more like Chinese characters. Broadly speaking, only consonants are written, and no spaces are placed between words. This may seem complicated, but it actually isn't too much of a hindrance. For example, youcanprobablyunderstandthiswellenough,[5] and y cn prbbl ndrstnd ths wll ngh s wll.[6] The real trouble comes when Istrtwrtngnlywthcnsnntsndwthtspcs.[7] To clear this up, Hieroglyphs have a feature that may be familiar to you if you've ever studied Chinese characters (or some related system). In Egyptology, this feature is known as the determinative, though you may know it by the term used to describe it in Chinese characters: the radical. In short, this is a component that appears next to a string of consonants and tells you something about their meaning: it is a semantic indicator rather than a phonetic one. To emulate its function, imagine you are reading a sentence and you see a string of consonants: "thrn." Next to it,

---

[5] "You can probably understand this well enough…"
[6] "You can probably understand this well enough as well."
[7] "I start writing only with consonants and without spaces."

there is a determinative that looks like a king; this tells you that "thrn" has to do with royalty in some way. Putting this together, you can come to the conclusion that a word containing "thrn" and having to do with royalty must be "throne." If the determinative told you the meaning had to do with "plants," then you would know that the word was "thorn" instead.

For another example, the consonants "lt" could indicate "late," "lit," or "lot," so you would rely on the determinative to specify. The majority of Chinese characters work the same way, with one component that tells you what the word rhymes with and a "radical" that tells you something about its meaning. Hieroglyphs are much the same, the only (pertinent) difference being that they lack vowels.

Of course, the Afroasiatic languages tend not to be very vowel heavy—(Modern Standard) Arabic only features three, /a i u/, with the added dimension of vowel length—so the amount of possible syllables that one has to consider when reading is not nearly as great as in some Indo-European languages, such as English or Danish (which has just an obscene amount). That doesn't make it much easier on language learners, of course, but writing systems are designed for those who already speak the language with little consideration given to those who are trying to learn it.

Hieroglyphs also feature purely semantic glyphs which are either abstract or literal depictions of the word they are referencing. The hieroglyph '   ' means "eye," a literal representation of an eye, albeit simplified slightly. Conversely, some words lack a determinative, written entirely using phonetic components.

From Hieroglyphs, there developed a cursive system known as Hieratic which, in turn, grew into Demotic. All three systems were eventually used concurrently for different purposes: Hieroglyphs for the classic engravings you're familiar with, Hieratic for religious texts, and Demotic for more mundane things. This is an oversimplification, but it will suffice.

I particularly like the look of Hieratic and Demotic, so in our next section we'll explore how one might design a language whose writing system is based off of them.

4 | DIACHRONICA

Now that we've touched on all the prerequisites, I can begin really talking about how one might go about making a language which includes these features.

However, a tutorial on how to make triconsonantal root systems would require far too much time, and others have already done a better job of it than I possibly can, so instead of rehashing what they've said here, I'll simply link you to a few of those I've found particularly helpful:

- [This thread on the old Zompist board](#) is a good starting point.
- Biblaridion has [a video on Nonconcatenative Morphology](#).
- You should probably also read [David Peterson's lovely breakdown](#) of his first constructed language, Megdevi, to learn what *not* to do if you're trying to make a realistic triconsonantal root language.

- There's [another (quite sparse) post by Jörg Rhiemeier](#) on some of the complexities and pitfalls of triconsonantal root systems.

- And, if I'm allowed my own two cents (it's my blog, you can't stop me), I'd recommend learning about [umlaut](#), [ablaut](#), [metathesis](#), and [metaphony](#) as these are among the various methods on can use to create the characteristic nonconcatenative morphology of triconsonantal systems. However, it is also useful (perhaps even necessary, in my experience) to make use of [paradigm levelling](#), [vowel reduction](#), and [shifts in stress and subsequently vowel length](#) in order to create even vaguely naturalistic morphology. I won't claim to be an expert on the subject; I'm bound to make some mistakes; but I think knowing about these features and processes makes one better equipped to take a crack at it.

Now that that's been said, we can begin to flesh out a sketch of a language that features those systems we've explored today. I've had an itch to create a language influenced by Armenian or Kurdish, and having recently read quite a bit on Ossetian, I think we'll draw on it as well.[8] Now, I know that all three languages I've just mentioned are Indo-European, but bear with me—the Semitic influences will soon be evident. So, phonologically, we'll be creating a little language that looks something like this:

---

[8] In retrospect, the phonology of our language has broad similarities to these languages, but its phonotactics give it a much different sound, perhaps even more akin to Yiddish (due to syllabic consonants) or Moroccan Darija (due to its consonant clusters), at times.

Consonants		Labial	Alveolar	Palatal	Velar	Uvular	Glottal
Nasal		m	n				
Stop	voiceless	p	t		k	q	ʔ
	voiced	b	d		g		
Affricate	voiceless		t͡s	t͡ʃ			
	voiced		d͡z	d͡ʒ			
Fricative	voiceless	f	s	ʃ	x ~ χ		h
	voiced		z	ʒ	ɣ ~ ʁ		
Approximant		v ~ ʋ	l	j			
Tap			ɾ				
Trill			r				

An explanation is probably in order: the basic consonants aren't anything particularly controversial, but it is important we discuss the origin of some of these beforehand.

Basically, the language will have once had an ejective series (derived from stops and affricates preceding the glottal stop) but it will have merged with the voiceless series by the time of the modern language. The distinction will still be reflected in writing, and it will contribute to the phonemization of the non-sibilant fricatives which we'll derive from lenited stops.[9] The affricates will result from palatalization; most the fricatives (other than the sibilants), from the aforementioned lenition of stops; and the trill will simply have arisen from /ɾ.ɾ/

---

[9] Phonemization is when a consonant (or any sound) comes to form minimal pairs (words or morphemes that differ only in one phoneme) with other consonants. To clarify, something like Japanese features the sound [t͡s], but it only ever occurs when /t/ comes before /u/, so it is considered an allophone of /t/.

sequences.[10] It's probably easier to show you all this change occurring, but before I describe the proto-language, I should touch on the modern language's vowels.

Vowels						
	Front		Central		Back	
	short	long	short	long	short	long
High	i	iː			u	uː
Mid	ɛ	eː	ə		ɔ	oː
Low				aː		

Again, this isn't anything groundbreaking: it's just a basic five vowel system with the addition of vowel length and an /a/ that has merged with the schwa, itself having arisen arisen from vowel reduction and epenthesis (the insertion of vowels to break up hard-to-pronounce sequences). I debated dropping the vowel length in some later stage of the language, but that would mean losing some important distinctions, so we'll hold onto it for now.

The language will feature a few syllabic consonants as well:

Syllabic Consonants			
Nasals		Liquids	
m̩	n̩	l̩	r̩

---

[10] I should also add that "sibilants" are, in our case, the sounds /s z ʃ ʒ/. These are often treated a little differently than the other fricatives, due to their relatively high intensity.

I'd briefly considered including some syllable sibilants, but that might be pushing it, so we'll leave that for some later project. I suppose it'd be fine: I'm not putting any syllabic stops, so we're not going *really* out-there with this phonology. Anyways, here's the proto-language:

Consonants		Labial	Alveolar	Palatal	Velar	Glottal
Nasal		m	n			
Stop	voiceless	p	t		k	ʔ
	voiced	b	d		g	
Fricative			s			h
Approximant		w	l	j		
Tap			ɾ			

I had originally planned on having the proto-language feature voiceless, aspirated, and voiced series of stops, with the latter two leniting to form the modern language's fricatives—as occurred in the evolution of Modern Greek—but given that this is supposed to be more reminiscent of the Semitic languages, I decided to follow Hebrew's evolution instead. Its non-sibilant fricatives evolved from postvocalic spirantization which was rendered phonemic by future sound changes: that is to say, after vowels, stops became fricatives, but the position of these fricatives was perfectly predictable, only becoming unpredictable (and thus phonemic, rather than allophonic) due to future sound changes.

As for the proto-language's vowels, it will only have three:

Vowels		
	Front	*Back*
High	i	u
Low	a	

When it comes to triconsonantal roots systems, the fewer vowels you have the easier your job becomes.[11] You could conceivably create a language that had both triconsonantal roots and a vowel system reminiscent of some of the more eccentric Germanic languages, but it would be obscenely difficult to keep it naturalistic. I haven't seen it done (well), so I won't attempt it here.

The bread and butter of the triconsonantal root systems I'm aware of is stress: it really makes or breaks the evolution of such systems. Because I know more about it than any of the other Semitic languages, we're going to borrow (and extend) some of the historical changes that Biblical Hebrew underwent in order to derive our own system.

In addition, we should talk about an oft neglected facet of phonology: phonotactics. Conlangers will often have their languages feature CVN syllable structures—in the vein of Japanese—or will simply wave their hands when it comes to this subject. I'm kinda a sucker for it though. What syllables a language allows is not entirely arbitrary, and there are certain through lines across various language families, but you have a good amount of freedom to determine the phonaesthetics of your language in the particular syllables it allows or disallows (and the sequences of syllables, as well).

---

[11] At least, the fewer vowels you start with, the easier it is.

The Afroasiatic languages, particularly the languages of Morocco, can get fairly wild with their syllables, and one of our influencing languages, Armenian, is similarly complex. Much like with phonology, it will be easier if our proto-language features a simple syllable structure, but we can get pretty wild with our modern one. I'm going to say, for simplicity's sake, that the proto-language only allows CVN syllables, where the onset (the first consonant) can be any of the language's consonants and the vowel can be any of its (three) vowels, but the only coda that is allowed is /n/.

I should touch on the sonority hierarchy before we discuss the modern language's syllable structure. In essence, every consonant has a certain sonority—a loudness or resonance—that tends to determine which consonants it can appear in clusters with. Broadly speaking, languages tend to like to structure their consonant clusters so that the most sonorant consonants are closer to the nucleus (the center of the syllable, usually the vowel). Thus, a language would probably prefer to feature syllables like /krast/ rather than /rkast/ or /krats/. Of course, as (I'm assuming) English speakers, we're probably not uncomfortable with /krats/; it's not that far from "cats," which seems like a perfectly reasonable syllable. Plenty of languages allow such exceptions, and plenty more allow some fairly wild consonant clusters.

If we stick with languages with triconsonantal roots, Moroccan Arabic (or Darija) has been influenced by the local Amazigh (Berber) languages and dropped quite a few of the short vowels that were present in Classical Arabic, resulting in

some wild consonant clusters.[12] For an example of a Berber cluster, we can look at Tachelhit's /lktab/. Similarly, one of our stated influences, Armenian, has some wild clusters so I'd like to have them in our language as well.[13][14] Given that we have the schwa, we could reasonably say that all our language's short vowels collapsed into the schwa and then were lost entirely, save in a few positions, and all the old long vowels shortened to form the modern vowel system. That seems plausible enough, and it would help simplify paradigms quite a bit, making deriving the triconsonantal root system that much easier.

To wrap back around to syllable structure, we've already said that the proto-language will feature a CVN structure, so we'll go ahead and finalize our modern language's structure and say that it will have a CCCCVCCC structure. Frightening, I know, but it really isn't that much more complex than English's. Broadly speaking, all consonant clusters will feature rising sonority—that is, each consonant will be more sonorant than the last—though two consonants of equal sonority will be allowed in sequence. This means we can have /ktɾa/ but not /pkta/. If the latter of these two consonants is a stop and this cluster is in the coda, the stop will have no audible release. However, geminate stops are not allowed in the coda; to remedy this, an /ɐ/ will be inserted after them.

---

[12] Speaking of which, you should definitely check out the band Tinariwen, a Tuareg group from Mali who have some pretty solid songs. I'd recommend Nànnuflày or Sastanàqqàm. I might as well also add Bombino, another Tuareg singer-songwriter, whose song Mahegagh I particularly enjoy. If you can't tell, I have a slight addiction to Tuareg music.

[13] Armenian syllable structure seems like a contentious topic, so I'll steer away from it, but I will say that the mountains really must do something to a person. Both the Caucasian languages and the Indo-European languages that ended up in the Caucasus (seem to) feature some wild consonant clusters, and if we look at Tibetan or Quechua they feature their own complexities. Maybe it's the air, but it sure seems like languages in the mountains tend toward some extreme.

[14] In retrospect, this comment on Armenian may make it seem like I support the notion that geography affects phonology, which to my knowledge is broadly rejected by linguistics.

Furthermore, drawing a little on Ossetian, we'll allow sibilants, in our case /s z ʃ ʒ/ to ignore the general sonority rule, appearing anywhere in a cluster, though they (along with the sonorants, /m n ʋ l j r ɾ/) will obey a rule dictating that two of them cannot appear one after the other. Thus, /skʃa/ is allowed, but /sʃka/ is not. Various historical gaps will arise in which consonants appear next to another, just as a function of the language's evolution, but we'll leave determining those as an exercise for the reader. This should wrap up our discussion of the modern language's phonotactics, and we can finally get around to talking about the sound changes that occur between the proto-language and the modern one.

I've already outlined some of them, but it won't hurt to lay them out chronologically so that we can trace how hypothetical words might change over time. I'll try to provide an explanation and an example for each so that it's a little easier to digest. Our first change will be this one:

1. The shift of stress to a universally penultimate position.

This isn't really a sound change proper; it could reasonably be the default state of the proto-language, but I figure it'll help if we write it down here. This rule means that /kataba/ would be stressed on the /ta/ syllable: /ka.ˈta.ba/ or /ka.tá.ba/. We'll mark stress via the apostrophe as, to my knowledge, it is the standard method. This stress shift won't become immediately important, but it is—in my experience—a pivotal part of the evolution of many realistic triconsonantal root

systems. It ensures that later, when we lengthen and reduce certain vowels, there is a rhyme or reason to it that makes derivation patterns more predictable.

Our second change will be the palatalization of stops and sibilants, yielding affricates and palatal sibilants. There is a kinda interesting reason why I want to do this now and not later, but it requires I give you a glimpse into some sound changes to come. In essence, we are going to have two phases of vowel reduction: the first will drop certain vowels between consonants that are themselves intervocalic, so /sakatana/ will become /saktana/. Then, we'll have postvocalic spirantization occur, meaning /saktana/ will become /saxtana/. Then, we'll do some shenanigans with vowel lengthening and dropping and whatnot, but that's not immediately important. What is important though is that I want to palatalize now, before some vowels are dropped, so that the resulting consonants are phonemic rather than allophonic. I've mentioned this distinction before, but basically phonemes are independent consonants or vowels that don't appear in predictable positions. For example, /k/ can appear in any syllable, /ka ki ku/. However, palatalization will only occur before /i/, meaning we will always be able to predict where the resulting sounds will occur. If /k/ becomes /t͡ʃ/ before /i/, then /t͡ʃ/ is an allophone of /k/, meaning we know that whenever /k/ appears before /i/, it becomes /t͡ʃ/. However, if that /i/ were to disappear later, then /t͡ʃ/ would no longer be predictable; in other words, it would become phonemic. For example, we might have two words /kataba/ and /kitaba/. After palatalization, these would be /kataba/ and /t͡ʃitaba/. However, if that first vowel were dropped, we would get /ktaba/ and /t͡ʃtaba/. One can no longer predict where /t͡ʃ/ will show up based on

the modern form of the words alone. Of course, if one knew the etymology of the word, then one could, but that's not what determines whether a consonant is phonemic or not (though it will be important for deriving words from roots, later on).

I should plop down the rule before I get carried away:

2. /t d k g s z/ become /t͡s d͡z t͡ʃ d͡ʒ ʃ ʒ/ before /i/.

This is a pretty standard example of palatalization, and I've already explained it quite a bit, so we can go ahead and jump to our next three sound changes:

3. Metathesis of initial #V.CV to #CV.V

4. Intervocalic /h/ is lost.

5. Adjacent vowels lengthen.

   a. Also, /ai/ → /e:/ and /au/ → /o:/

6. Metathesis of vowel length: /CV(C).CV:/ → /CV:(C).CV/

Of these changes, the first one (3) is perhaps the one I'm least comfortable with. Metathesis is often sporadic, though some languages feature it systematically; here, it serves to bring some of those noun class markers into the stem so that they can lengthen adjacent vowels and affect the now preceding consonant. If you'll permit me this, I'll appreciate it.

With these other rules, a word like /ka.ta.ha/ becomes /ka.ta.a/ which becomes /ka.ta:/ and finally /ka:.ta/. If we use /ha/ as a common suffix in the proto-language, it'll mean that modern words which once had it will now feature elongation of their final vowel, not to mention that this elongation will have effects on later shifts in stress placement (and subsequent vowel reduction). We could also have stops assimilate to one another, so /sakta/ would become /satta/, then we could have these geminates shorten, lengthening the preceding vowel, yielding /sa:ta/. I won't do this as it would complicate derivation a little, and I particularly like sequences of stops, but it is something to consider.

Our next changes are going to be two instances of umlaut, particularly vowel-raising caused by /i/ and vowel-lowering caused by /a/. This will cause the following changes:

7. Umlaut raises /e a u o/ to /i e i e/ when the next syllable contains /i/.
8. Umlaut lowers /i e u o/ to /e a o a/ when the next syllable contains /a/.

Later down the line, this will become important, but for now keep in mind that a word like /ka.ta.bi/ has become /ka.te.bi/. Our next few changes will ensure that this sound change is phonemic rather than allophonic:

9. Unstressed short vowels reduce to schwa between intervocalic consonants (except the glottal stop).
10. The schwa is lost entirely.

I could just as easily write this as one rule—unstressed, short vowels are lost between intervocalic consonants—but for illustrative purposes it is good to know that sounds usually go through some medial phase, reduced to a schwa in the case of vowels, before being dropped entirely.

We've already discussed it a bit, but this means that a word like /katabana/ becomes /katbana/. Later, when consonants undergo postvocalic lenition, this will become /kaθbana/. If we hadn't done this medial vowel reduction, all non-initial stops would lenite except after /n/, which isn't quite what I'm going for. Instead, having two separate occurrences of vowel reduction allows us to have some more variable syllable structures in the modern form.

It should also be noted that this doesn't affect long vowels (although currently, if I'm not mistaken, they only appear penultimately). For example, if we run through the sound changes so far, a proto-word, /panataha/, features penultimate stress /pa.na.'ta.ha/; since there is no /i/, we can skip palatalization (and umlaut); the /a.ha/ becomes /a:/ and the vowel length gets metathesized onto the (new) penultimate syllable, yielding /pa.na:.'ta/. Since there are no short, unstressed vowels between intervocalic consonants, we can skip vowel reduction as well. All these changes have resulted in semi-variable stress. Words will either feature penultimate stress and a short penultimate vowel or will feature final stress and a long penultimate vowel.

We maintained vowels before the glottal stop, but now we're going to undo that. *This* is where we're going to get our ejectives from:

11. Short, unstressed vowels reduce and are lost between any consonant and the glottal stop.

12. Stops, affricates, and sibilants become ejectivized (and devoiced, if applicable) preceding the glottal stop.

13. The glottal stop is lost after all consonants.

If we run through all the sound changes again with the hypothetical word /tiʔanaha/, it goes through the following phases: /ti.ʔa.na.ha/ → /ti.ʔa.ˈna.ha/ → /t͡si.ʔa.ˈna.ha/ → /t͡si.ʔa.ˈna.a/ → /t͡si.ʔa.ˈnaː/ → /t͡si.ʔaː.ˈna/ → /t͡sʔaː.ˈna/ → /t͡sʼaː.ˈna/. In short, /ti.ʔa.ˈna.ha/ → /t͡sʼaː.ˈna/. Later, we'll lose these ejectives, but they will have some lingering effects on our phonology.

      With our ejectives now distinguished from our plain stops, we can have the latter undergo postvocalic lenition, finishing off the final step of the first phase of our language's evolution.

14. All stops lenite after vowels: /p t t͡s t͡ʃ k b d d͡z d͡ʒ g/ → /f θ s ʃ x v ð z ʒ ɣ/.

With this change, we've reached the beginning of the second stage of the language. From here, the name of the game is stress shifts, vowel reduction, and epenthesis. This phase is much less consonant focused than the previous one, but it will involve some discussion of terminology relating to stress. Before we begin, I should note that our current phonology looks like this:

Consonants		Labial	Dental	Alveolar	Palatal	Velar	Glottal
Nasal		m		n			
Stop	plain	p		t		k	ʔ
Stop	ejective	p'		t'		k'	
Stop	voiced	b		d		g	
Affricate	plain			t͡s	t͡ʃ		
Affricate	ejective			t͡s'	t͡ʃ'		
Affricate	voiced			d͡z	d͡ʒ		
Fricative	voiceless	(f)	(θ)	s	ʃ	(x)	h
Fricative	ejective			s'	ʃ'		
Fricative	voiced	(v)	(ð)	z	ʒ	(ɣ)	
Approximant		w		l	j		
Tap				ɾ			

Vowels				
	Front		Back	
	short	long	short	long
High	i	iː	u	uː
Mid	e	eː	o	oː
Low	a	aː		

Consonants in parentheses are currently only allophones of other consonants, but
we'll remedy that soon enough. If we make a chart of some hypothetical words

and inflections, we can perhaps get a sense of the path we've followed thus far and where we're going:

Word Evolution		
Proto	Phase One	
	Phonemic	Realization
/simata/	/ʃi.'ma.ta/	[ʃi.'ma.θa]
/simataha/	/ʃi.maː.'ta/	[ʃi.maː.'θa]
/siʔimataha/	/ʃˁi.maː.'ta/	[ʃˁi.maː.'θa]
/kataba/	/ka.'ta.ba/	[ka.'θa.va]
/katabaha/	/ka.taː.'ba/	[ka.θaː.'va]
/kaʔatabaha/	/kˀa.taː.'ba/	[kˀa.θaː.'va]

We'll begin the second phase with a series of changes—modeled after a few that Hebrew underwent—intended to further reduce vowels and cause certain lengthenings to occur.

15. Stress shifts to pretonic heavy syllables preceding tonic light syllables.

16. Loss of all short, unstressed final vowels.

Heavy syllables are considered those that have either a coda or a long vowel; a pretonic syllable is the syllable right before a tonic syllable, and a tonic syllable is a stressed syllable. If we look at a previous example, /ka.taː.'ba/, these two

changes would cause it to become /ka.'ta:b/. Next, we'll go ahead and drop certain

unstressed, short vowels:

17. All short, unstressed vowels are reduced to schwa except when they
    precede another vowel.
18. The schwa is dropped except after ejectives or where its loss would result
    in illegal clusters.
19. /k'/ → /q'/
20. Loss of ejectivization.

Now, if we replicate that chart from earlier, we can see how our derivations are

coming along:

Word Evolution		
Proto	*Phase One*	*Current Form*
/simata/	/ʃi.'ma.ta/	/'ʃmaθ/
/simataha/	/ʃi.ma:.'ta/	/'ʃma:θ/
/siʔimataha/	/ʃˀi.ma:.'ta/	/ʃə.'ma:θ/
/kataba/	/ka.'ta.ba/	/'kθav/
/katabaha/	/ka.ta:.'ba/	/'kθa:v/
/kaʔatabaha/	/k'a.ta:.'ba/	/qə.'θa:v/

We're almost at the end now. All we need to do is merge the schwa with /a/, /w/ with /v/, /θ/ and /ð/ with /t/ and /d/, and shift a few other sounds and we're done. In sequence, that's:

21. /e o {a ə}/ → /ɛ ɔ ɐ/
22. /w/ → /v/
23. /θ ð/ → /t d/
24. /x ɣ/ → /χ ʁ/
25. /ɾɾ/ → /r/

This wraps up the sound changes from our proto-language to its modern form. I failed to mention this earlier, but these sorts of changes would likely take a good amount of time to occur, just given *how* different the proto-language is compared to the modern one. I have this many changes occur because, in my experience, it lends itself well to the sort of complex morphology I'm going for, but if this were meant to be a highly regular, agglutinative language with quite simple phonotactics, you wouldn't have (nor would you probably want) to go through this many changes.

Let's take a look at a few words and how they evolved from the proto-language. This is assuming they go relatively unaffected by other processes throughout the duration of these changes.

Word Evolution

Proto	Middle	Modern
/simata/	/ʃi.'ma.ta/	/'ʃmɐt/
/simati/	/ʃi.'me.t͡si/	/'ʃmɛs/
/simataha/	/ʃi.ma:.'ta/	/'ʃma:t/
/simatiha/	/ʃim.'ti.a/	/'ʃimt/
/siʔimataha/	/ʃʼi.ma:.'ta/	/ʃɐ.'ma:t/
/kataba/	/ka.'ta.ba/	/'ktɐv/
/katabi/	/ka.'te.bi/	/'ktɛv/
/katabaha/	/ka.ta:.'ba/	/'kta:v/
/katabiha/	/kat.'bi.a/	/'kɐtb/
/kaʔatabaha/	/kʼa.ta:.'ba/	/qɐ.'ta:v/

For some of these, we can begin to see how we might create regular derivations. For example, the standard way to derive the present tense conjugation for verbs might be: CCaC. If we treat the old /-i/ as a past tense marker, then forming the modern past tense is as easy as swapping out that /a/ for an /ɛ/.

Let's say, hypothetically, that our root for "writing is M-K-T. We could take that root and insert the vowel in its proper place: /MKɐT/. However, certain phonemes will take different forms in different situations: in this case, that 'K' underwent postvocalic spirantization historically, so in this position it is /χ/. Thus, our word becomes /m̩χɐt/.

To continue with this, let's say we had a historical suffix /-la/ that formed the polite form of any given verb tense. Thus, the old present tense, polite ending would have been /-ala/. If we look at where stress would have fallen historically,

we can begin to see how this might ripple down into the modern form. The plain form of the verb would have been /ma.ˈka.ta/ while the polite form was /ma.ka.ˈta.la/. This affected which vowels were deleted, resulting in the medial forms of /ma.ˈxa.ta/ and /max.ˈta.la/ respectively. The second phase of our sound changes would further alter these words, resulting in the aforementioned plain form of /m̩.ˈχɐt/ but a polite form of /ˈmɐχ.tl̩/. Interestingly, this would make it so that the polite forms of the past and present tense are identical. Here's a chart:

Hypothetical Conjugations		
	Present	*Past*
Plain	CCɐC	CCɛC
Polite	CɐCCl̩	CɐCCl̩

We have two options, given this overlap: leave it as is or find some new way to introduce a distinction between the tenses in their polite forms. Speakers might make use of some other method, maybe the plural form, to mark the polite past tense, or they would simply shrug and make no distinction between these verb forms—that's a perfectly realistic situation. Fusional languages often have different conjugations (and declensions) that are identical; so as long as context or the use of other constructions can disambiguate, it isn't particularly a problem. That being said, I do intend to draw influence from Coptic which has a particularly interesting way of marking tense that may come in handy here.

Coptic has a feature called nominal TAM—that is, it marks tense, aspect, and mood on the noun rather than the verb (sometimes). We're going to make use

of this by having verbs mostly inflect for formality and certain non-finite forms, while tense and other things like mood and maybe even evidentiality will (often) be marked on the noun. But first, I should explain how this might come about and how we're going to have it play out in the language we're making here (which I should really give a name, so I can stop referring to it so obliquely). I really like that syllabic /l/, and since we want to have a fancy writing system it'd make some sense to derive the name from something to do with writing. We'll go ahead and steal that third-person, polite form of the verb meaning "to write" and use it as the language's name: Makhtl.

Okay, so Makhtl is going to feature nominal TAM as well, but we have to decide here and now something about the way the language is going to be structured. We have a few options, but broadly we can break them down into two camps: head-initial or head-final. The latter is like English, placing the "head" of a phrase at the beginning, while the latter is like Japanese, which does the reverse. In more concrete terms, this determines some things about how the language orders its adpositions (words like "in," "at," "to," etc), how it orders its verb and its object, and whether it is mostly prefixing or suffixing.

Most of the languages we've listed as influences are head initial. Coptic, for example, almost exclusively uses prefixes, while something like Turkish almost exclusively uses suffixes. I'd venture to guess that head-final conlangs are more common since the English-speaking conlanging community is likely to find head-final languages unfamiliar and thus more interesting to work with, so that leans me towards choosing a head-initial structure, but I've also been working on

too many head-initial languages recently so I might as well give in to the temptation and go with a primarily suffixing, head-final language. This has certain implications, as I said, so we should go about outlining them, but first we should know what forms these words will take, and that means finally outlining our triconsonantal root system.

5 | ROOTS & DERIVATIONS

You know, I almost forgot this essay had subsections. So far, this has run about six and a half thousand words, two thirds of which can be found in the last section alone; I think I oughta break it up a little. Then again, I imagine this section will run quite long as well, so really does it make a difference?

Well, it is about time we got around to laying out the regular derivations one uses to derive most nouns and verb forms. We'll treat adjectives as a subset of verbs, as quite a few languages do.

I would be remiss not to mention (again) that, in this regard, I owe something of an original inspiration to Tiramisu's *Tutorial: Making a Realistic Triconsonantal Language*. For a long time, I altogether avoided even thinking about making a triconsonantal root language; like many others, one of my first conlangs was an attempt at it and it went about as well as anyone would expect. I knew diddly-squat about language evolution and even less about how the Afroasiatic languages in particular came into their respective forms, and I thought you could just slap a few vowels between some consonants and call it a day. After realizing how stupid of an assumption that was, I avoided triconsonantal roots like

the plague, until honestly very recently, when I finally worked up the will to read some actual grammars of Afroasiatic languages and research their historical sound changes in an attempt to design something even vaguely realistic.

I have always loved the sound of these languages, but I could never replicate their underlying systems in a way that wasn't ham-fisted (and who knows if I even have now, you can feel free to lambast me via email if I'm still off the mark).

Of course, I want Makhtl to have some character of its own, hence my deviation in various ways from the Afroasiatic model. For one, I'm drawing on the Bantu noun class system and Turkic morphology, but still the heart of this language is its triconsonantal root system.[15]

Most roots will be triconsonantal, though a number of biconsonantal and quadconsonantal roots will also be present. In the following chart, I'll outline the standard derivation pattern, but I need to explain first that the subscript 'L' marks a consonant as 'lenited.' Sometimes, consonants may appear to be in a position where they would've been lenited, but they are only postvocalic due to later epenthesis, meaning they weren't affected by historical lenition. In other places, those that come after vowels underwent historical lenition; it's only that the vowel that triggered it has since been dropped. This lenition process is important as it maintains the distinction between certain forms. Furthermore, a subscript 'P' will indicate palatalization, whereby /t d k g s z/ become /t͡s d͡z t͡ʃ d͡ʒ ʃ ʒ/. These

---

[15] If you're interested in a system (though notably different), outside the Afroasiatic family, you should read Dianne Friesen's *A Grammar of Moloko*—a language with a neat phonology and neater phonotactics. Anyways, we should continue on to derivations.

palatalized forms can also be affected by lenition, in which case we'll feature a subscript 'PL.'

Plain Verb Template				
Voice	Type			
	Stem	Active Participle	Passive Participle	Infinitive
Plain	$CC_L ʁ C_L$	$CC_L ɛ C_L$	$CC_L e{:}C_L$	$CuC_L ʁ C_L$
Causative	$juCʁCC_L$	$juCʁCC_L$	$iCCe{:}C_L$	$ju{:}CCʁC_L$
Reciprocal	$mo{:}CCʁC_L$	$mo{:}CCɛC_L$	$mo{:}CCe{:}C_L$	$mu{:}CʁCC_L$
Reflexive	$Ca{:}C_L C_L$	$Ca{:}C_L C_L$	$Ca{:}C_L e{:}C_L$	$a{:}CCʁC_L$
Causative Reflexive	$jʁCCC_L$	$jʁCCC_L$	$iCCe{:}C_L$	$ja{:}CCʁC_L$

Thus, the plain form, stem form of "to write" is /mχʁt/, while its causative, passive participle form is /imke:t/. There are surely going to be a variety of irregular verb forms, but for the most part we'd expect roots to adhere to this pattern.

As for the polite form, here's that chart:

Polite Verb Template				
Voice	Type			
	Stem	Active Participle	Passive Participle	Infinitive
Plain	$CʁCCl$	$CC_L C_P ɛl$	$CC_L C_P ʁhɛl$	$CuC_L Cl$
Causative	$iCCCʁl$	$iCCC_{PL} ɛl$	$iCCɛC_{PL} l$	$jo{:}CCCʁl$

Reciprocal	mo:CCCɐl	mo:CCC_{PL}ɛl	mo:CCɛC_{PL}l	mu:CCCɐl
Reflexive	Ca:CCɐl	Ca:C_LC_Pɛl	Ca:C_LC_Pɐhɛl	a:CCCɐl
Causative Reflexive	iCCCɐl	iCCC_{PL}ɛl	iCCɛC_{PL}ɐl	ja:CCCɐl

If we take a look at our root, M-K-T, again, its polite, plain, stem form would be /mɐχtl/, but if we shift that into the polite, causative-reflexive, passive participle form it becomes /imkɛsɐl/.

These plain and polite forms will serve as the foundation onto which we'll attach various clitics and affixes and whatnot, but we'll get around to that once we've fleshed out our nouns a little more, as these also bear some of the burden of marking tense, aspect, and mood.

I didn't mention this earlier, but it deserves some explanation: these forms are derived from auxiliary verbs or particles that are semantically bleached and worn down into affixes that then, due to sound changes, irreversibly fused with the stem. *Everything comes from somewhere* when it comes to language development, and while surely there are some exceptions, it serves one well to treat this as a hard rule.

I also didn't mention *this* earlier, but if any of these forms would cause an illegal cluster, epenthesis occurs, inserting an /ɐ/ to fix the disallowed sequence. However, if the cluster involves a consonant that can serve as a syllable nucleus, it'll get reanalyzed as a separate syllable altogether, hence why we don't insert an /ɐ/ into /mχt͡sɛl/ even though /mχ-/ is a disallowed sequence according to our phonotactics. Instead, the /m/ is treated as a syllable nucleus: /m̩χ.t͡sɛl/. If that /m/

were to be an /v/, epenthesis would be required, turning the word into /vɐχ.t͡sɛl/. Since this is going to become pertinent soon, I should make one more note regarding the origin of roots.

Often, in dealing with triconsonantal root languages, it is tempting to consider roots as something primordial, a platonic form floating in the aether, eternal, but really they come from the application of patterns that exist with one group of words onto other groups of words that might not have featured those patterns previously.

For example, we have our root, M-K-T, that probably came from the proto-language's words for "chisel," /maka/, and "to do," /ta/. At some point, speakers needed a word for writing, and since we're meant to be mirroring Ancient Egyptian's writing system, we'll say that they started by chiseling into stone. So they took "to do," and they used it with "chisel," and it yielded /makata/, "to chisel." Over time, as new methods for writing developed, this word broadened in use until it could apply to writing in general. It lost its easily discernible relation to the word for "chisel" and was reanalyzed as its own, separate semantic domain.

Now, speakers want a word for the nifty new brush they're using to write on papyrus or whatever they've gotten their hands on, so they look at some other words and see that often, creating the instrument from a verb involves the template /t͡se:CCɐC/, so they extend this to the verb, /mχɐt/, yielding /t͡se:mkɐt/. In essence, they extended a derivational method from another set of words onto a verb which previously didn't need an instrument—they already had "chisel"—and

in doing so pushed the system that much closer to having the broadly applicable derivational systems that characterize triconsonantal root systems. If you extend this to many, many more verbs (and you begin deriving verbs from nouns that previously had no related action, using similar processes) then you can begin to see how one could look at the resulting system and come to the conclusion that triconsonantal roots underpin everything.

What we need to do now though is talk about deriving nouns. Essentially, each of the aforementioned voices we just displayed in those charts will have their own forms of the following: verbal nouns, locations, agents, patients, and instruments. Thus, our (blank) chart looks something like this:

Noun Template					
Voice	Type				
	Verbal Noun	Location	Agent	Patient	Instrument
Plain					
Causative					
Reciprocal					
Reflexive					
Causative Reflexive					

I'm going to be a bit more granular about this—going through each type, one by one, so that you can get a sense for where these come from.

First, we have the verbal noun which is just a nominalized form of the verb, akin to our "-ing" ending, as in "running" or "playing." In both Korean and

Japanese, there are forms of nominalization that come from the word for "thing," as in an event or affair, so we're going to derive our verbal nouns from this. I suppose we might understand this as "the event / affair / thing of [verb]." It would make sense then for this to resemble a genitive construction, so historically it'd look something like /[verb] ina/, where /ina/ is the construct state of the word /i/, with that aforementioned meaning of "thing" or "event." Over time, this ending would erode and eventually get suffixed onto the verb, becoming our modern form /-ɛn/. Of course, this also affects stress, and thus, which vowels are reduced, so it isn't as simple as slapping that ending on, at least not by modern times.

Noun Template					
Voice	Type				
	Verbal Noun	Location	Agent	Patient	Instrument
Plain	CC$_L$C$_P$ɛn				
Causative	iCCC$_{PL}$ɛn				
Reciprocal	mo:CCC$_{PL}$ɛn				
Reflexive	Ca:C$_L$C$_P$ɛn				
Causative Reflexive	iCCC$_{PL}$ɛn				

Thus, the reflexive, verbal noun form of M-K-T is /ma:χt͡sɛn/. Looking back at the polite forms of the verbs, we could honestly describe this as clipping that final /-(ɐ)l/ off the polite form and adding /-ɛn/ in its place. Ah, but we have plenty more noun forms to cover so I shouldn't linger too long on this one.

Anyways, our next one is location. This is a little trickier. We're going to employ a particular method for these next two that'll need some explanation. In essence, the location and agent form of roots will derive from that old noun class system we talked about. In fact, I'll go ahead and bang both of them out here:

Noun Template					
Voice	Type				
	Verbal Noun	Location	Agent	Patient	Instrument
Plain	$CC_LC_P\varepsilon n$	$CuC\text{ɐ}C_L$	$CiC_L\text{ɐ}C_L$		
Causative	$iCCC_{PL}\varepsilon n$	$jo{:}CC\text{ɐ}C_L$	$j\varepsilon CC\text{ɐ}C_L$		
Reciprocal	$mo{:}CCC_{PL}\varepsilon n$	$mu{:}CC\text{ɐ}C_L$	$mi{:}CC\text{ɐ}C_L$		
Reflexive	$Ca{:}C_LC_P\varepsilon n$	$a{:}CC\text{ɐ}C_L$	$a{:}CC\text{ɐ}C_L$		
Causative Reflexive	$iCCC_{PL}\varepsilon n$	$ja{:}CC\text{ɐ}C_L$	$ja{:}CC\text{ɐ}C_L$		

As you can see, some reflexive and causative-reflexive forms derive identically, meaning the words for "one who causes theirself to be [verbed]" and "the place where one causes theirself to be [verbed]" are the same. Plenty of roots probably won't feature causative-reflexive forms (it'll be among the rarer voices), so I'm not particularly worried about this overlap.

These two forms derive from the /u-/ and /i-/ noun class prefixes that were historically used to mark the "location" and "person" classes respectively. This also means that these two classes would likely retain (at least, for some words) the old method of pluralization: turning these prefixes into /uku-/ and /ihi-/

respectively. We'll cover this in greater depth when we get around to noun morphology, but remember this point as it'll be important.

We might as well get the next two done in one fell swoop as well. These are the patient and the instrument forms: the recipient of the action and the means by which it is completed. We actually touched on the latter of these earlier: it derives from an old noun class prefix, /tihi-/, which itself was just once a word for a tool. Thus, we can plug that into our template:

Noun Template					
Voice	Type				
	Verbal Noun	Location	Agent	Patient	Instrument
Plain	$CC_LC_P\varepsilon n$	$CuC\text{ʁ}C_L$	$CiC_L\text{ʁ}C_L$		$\widehat{ts}e{:}CC\text{ʁ}C_L$
Causative	$iCCC_{PL}\varepsilon n$	$jo{:}CC\text{ʁ}C_L$	$j\varepsilon CC\text{ʁ}C_L$		$ise{:}CC\text{ʁ}C_L$
Reciprocal	$mo{:}CCC_{PL}\varepsilon n$	$mu{:}CC\text{ʁ}C_L$	$mi{:}CC\text{ʁ}C_L$		$mi{:}se{:}CC\text{ʁ}C_L$
Reflexive	$Ca{:}C_LC_P\varepsilon n$	$a{:}CC\text{ʁ}C_L$	$a{:}CC\text{ʁ}C_L$		$\widehat{ts}a{:}CC\text{ʁ}C_L$
Causative Reflexive	$iCCC_{PL}\varepsilon n$	$ja{:}CC\text{ʁ}C_L$	$ja{:}CC\text{ʁ}C_L$		$ise{:}CC\text{ʁ}C_L$

Again, the causative and causative-reflexive forms overlap in their instrument forms, but we're not particularly worried about the ambiguity. I really should've titled this essay: *How I Learned to Stop Worrying and Love the ~~Bomb~~ Ambiguity Arising from Overlapping Derivational Patterns*. Sorry, the quality of my humor is one-to-one with the amount of coffee I've had, and I have not nearly had enough.

Anyways, that leaves the patient. If we use the agent as its base, we can derive it through the regular application of some suffix that—in our historical narrative—will have arisen well after the agent form came into being; we'll use a /-aːt/, from an old word having to do with states and conditions. Our chart would then look something like this:

Noun Template					
Voice	Type				
	Verbal Noun	*Location*	*Agent*	*Patient*	*Instrument*
Plain	$CC_LC_P\varepsilon n$	$CuC_L\text{ʁ}C_L$	$CiC_L\text{ʁ}C_L$	$CiC_L\text{ʁ}C_L\text{aːt}$	t͡seːCCʁC_L
Causative	$iCCC_{PL}\varepsilon n$	joːCCʁC_L	jɛCCʁC_L	jɛCCʁC_Laːt	iseːCCʁC_L
Reciprocal	moːCCC$_{PL}\varepsilon$n	muːCCʁC_L	miːCCʁC_L	miːCCʁC_Laːt	miːseːCCʁC_L
Reflexive	Caː$C_LC_P\varepsilon$n	aːCCʁC_L	aːCCʁC_L	aːCCʁC_Laːt	t͡saːCCʁC_L
Causative Reflexive	$iCCC_{PL}\varepsilon n$	jaːCCʁC_L	jaːCCʁC_L	jaːCCʁC_Laːt	iseːCCʁC_L

Because the patient forms come from the agent forms, they're likely to feature those irregular plural forms I mentioned earlier.

For illustrative purposes, we might as well take a look at all the forms of M-K-T, though keep in mind that some of these aren't actual words, only hypothetical forms of the root:

Noun Template					
Voice	Type				
	Verbal Noun	*Location*	*Agent*	*Patient*	*Instrument*

Plain	m͡χt͡sɛn	muχet	miχet	miχeta:t	t͡se:mket
Causative	imksɛn	jo:mket	jɛmket	jɛmketa:t	ise:mket
Reciprocal	mo:mksɛn	mu:mket	mi:mket	mi:mketa:t	mi:se:mket
Reflexive	ma:χt͡sɛn	a:mket	a:mket	a:mketa:t	t͡sa:mket
Causative Reflexive	imksɛn	ja:mket	ja:mket	ja:mketa:t	ise:mket

Now that we have our basic noun derivations, we can begin talking about nominal morphology: how these nouns inflect for state, number, and various other important features.

6 | NOUNS

We have something of a foundation: we know Makhtl is going to be a nominal TAM, head-final, semi-fusional but mostly agglutinative language, but now we have to decide a few more things: firstly, just *how* agglutinative are we talking.

I'll admit to a particular weakness for nominal morphology reminiscent of Turkish. Possessive affixes are a particular feature that shows up in way too many of my languages—they're just too damn good—and conlangers in general tend to go a little too heavy on our noun cases. There are a few languages, like Indonesian, which feature possessive affixes without extensive case systems. In fact, Indonesian lacks noun cases entirely, making it quite interesting in that regard. I could use a Japanese-esque system; that language has quite complicated verbal inflection but its nouns never change their form (though they do take particles for case and whatnot). If we look at Coptic, it uses prepositions in a manner akin to case markings, but otherwise it has lost case marking, so we'll

probably emulate this system in Makhtl with the addition of some Turkish influence.

Most affixes that attach to Makhtl's nouns will have formed relatively recently and are thus expected to be regular, though there will be stem alternations for different affixes that one will have to watch out for. One side effect of our nominal TAM marking is that we could get away with leaving the explicit distinction between the subject and the object unmarked, as whichever takes the TAM marking would be the subject, but since nominal TAM is meant to be a relatively recent innovation (and since I suspect it will prove more complicated than I currently expect), it would make more sense for there to remain some indication of a subject-object distinction from before these markings evolved.

Here's what I'm thinking: we'll have some inflections that are older than the rest. These will have evolved during the first phase of the language's development, making them considerably more fusional than the other affixes and clitics. The rest—cases, TAM marking, and possessive suffixes—will attach on to one of these stems (depending on the word's syntactic function) and thus we'll have the semi-fusional, semi-agglutinative system that we desire.

Before we do that though, we should explain what these states actually are. Many of the Semitic languages feature three noun forms called "states," that are a little like cases but function instead, usually, to mark definiteness and possession. Aramaic in particular has the absolute, emphatic, and construct states which function like an indefinite, definite, and possessed form respectively. Specifically, the absolute state is the default form of a noun and it marks

indefiniteness (*a* store, *a* job, *a* person, etc). The emphatic state, on the other hand, is used for definite nouns (*the* store, *the* job, *the* person). And lastly, the construct state (which has a variety of additional uses in various languages) is generally used to mark the possessee in a genitive phrase. I should probably explain: a genitive phrase is a phrase which subordinates one noun to another, usually denoting possession but sometimes used for other things like origin or composition. Thus, if I wanted to say "the person's stone" in Aramaic, I would put "stone" in the construct state and "person" in the emphatic, like:

person-EMPH stone-CON.

Given this feature's pervasive presence in the Semitic language family, I should probably include it here. To explain the reasoning behind how I'm going to have these states manifest, I should probably construct a sort of narrative describing the grammatical evolution of the language. We'll say, for now, that in its proto-language form it was much less rigidly head-final, featuring a mixture of prefixes and suffixes, but over time it has shifted much more towards head-finality, though some of its old mixed features remain. I wanted to get *that* out of the way so that I could plop this chart down and tell you where each form came from without it seeming like I'd forgotten this was supposed to be a mostly suffixing language.

If we go back to our agent nouns again, we can create a chart for their respective forms (in the plain voice). For this particular category, the states would look something like this:

Agent Noun States			
#	**State**		
	Absolute	*Emphatic*	*Construct*
SG	$CiC_LɐC_L$	$CiC_Le{:}C_L$	$CiC_LɐC_L\underset{.}{n}$
PL	$i{:}CCɐC_L$	$i{:}CCe{:}C_L$	$i{:}CCɐC_L\underset{.}{n}$

If we take our good old root for "writing" and fit it with this templates, the result is:

States ~ Plain Agent Derivation of M-K-T			
#	**State**		
	Absolute	*Emphatic*	*Construct*
SG	miχɐt	miχe:t	miχɐt\underset{.}{n}
PL	e:mkɐt	e:mke:t	e:mkɐt\underset{.}{n}

Neat, yeah? We have /miχɐt/ which means "a writer" and /e:mke:t/ which means "the writers." This is sure to have some overlap with the other voices. In fact, let's go ahead and run through each of the states for each of the voices:

States ~ Causative Agent of M-K-T			
#	**State**		
	Absolute	*Emphatic*	*Construct*
SG	jumkɐt	jumke:t	jumktɐn
PL	ju:mkɐt	ju:mke:t	ju:mkɐt\underset{.}{n}

States ~ Reciprocal Agent of M-K-T		
#	State		
	Absolute	Emphatic	Construct
SG	mi:mkɐt	mi:mke:t	mi:mktɐn
PL	me:mkɐt	me:mke:t	me:mkɐtn̩

States ~ Reflexive Agent of M-K-T		
#	State		
	Absolute	Emphatic	Construct
SG	a:mkɐt	a:mke:t	a:mkɐtn̩
PL	e:mkɐt	e:mke:t	e:mkɐtn̩

States ~ Causative-Reflexive Agent of M-K-T		
#	State		
	Absolute	Emphatic	Construct
SG	ja:mkɐt	ja:mke:t	ja:mkɐtn̩
PL	je:mkɐt	je:mke:t	je:mkɐtn̩

This paradigm only applies to this category of nouns; the others will feature more concatenative methods of marking states.

I would expect this old noun class system to hold on most strongly for nouns referring to people, only because we seem to tend to have more granular distinctions for animate classes of nouns than inanimate ones. I'm sure exceptions to this can be found, and if I'm way off the mark, let me know, but for now I'll run with the assumption that other noun classes will have collapsed into a much

more uniform paradigm. Similarly, very old, basic words (such as "water" or "fire") will be more likely to have unique forms.

Just to introduce a little more flavor into the system, we're going to say that Makhtl distinguishes between two grammatical genders: animate and inanimate.[16] The former obviously includes all animate things—humans, animals, etc—as well as miscellaneous other words, which take states via the methods we described above; meanwhile, the inanimate gender will include all other derivations and operate a little differently. In short, nouns of the inanimate gender will distinguish between collective and singulative forms rather than singular and plural. What this means is that the default, unmarked form of a noun is understood to refer to a collection of the noun (or that plurality is not particularly important, contextually); for example, the default form of the word is "pencils" rather than "pencil," and one must use a singulative marker to (specifically) refer to only one "pencil."

"Inanimate" Nouns ~ State Suffixes			
#	State		
	Absolute	*Emphatic*	*Construct*
COL	—	-e	-n̩
SGV	-oːt	-eːt	-aːtn̩

---

[16] I'd originally had Makhtl feature a human-inhuman distinction, the sort which occurs in the Dravidian languages, such as Tamil, though theirs is a semantic distinction rather than a strictly morphological one. I've since decided that it'd make more sense to change this to a plain animate-inanimate distinction, for illustrative purposes.

If we take a look at one "animate" gender word and one "inanimate" one, we can see the difference more plainly: /miχeːt/ means "the writer" and /tɐχeːt/ means "the seed." The former comes from the root, M-K-T, while the latter comes from T-K, referring to sowing or planting.[17]

These two classes will determine a bit about the way that their respective nouns are marked for their syntactic function—speaking of which, we should finally get back around to the rest of our nominal morphology.

At some point in the past, Makhtl featured more complex case markings, but these have largely worn away by modern times, replaced by what used to be postpositions which have now been suffixed onto the noun. Effectively, Makhtl has reintroduced cases after a period of caselessness, though these new ones are not so old as to have lost much of their locative uses. Similarly, certain quirks arise from these cases' origins. For example, Makhtl uses one of these cases to mark the direct object of a verb, so we'll call it the accusative case, but this originated as a form of differential object marking, so it is only used with animate nouns.[18] Inanimate nouns lack an accusative case; in sentences where an inanimate noun is the direct object, Makhtl must make use of a bit more rigid of a word order to avoid ambiguity.

And finally, these cases attach to whichever element is the last in a noun phrase; this means that if a noun is followed by an adjective, the case will get

---

[17] Just as a side note, this root came from the word for "seed" and was expanded beyond its original meaning to include other terms for farming and whatnot.

[18] Differential object marking can be found in languages like Spanish, where the preposition "a" is used to mark certain direct objects—those that are both human and specific. Our use of this is a little broader, lacking the specificity requirement, but it is still fundamentally similar.

attached to the adjective instead of the noun. I ought to run through these cases

now, before I get carried away:

Cases			
Case	Class	Suffix	Uses
Nominative	*Animate*	—	- This is the default, unmarked form of all nouns; it is used when the noun is the subject of the sentence.
	Inanimate		
Accusative[19]	*Animate*	-(ɐ)ʃ	- The accusative is used to mark direct objects of transitive verbs. - It marks the subject of certain stative verbs, mostly having to do with experiences.
	Inanimate	—	
Dative	*Animate*	-iːl	- The dative case is used to mark indirect objects, the beneficiary or recipient of the action. - It is also used to indicate motion towards or to something, akin to an allative case.
	Inanimate		
Genitive	*Animate*	-ɛv	- The genitive case is used to mark possession, origin, description, or composition. The subordinated noun—the possessee, origin, etc—comes after the subordinator.
	Inanimate		
Instrumental	*Animate*	-(ɐ)r	- The instrumental indicates that the noun is the means by which the action is accomplished: its instrument.
	Inanimate		
Comitative	*Animate*	-t͡si	- The comitative indicates that the subject completes the verb alongside or in conjunction with the noun marked in the comitative.[20]
	Inanimate		

---

[19] Coincidentally, I hadn't realized that Turkish does something very similar with its accusative case. I was worried that this sort of thing was a little unnaturalistic, but now my fears are assuaged.
[20] In English, we use "with" for both the instrumental and comitative, but some languages like Russian distinguish between these two uses.

These cases will serve as our "core" cases. There are a number of other locative cases, but we'll leave those for later. Right now, we ought to talk about word order.

7 | WORD ORDER & SENTENCE TYPES

A number of Semitic languages have "verbless" sentences or clauses—they lack what we call the copula, the sort of verb that equates one thing with another. English has one copula, "to be," and we employ it quite a lot. Other languages, like Spanish, have a couple that are used in different situations. Some, including Russian, drop that copula in particular situations (particularly, the present tense, though I don't know Russian so call me out on this one if I'm mistaken). Arabic is much the same, dropping its copula in the present tense. Since this is a neat feature, we're going to include it in Makhtl. It doesn't require too much of an explanation beyond what I've just said, but if anything requires clarification I'll try to fill you in as we go.

This takes us to our second discussion: word order. Many of the Semitic languages feature two options when it comes to word order: VSO and SVO, verb-subject-object and subject-verb-object respectively. While I love VSO, the problem is that it is (almost) entirely restricted to head-initial languages, but earlier I decided to be all fancy and go with a head-final tendency, and now that's biting me in the butt. We could use an SVO word order like English or some of the Arabic dialects, but that's also probably the most familiar order to anyone reading this, and while SOV is *really* common, unless you've studied a language

that features it, you'll probably find it more entertaining if I go that route. So, despite being almost the opposite of Middle Egyptian, Classical Arabic, and Biblical Hebrew—we're going to go with SOV. It's fun, in its own ways. For example, if I wanted to say some gibberish sentence, it'd look like this:

Mikhētsh mēsīl mselqef.

/mi.ˈχeːtʃ meː.siːl m̩.sɛl.qɛf/

mikhēt-sh	mēs-īl	msel-q-ef
writer\EMPH.SG-ACC	water\EMPH.SG-DAT	give\POL-PST-1SG

I gave the writer the water.

I've foreshadowed one of our future topics: person agreement on verbs via clitics. We could try some polypersonal agreement, cause that's always fun, but I think we might stick with just marking the subject. But I'm getting ahead of myself: we're still talking about word order.

You might be wondering how we'd order verbless sentences (those with a zero copula), but it's honestly pretty simple. While most sentences are likely to be fairly free with their word order, since we have noun cases, these won't be as easy. Essentially, we're going to designate two parts of verbless sentences and clauses: the subject and the predicate. This structure extends to verbal sentences as well, but our rule for Makhtl will be that the predicate (specifically the verb) always appears at the end of the sentence. The object will be able to move a little bit, if we want to emphasize it, but the verb will rigidly sit at that clause-final position. For verbless sentences, the thing that the subject is being equated to will

fill this position, so the first of these examples will be grammatical but the latter won't be:

Łter makh ōt.

/l̩.tɛr mɐχ oːt/

lter   makh   ōt

PROX   chisel   one

This is one chisel.

* Makh ōt lter.

/mɐχ oːt l̩.tɛr/

makh   ōt   lter

chisel   one   PROX

* This is one chisel.

This second example would be grammatical if you meant something like "one chisel is this," but Makhtl will disallow the use of the demonstrative pronouns in such a position.

      I considered, here, making Makhtl an ergative-absolutive language, but we've already got enough on our plate—I don't want to overburden the language with neat features and end up diminishing the whole. That being said, I am going to draw from Basque regarding topic-prominency, so call me a hypocrite if you will. This subject requires a little bit of an introduction.

Topic-prominent languages include Japanese, Korean, and Lakhota, and are generally characterized by the way they structure their sentences with regard to the topic of discourse. The topic is generally what is being talked about, and the comment is what is being said about it. A language like Japanese or Korean has an explicit marker that it puts after the topic, but Basque doesn't have this: instead, it only moves its topic to the front of the sentence.[21]

Basque is a subject I could kinda talk endlessly about, so I need to show a little restraint, but I will say that: A) you should read about it, if you haven't already—it's very different from any of the surrounding languages, being the only language isolate in Europe—and B) it has some neat rules that I think we can include in Makhtl. These rules are:

1. The topic comes at the beginning of the sentence.
2. The comment comes right before the verb.

Thus, Makhtl will move its topic to the front of the sentence and will *tend* to take the emphatic state. Since we have noun cases, we can pretty freely move words around without making things that ambiguous. If we take a look at some example sentences, we'll see this in action:

Djātnenakh kahīl djatnelqef.

/d͡ʒaːt.nɛ.nɐχ kɐ.hiːl d͡ʒɐt.nɛl.qɛf/

---

[21] I say "only" because Japanese, at least, also has its topic at the front of the sentence, it just also has its marker to doubly indicate the topic. I'm honestly not sure how many topic-prominent languages feature a marker and how many opt for movement alone; I should really read up more on that.

djātnen-akh   kah-īl   djatnel-q-ef

clothing-3SG   3SG-DAT   weave\POL-PST-1SG

(As for) his clothing, I weaved it for him.

Kahīl djātnenakh djatnelqef.

/kɐ.hiːl d͡ʒaːt.nɛ.nɐχ d͡ʒɐt.nɛl.qɛf/

kah-īl   djātnen-akh   djatnel-q-ef

3SG-DAT   clothing-3SG weave\POL-PST-1SG

For him, I weaved his clothing.

As you can probably tell, both the word for "clothing" and "weaving" come from a root, DJ-T-N, which has to do with weaving. The word for clothing comes from the reflexive form, which came to mean "to clothe oneself" via semantic drift. Due to its reflexive origin, it works a little differently than our word, "to wear." In Makhtl, the article of clothing is indicated with the instrumental case, for example:

Mēsr djātnqef.

/meːs.r̩ d͡ʒaːt.n̩.qɛf/

mēs-r   djātn-q-ef

water\EMPH.SG-INSTR   self_clothe-PST-1SG

I clothed myself with the water. / I wore the water.

This is only poetic because I lack any words for articles of clothing, and while I could go about making one, I'd rather get along with our discussion of word order. But, as a quick side note, I want to add that these subject clitics that I keep applying to the verbs can be dropped if: A) the subject is explicitly stated in the sentence or B) it is evident from context who is doing the action. For example:

Mēsr nefaq djātn.

/meːs.r̩ nɛfɐq d͡ʒaːt.n̩/

mēs-r   nef-aq   djātn

water\EMPH.SG-INSTR   1SG.NOM-PST   self_clothe

As for the water, I wore it. / I wore the water.

Something that's neat about Coptic is the relative fluidity with which its nouns can be turned into clitics that attach to the verb or other nouns, serving as subject markers or possessive prefixes. These clitics bring with them the same tense-aspect-mood prefixes that they would feature in their independent forms, but—if I'm not mistaken—we run into certain complications when we try to replicate this in Makhtl.

In researching clitics, I came across Ethelbert Kari's article on the clitics and affixes of Degema, a language of Degema Island, Nigeria, in which he cites Arnold Zwicky and Geoffrey Pullum's article on the differences between the two, between clitics and affixes. Having now managed to get my hands on the latter article, I can present to you, hopefully in more digestible terms, the differences they observe between clitics and affixes:

1. Clitics tend to "care" less about what they're attaching to, able to grab onto whichever word happens to fall in the right position. In contrast, affixes tend to only attach to one kind of word, be it a noun or a verb or whatnot.

2. Affixes tend to have arbitrary gaps in which words they can attach to and the conditions under which they can appear.

3. Affixes tend to care more about the phonological shape of the word they're attaching to—what phonemes it has—while clitics are much less discerning.

4. Affixes have particular semantic baggage and idiosyncrasies—their meaning can be a lot more variable depending on the particular word they're attaching to, while clitics tend to be pretty constant in this regard.

5. Syntactic rules can shuffle words with affixes around much easier than they can words with clitics. This is kinda related to the first difference, seeing as clitics cling much more to a particular position in a clause while affixes cling much tighter to the particular word they're modifying.

6. And finally, clitics can attach to other clitics, but affixes (usually) cannot attach on the outside of a clitic. That is: [stem]+[clitic]+[clitic] is allowed, but [stem]+[clitic]+[affix] is not.

Note, however, that this doesn't hold up in some languages, though it does seem to be the case most of the time.

With these rules, we run into some problems. If our nominal tense-aspect-mood markers are clitics, then we'd expect them to stay in a particular syntactic position and attach to whichever word precedes them. In Coptic, they do seem to do this. If the subject itself becomes a clitic on the verb, the TAM clitic simply attaches right in front of it. We either have [TAM clitic]+[subject] [verb] or [TAM clitic]+[subject clitic]+[verb].

However, because of the way we've ordered Makhtl, we can't really do that. With our SOV word order, any TAM clitics that would attach to the subject are left stranded when it is dropped, and it's difficult to justify them moving all the way to the back of the verb (due to rule five). To remedy this, we'll probably have to say that these markers occur just after the topic, and if the topic happens to be dropped, then they just act like independent particles. That isn't entirely strange, if we look at other head-final languages like Japanese.[22]

彼は今食べています。

kare	wa	ima	tabe-te	i-masu
he	TOP	now	eat-CONJ	is-POL

He is eating right now.

---

[22] I'm most familiar with Japanese, hence why I keep falling back on it for references to head-final structures.

I want to be clear that this isn't a great gloss of this sentence, I'm just lazy and we're only really interested in the first three words. Makhtl lacks that topic particle, so we can ignore it, but if we imagine a hypothetical change where, over some years, 今, /ima/, came to cliticized onto 彼, /kare/, then it wouldn't be far fetched for this (hypothetical) sentence to look something like this:

Kareima tabete imasu.

kare=ima   tabe-te   i-masu

he=PRS   eat-CONJ   is-POL

He is eating.

If many, many years pass and this clitic becomes an affix, then we've got something like our desired nominal TAM going on. However, this doesn't seem to be the method by which any natural language has developed nominal TAM, so I'm a little wary of running with it. What I really want is to figure out how Coptic developed its nominal TAM and use that as a guide. I'll report back with what I find.

8 | COPTIC & NOMINAL TAM

Dr. Chris H. Reintges, from the Université Paris 7, writes that "the presence of nominal features" in Egyptian verbs caused them to "no longer [be] compatible with the exponents of tense, aspect, and mood distinctions." In turn these became "externalized outside of the verbal domain as auxiliary-like conjugation bases." If we take this as our model, it seems like the erosion of verb structures facilitates,

to some degree, the shift (or the potential to shift) to nominal TAM. What this means for our purposes is that Makhtl should—in its changes from the proto-language to the modern one—see the loss of its old tense, aspect, and mood inflections, as well as any person or number marking that it featured, in order to facilitate the conditions necessary for nominal TAM to arise.

Reintges gives us a description of the changes in these structures between Old Egyptian and Coptic, and while they're a little heady—knee-deep in X-Bar Theory—he gives us a nice summary near the end of his paper, saying that the "shift from synthetic to analytic morphology" correlated with a change to verbs which saw the main verb lose its marking for "finiteness and tam marking," which in turn caused nominal TAM to arise as "the sole representation of finiteness and core propositional features."[23] In other words, as verbs become increasingly analytic, opportunities arise for novel methods of TAM representation to appear, and due to the features Reintges talked about, this came to be nominal TAM.

Our historical narrative will go something like this. Old Makhtl featured an SVO word order which, with the loss of verb and noun inflection, saw a rise in the use of auxiliary verbs to mark tense, aspect, and mood. As the language reintroduced cases and became progressively more head-final, the main verb shifted back to a final position, but these auxiliaries remained in a medial position having cliticized to pronominal subjects; with non-pronominal subjects, they moved to a postverbal position along with the main verb. Thus, we have two kinds of verbal sentences: those with a pronominal subject, in which TAM

---

[23] Reintges, Chris H. "Increasing Morphological Complexity."
https://conf.ling.cornell.edu/DiGSX/abstracts/Reintges.pdf

markers are cliticized to the end of the subject, and those with non-pronominal

subjects, in which TAM markers are cliticized to the end of the verb. This is

motivated somewhat by Rachel Nordlinger and Louisa Sadler's description of

nominal TAM in their paper on "Tense as a Nominal Category," published by UC

Berkeley.[24]

I'll admit to being a little out of my depth on some of this stuff, and if

anyone has any better recommendations on how to derive a head-final, SOV

language with nominal TAM, I'd be glad to hear it. While this seems reasonable

to me, there's plenty about theories of grammar that I don't know. I also want to

stress that every paper I've read on the subject points to there existing a quite

lively debate about the underlying mechanisms for nominal TAM, so keep that in

mind.

Anyways, we've (kinda) justified (to varying degrees) most of the

structures I wanted for the language, but I should go ahead and run down a list of

less complex rules for Makhtl's word order:

- Adjectives follow nouns.

- Postpositions follow nouns.

- There are three ways to mark possession that differ in word order, but

  broadly speaking the possessor precedes the possessee.

- Auxiliary verbs follow the main verb.

- Relative clauses follow the nouns they modify.

---

[24] Nordlinger, Rachel, and Louisa Sadler. "Tense as a Nominal Category." LFG00 Conference (2000).

With these rules, we'd expect the average sentence to look something like this:

Nefev mēs shvelr zinētsh tēyef nefaq msatta.

/nɛ.fɛv meːs ʃvɛl.r̩ zi.netʃ teː.jɛf nɛ.fɐq m̩.sɐt.tɐ/

nef-ev	mēs	shvel-r	zinēt-sh
1SG-GEN	water\EMPH.SG	cold\APRT-INSTR	man\EMPH.SG-ACC

tēy-ef	nef-aq	msat-ta
know\PPRT-1SG	1SG.NOM-PST	clean-POT

I could have cleaned the man I knew with my cold water.

As for my cold water, I could have cleaned the man I knew with it.

You'll notice that we don't have an explicit relativizer; instead, participles are used: the active for when the noun being modified is also the one doing the action, and the passive when its the object of that action (or in any other situation).

We also get to see one of those TAM markers that only ever shows up on the verb: /tɐ/. It is the potential voice, indicating that the subject has (or in this case, *had*) the potential to do the action: as in, they are able to do it. It actually doesn't come from the verb /tɐ/, "to do," but instead from a similar root as that verb in our relative clause /tɐj/, which means "to know." The root had to do with one's knowledge or ability, hence why it ended up appearing in both these places. Also, I translated this into English's past perfect tense only because "I could…" is often used to indicate a future potential rather than a past one.

And lastly, I considered using the postposition form of the instrumental case, /era/, to show where it would go in the sentence, but then I'd lose out on the

ability to show that cases get added to the ends of adjectives if they come after

nouns, and that honestly seemed more interesting.

We've basically covered the broad rules of word order that are most

necessary, at least for now, and while there are certainly many more pages I could

waste talking about quirks and exceptions, I'd rather move on to our writing

system—the whole point I began this essay in the first place.

9 | LGCNSNNTL

So we've finally reached hieroglyphs. Thirteen thousand words, and we've finally

reached the original topic I wanted to talk about. My crippling inability to say

anything briefly is definitely not to blame.

Before we start outlining our own system, it'll serve one well to know that

Hieroglyphs—and, by extension, Demotic and Hieratic—make use (broadly) of

three kinds of phonetic components: uniliterals, biliterals, and triliterals. These

indicate one, two, or three consonants in one sign respectively. For some

examples, we have   which signifies /m/,   which signifies /pd/, and   which

signifies /   nẖ/.[25] We're going to copy this structure for Makhtl, but first we

should talk about conservative orthography.

Middle Egyptian, like English, had a fairly conservative writing system,

especially as the spoken language continued to evolve. We're going to emulate

this to a degree: the language will broadly spell words as they were after the first

phase of its evolution, albeit attempting to keep all the consonants in

triconsonantal roots uniform across the derivations. This means that a word like

---

[25] I use the //s here only so that the /   / isn't difficult to read, as it would be between quotations marks.

/tɐ/ will be written ⟨h-t⟩ as it historically featured an /h/ that has since been dropped. Similarly, the word for water, /msi/, will be written ⟨m-t⟩, reflecting the underlying biconsonantal root.

Most words will be relatively similar to their current forms; the name of the language, /mɐχt.l̩/, will be written ⟨m-x-t-l⟩, likely with two biliterals or one uniliteral and one triliteral. It'd also have a determinative having to do with speech, attached to these literals, which leads us into our next point.

Oftentimes, Middle Egyptian marked its determinatives with a little line below or beside it. We'll go ahead and do something similar. This allows the reader to know when they're reading something that is meant to be read semantically versus phonetically, and it'll clear up word boundaries nicely (since we won't have any spaces).

And lastly, while structurally this is going to be akin to Hieroglyphs, as I said before I'm a sucker for the aesthetics of Demotic and the Arabic script, so we'll be drawing quite a bit on those two for our visual influence. I also *really* like the look of Hebrew, but my plate is already full with those aforementioned scripts, so I can't really afford to fit that in. Alas, I must show some restraint.

Neat, we have the foundations of our script.

Now it's time to get to work.

To set a limit on the length of this project, I'm going to say that we'll devise enough of the script to be able to write the sentence I used as an example earlier:

Nefev mēs shvelr zinētsh tēyef nefaq msatta.

/nɛ.fɛv meːs ʃvɛl.r̩ zi.netʃ teː.jɛf nɛ.fɐq m̩.sɐt.tɐ/

nef-ev	mēs	shvel-r	zinēt-sh
1SG-GEN	water\EMPH.SG	cold\APRT-INSTR	man\EMPH.SG-ACC

tēy-ef	nef-aq	msat-ta
know\PPRT-1SG	1SG.NOM-PST	clean-POT

I could have cleaned the man I knew with my cold water.

This will hopefully allow us to show off determinatives, uniliterals, biliterals, and triliterals without spending forever expanding upon a system that is ultimately only supposed to serve as an example of the foundation of such scripts. The problem with creating logographic scripts for conlangs is that it takes forever; whereas many languages make use of a few dozen glyphs for their entire system (or, at least, most of it), logographic systems sometimes require a few hundred components that combine to form tens of thousands of characters. Middle Egyptian seems to feature fewer atomic components than something like Chinese, at least in its Hieroglyph stage, but it still had quite a few. Since I'm not inclined to sit here creating hundreds and hundreds of characters for an example language, I'm going to break down all the components I'll need to create here and now, going word by word.

⟨Nefev⟩ is the first-person, singular pronoun, ⟨nef⟩, with the genitive case ending. Thus, we'd expect it to be made up of a biliteral ⟨n-f⟩, possibly without a determinative since its a pronoun, and a uniliteral ⟨v⟩.

⟨Mēs⟩ comes from ⟨msi⟩, meaning "water," the kind of word we'd expect to have a pictogram for instead of a more complex glyph. However, this form of the word is going to feature a reduced form of the uniliteral, ⟨t⟩—which we'll indicate with ⟨t̠⟩—to mark it as emphatic. This is the glyph used to mark the singulative, emphatic form of inanimate nouns, and it is used here because otherwise the absolute and emphatic forms of "water" would be identical. Plurality, on the other hand, is always marked with three dots after the word; we'll indicate here with ⟨∴⟩. Thus, the plural, emphatic form of "water" is ⟨m-t t̠ ∴⟩, read as /aːmeːs/. The plural, emphatic form of "writer" is ⟨m-k-t t̠ ∴⟩, read /eːmkeːt/. In this sentence though, water is not plural, so it'll only be ⟨m-t t̠⟩.

⟨Shvelr⟩, much like the first word, features a uniliteral case marker, ⟨r⟩, attached to what could either be a triliteral, ⟨sh-v-l⟩, or a biliteral and a uniliteral, ⟨sh-v l⟩. We'll go with the latter, so this word will be ⟨sh-v l r⟩.

⟨Zinētsh⟩ is the singular, emphatic form of ⟨zinat⟩, which could be written with a pure logogram, depicting a man. This component would, if used in other words, serve as the triliteral for ⟨z-n-t⟩ (and really it kinda serves that purpose here, just without a determinative). Again, since this is emphatic, it'll feature that particular, reduced uniliteral, ⟨t̠⟩, and this word has the accusative case marker ⟨sh⟩, so altogether that makes ⟨z-n-t t̠ sh⟩.

Our next word is ⟨tēyef⟩, which is a form of the word ⟨tay⟩ with the first-person, singular subject marker attached. The verb comes from the root, "h-t-j," which itself comes from two roots, "h-t + j." If we use this old origin, we

can arrive at the written form (for the root) ⟨h-t y⟩, and with the added ⟨f⟩ that becomes ⟨h-t y f⟩.

⟨Nefaq⟩ is the same as earlier ⟨nefev⟩ but with a ⟨q⟩ instead of a ⟨v⟩. Thus, we need to add that to our list of necessary uniliterals. And finally, ⟨msatta⟩ comes from the same root as "water," ⟨m-t⟩, with the addition of a uniliteral ⟨t⟩. In fact, we'll have two of those uniliterals, as the potential voice clitic also takes the form of a ⟨t⟩. Since this word has to do with water, we'll reuse that pictogram with the added determinative marker, to reinforce that it is being used for its semantic value, along with to uniliteral ⟨t⟩s.

Altogether, we need nine unilaterals, if we include the reduced form of ⟨t⟩ that'll serve as the emphatic marker. These unilterals are: ⟨v⟩, ⟨l⟩, ⟨r⟩, ⟨sh⟩, ⟨y⟩, ⟨f⟩, ⟨q⟩, ⟨t⟩, and ⟨t̲⟩. We'll also need four biliterals, one of which will be used solely as a pictogram and another which will serve both as a semantic and phonetic component of a larger word; these are: ⟨n-f⟩, ⟨m-t⟩, ⟨sh-v⟩, and ⟨h-t⟩. And finally, we have one triliteral that is also used as a pictogram: ⟨z-n-t⟩. Oh, and we're going to need a determinative as well: one having something to do with the cold or perhaps statuses. Luckily for me, I picked a couple words that'd probably be basic enough to have pictograms—"man" and "water"—and that really reduces my workload quite a bit. Other than that though, we need to determine where the rest of these will come from.

Here is the finished product:

Nefev mēs shvelr zinētsh tēyef nefaq msatta.

In retrospect, I probably should've toned down the Arabic influence a little bit, but the deed is done. This is probably a lot to look at, so we'll take it one word at a time.

Firstly, we have ⟨nefev⟩, which comes from the biliteral ⟨n-f⟩ and the uniliteral ⟨v⟩. As something of a tribute to Hieroglyphs, I based ⟨n-f⟩ off of a hawk which, after its evolution, you can see only really vaguely see. The tail below the line is what remains of its feet, and the curve up to the head once was the curve along its back. The uniliteral, ⟨v⟩, comes from a pictogram for "mouth," hence its shape. You'll see a similar shape in the biliteral for ⟨ʃ-v⟩, but this one comes from an eye, hence the little protrusion afterwards which once was its iris, but we'll get around to that one soon enough.

The next word is ⟨mēs⟩, which uses as the pictogram for "water" as its semantic and phonetic component. This is indicated via the loop below the glyph. You may still be able to see the glyphs resemblance to waves or ripples on water. As for the semantic component attached afterwards—the ⟨t̲⟩ we talked about earlier—it is a

reduced form of the triliteral for "writing," from the root M-K-T that we've used several times thus far.

Easily the longest of our words, ⟨shvelr⟩ owes its length to its relative semantic complexity; it is perhaps not as basic as "water" or "man," and thus requires the semantic component—indicated by that little curly mark under the last section of the word—which derives from the triliteral for "house," once used as a determinative for shelter from the elements which has now been extended to apply to the elements themselves, hence why we use it here to help the reader know how to read the preceding components. The initial biliteral, ⟨sh-v⟩, derives from an eye, hence its similarity to the glyph for "mouth." Similarly, you can see the origin of the next uniliteral quite plainly. This one, ⟨l⟩, comes from a serpent, which here looms over the next uniliteral, ⟨r⟩, which is a little more obscure but which ultimately derives from a pictogram for a knot, related to the root ZH-T-R, which was taken to just indicate its last element, ⟨r⟩.

We find one of our next pictograms in this one, ⟨zinētsh⟩: a somewhat heavily abstracted image of a person kneeling in prayer. It bears the little curl telling the reader it is a determinative, though it might just as easily be used as a triliteral for the root Z-N-T. It features the emphatic marker, as well as the uniliteral, ⟨sh⟩, which once was the image of a tree; its trunk has now shifted down below the line, its branches becoming part of the characteristic through-line of the script. It can also serve as a pictogram for ⟨hesh⟩, the word for tree, in which case it would take the determinative marker.

Second to last, we have a repeat of the first-person pronoun, this time with the past tense marker /q/ attached. This uniliteral descends from an ideogram for the word /qɐv/, meaning "direction" or "way." These days, the resemblance is not quite as strong, but it used to resemble a vertical line with an arrow pointing towards something, as if you'd taken 上 and flipped it on its side. Again, here it has been used for its initial consonant, /q/.

And finally, we have the character for "water" again, still used for both semantic and phonetic purposes, but now it has an additional uniliteral repeated twice after it: /t/, derived from a pictogram for a hammer. This was is used for the for word, "to do," which is /tɐ/, or for its sole consonant, /t/.

I'm quite glad I limited myself to these words and ensured that our sentence featured quite a few basic words that would be more likely to feature simple logograms or pictograms. In non-contrived situations, I'd expect most words to resemble ⟨shvelr⟩, which again featured a biliteral, two uniliterals, and a determinative. This is, by far, the most common variety of Chinese character—one semantic component and another phonetic component—and while Hieroglyphs and its descendents function a bit differently, for anything more complex than "man" and "water," the reader is likely to need a guide if they're to find any meaning, wading through so many consonants.

This essentially closes this essay. Hopefully by running through an example or two I've provided a sort of framework from which you can grow your own languages, and if you have any bits of advice, comments, or complaints, feel free to reach out. Any logographic writing system, in my experience, that is intended for a naturalistic language will take forever and a day to make, so understand that before you set out, and you'd do well to read up a bit about real life languages which used (or still make use) of such systems beforehand.

I intend to write a few more of these soon, and though I'd planned on my next one being about Chinese characters, the amount of words that I'll need to properly explore that topic is really setting in—I may take a break to talk about poetry in the Basque Country and Ancient Greece, or perhaps I'll finally scratch that itch to create a zero-marking language in the vein of Indonesian or Classical Chinese. In any case, I hope I've entertained you; stick around, perhaps, and see what I ramble about next.

Thanks for reading.