

UNIT 7 EXOTIC WORD FORMATION

7.1 Is Morphology Just About Stems and Affixes?

(1) Morphology is a theory of how related words are related:

Prefix + Stem:	joy – en joy	Suffix + Stem:	joy – joy ful
	write – re write		write – writ- er
German:	Enkel – U renkel		Enkel – Enkel in

- All the groups of words we have looked at have involved putting together discrete, identifiable pieces, e.g., roots, prefixes and suffixes. The technical term is **CONCATENATION**.
- There are (or appear to be) other ways in which words can be related to each other, for which it is harder to clearly distinguish a discrete affix. These are traditionally called **NON-CONCATENATIVE MORPHOLOGY** or **PROCESSES**. They come in (at least) two kinds:

Cases where we can point to the actual pieces that are involved, but where we can't draw a single neat dividing line between affix and stem.

Cases where we can describe the change, but can't isolate the phonemes involved.

- **REDUPLICATION**

(2) Tagalog verbs (spoken in the Philippines)

present	future	
takbuh	tatakbuh	'run'
lakad	lalakad	'walk'
pili?	pipili?	'choose'
sulat	susulat	'write'
hanap	hahanap	'seek'

OBSERVATION: The future tense involves repeating the first consonant and vowel of the stem/root.

- DISCONTINUOUS MORPHEMES (Root and Pattern)

(3) Classical Arabic (all forms 3sg masc subject)

	“write”		“earn”	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
present (X-s)	katab	kutib		
cause to X	kattab	kuttib	darraj	durrij
X for ea. other	kaatab	kuutib	daaraj	duurij
make X	?aktab	?uktib	?adraj	?udrij

OBSERVATIONS: All of the verbs meaning “write” have the consonants *k-t-b*.
 All of the verbs meaning “earn” have the consonants *d-r-j*.
 All of the active forms have (only) the vowel “a”.
 All of the passive forms have the vowels “u”, “i”, in that order.
 All of the “cause to X” forms have the middle consonant doubled.
 (None of the other forms do).

DILEMNA: It looks like the “morphemes” are: *k-t-b* = ‘write’, *u-i* = ‘passive’ etc..., but they are discontinuous, or interleaved.

- TRUNCATION

(4) French Adjectives

masculine feminine

kur	kurt	‘short’
frwa	frwad	‘cold’
nɛʒø	nɛʒøz	‘snowy’
frɛ	frɛʃ	‘cool’
for	fort	‘strong’
lõ	lõg	‘long’
norveʒjẽ	norveʒjɛn	‘norwegian’
bõ	bon	‘good’

(5) Analysis ? If the masculine is underlying, then the feminine “suffix” would be unpredictable, -t / kur__, -d / frwa__, -g / lõ__ etc...

But if the feminine is underlying, then the masculine “affix” must be a deletion rule (delete last consonant of feminine)! = TRUNCATION

BIG PICTURE DILEMNA:

- It looks like we can just write “transformations”, i.e., rules that take some string of phonemes and transforms it into another string of phonemes.
- But if we can do that, we can describe anything.
- And if we can describe anything, we can predict... nothing.
- And therefore we can explain... nothing.

To show that *non-concatenative morphology* really is *concatenative* (just with more going on).

(Recall that we have seen evidence that ABLAUT in English is really stem allomorphy, and not an instance of non-concatenative morphology).

(6) For the future tense in Tagalog, the part of the stem that is repeated is the first consonant and first vowel. They are repeated at the front of the word (like a prefix). And they are repeated in the order they show up on the verb.

takbuh	tatakbuh	'run'
lakad	lalakad	'walk'
pili?	pipili?	'choose'
sulat	susulat	'write'
hanap	hahanap	'seek'

In a sense, then, there is a prefix in these forms. It is a sequence of a Consonant and a Vowel, and it “means” future. It just has no specification of the consonant or vowel, and needs to “borrow” them from the stem (we’ll see the mechanics of this in more detail later).

$$(7) \quad \overset{\text{---}}{\text{C V}} \text{ sulat} \rightarrow \text{ susulat} \quad \overset{\text{---}}{\text{C V}} \text{ takbuh} \rightarrow \text{ tatakbuh}$$

- What are these Cs and Vs ?
- How does this copying work?

7.2 Autosegmental Phonology and The Skeletal Tier

Some phonological rules refer directly to features (such as $[\pm\text{voice}]$, $[\text{dorsal}]$, etc...). Others, as we've seen, crucially refer to consonants (C) or vowels (V), regardless of the features involved. For example, syncope in Icelandic (PS 3) was stated:

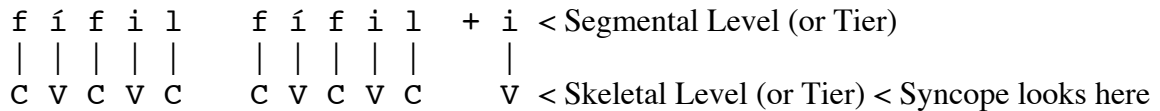
(8) Syncope: V → ∅ / ____ C V
 [-stress]

(9) This interacted with other MS/cyclic rules, and was responsible for alternations such as:

akur	‘acre’	
akri	‘acre [DATIVE]’	from: / a k u r / + AFFIX
fifil	‘dandelion’	
fifli	‘dandelion [DATIVE]’	from: / f i f i l / + AFFIX

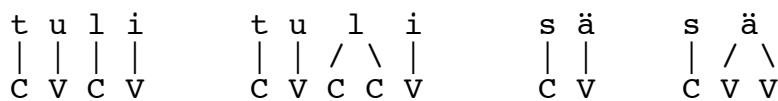
In order for these rules to operate correctly, it is useful to have a representation which is split into two parts: one with all of the features or SEGMENTS, the other with just the SKELETON, i.e., the Cs and Vs:

- (10) ‘dandelion’ ‘dandelion [DATIVE]’



Phonological rules stated in terms of Cs and Vs often take long vowels or geminate consonants to behave as sequences of two vowels or two consonants, even though in some sense they have only one set of features. We can represent this by allowing many-to-one mappings between the skeletal and segmental levels:

- (11) Finnish: tuli ‘fire’, tulli ‘customs’; kuka ‘who’, kukka ‘flower’
 meri ‘sea’, Meeri ‘Mary’; sä ‘you’, sää ‘weather’



- Why is this necessary?

- (12) Finnish “illative” case (means “to X”)

auto ‘car’	autoon ‘to the car’
koulu ‘school’	kouluun ‘to school’
päivä ‘day’	päivään ‘to the day’
kuva ‘picture’	kuvaan ‘to the picture’
bussi ‘bus’	bussiin ‘to the bus’

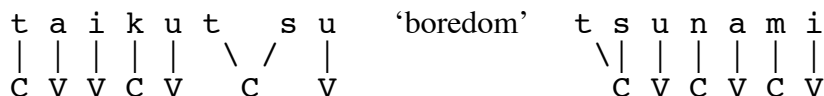
OBSERVATION: The suffix meaning “to (the) X” [called the Illative Case] is easy to figure out. It is the last vowel of the stem, followed by N. By allowing multiple-linking, we can easily diagram what is going on:

- (13)
- | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a | u | t | o | + | n | ➔ | a | u | t | o | n | |
| | | | | | | | | | | | | |
| V | V | C | V | | V | C | V | V | C | V | V | C |

We can also have multiple association working the other way, with a single skeletal slot (C or V) linked to more than one segment. What would this look like?

- (14) Affricates (and in some languages diphthongs) count for some phonological purposes as if they are a single unit, even though they clearly have two pieces. For example, Japanese does not allow consonant clusters in onset position, but it does allow the affricate “ts”.

For example, the word “grizzly” is borrowed as *gurizuri* since *gr* and *ʒl* are impossible.
 But:



Important: The NO CROSSING LINES principle.

While many-to-one and one-to-many mappings are possible, the lines which link the segmental tier to the skeletal tier MAY NEVER CROSS.

(15) French Underlying forms:

kurt	frwad	nɛʒøz	frɛʃ	fort	norvezjɛn	bon
	\				\	
CVC	CC V	CVCV	CCV	CVC	CVCCVC V	CV

Affix: Feminine = C i.e., an **un**-specified consonant.

Masculine = Ø (or no affix)

Derivation:	(U.R.)		(Affixation)		output:
	k u r t		k u r t		
		→	\	→	[kurt]
	C V C + C		C V C + C		

(16) French nasalization

norvezjɛn	bon	n = [+nasal]
\		
CVCCVC V	CV	

The feature “nasal” can link to both consonants and vowels.

Feminine:	(U.R.)		(Affixation)		output:
	b o [nas]		b o [+nas]		
		→	\	→	[bon]
	C V		C V C		
Masculine:	(U.R.)		(Affixation)		output:
	b o [nas]		b o [+nas]		
		→	/	→	[bõ]
	C V		C V		

Aside (if you’ve taken phonology)...

Writing C’s and V’s on the skeletal tier involves some redundancy. It is a good first approximation, and we will continue to use it. Phonological theory provides a theory of syllabification which allows one to largely predict the distribution of C’s and V’s on the skeleton from the structure of syllables (as well as providing accounts of phenomena such as compensatory lengthening). We might then want to write only “X”’s in place of the C’s and V’s (in which case, we could call them ‘timing slots’). Since there is a direct translation from the CV theory into the more articulated X-slot theory, we will continue to use the more transparent CV notation in this class.

7.2.1 Reduplication... Technicalities

We want to maintain the No Crossing Lines constraint, but we also want to say that the future morpheme in Tagalog (p.83, ex (2), repeated below) is a prefix with just skeletal information C V (i.e., it has no segmental information). The key is in the phonological rules that link the unlinked CV skeletal units to segments.

First, link up all the Cs and Vs in the root. Then add the affix and:

1. Copy the segments of the entire root over the affix.
2. Link the segments to the skeleton one to one, left to right.
3. Delete all unlinked / leftover elements.

Illustrations:

(17)	takbuh	tatakbuh	‘run’
	lakad	lalakad	‘walk’
	pili?	pipili?	‘choose’
	sulat	susulat	‘write’
	hanap	hahanap	‘seek’

Initial version (p.84):

(18) $\overbrace{C\ V} \text{ s u l a t} \rightarrow \text{ s u s u l a t} \quad \overbrace{C\ V} \text{ t a k b u h} \rightarrow \text{ t a t a k b u h}$

(19) Applying the rules:

1. Copy the segments of the entire root over the affix.

s	u	l	a	t	s	u	l	a	t
									“write”
C	V	+	C	V	C	V	C		
“future”									

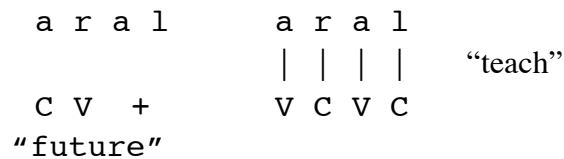
2. Link the segments to the skeleton one to one, left to right.
3. Delete all unlinked / leftover elements.

s	u	(l a t)	s	u	l	a	t
		↗					
		↗					“write”
C	V	+	C	V	C	V	C
“future”							

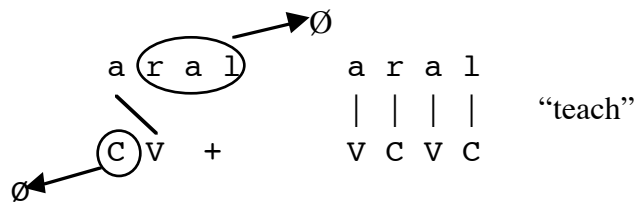
↗ ∅

(20) Why the additional complexity? First: vowel-initial roots *aral* ‘to teach’

1. Copy the segments of the entire root over the affix.



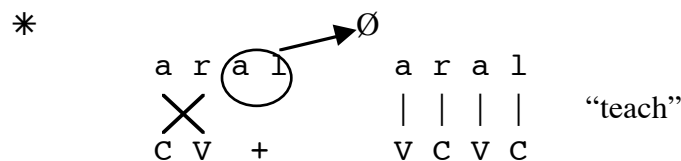
2. Link the segments to the skeleton one to one, left to right.
3. Delete all unlinked / leftover elements.



4. Output: aaral

Extra Phonology: $\emptyset \rightarrow ? / V_V$
True Output: a?aral

(21) Note the relevance of NO CROSSING LINES



Incorrectly predicted output: *ra(?)aral

The key to understanding this is another set of reduplication phenomena:

(22) FULL REDUPLICATION, Indonesian

Noun	All sorts of Nouns	
orang	orang orang	‘man’, ‘all sorts of men’
anak	anak anak	‘child’, ‘all sorts of children’
mangga	mangga mangga	‘mango’, ‘all sorts of mangoes’

The morpheme meaning “all sorts of” is a complete copy of the stem, regardless of the stem’s CV shape. Think of it this way:

(24) Greek:

Even Less Specified:

Tagalog:

CV- totally unlinked
= partial reduplication

Least Specified of All:

Greek:

perfectives (see below)

Indonesian: <copy> - full reduplication

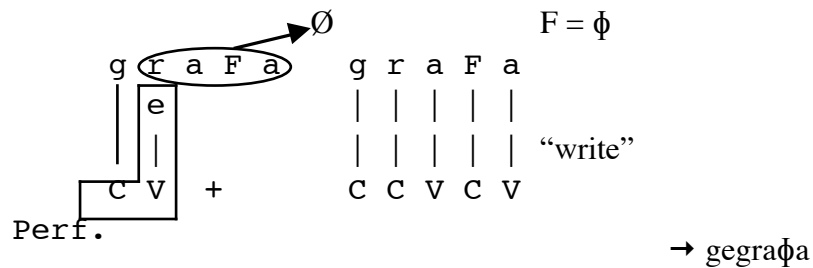
(24) Greek:

thæka	tethæka	'place'
klæka	kekklæka	'call'
luka	leluka	'lose'
bouleka	bebouleka	'consider'
grapha	gegrapha	'write'

Perfective: e- Like Finnish, part of the affix information is pre-specified
 | but some must come from the stem to which it attaches.
 C V Unlike Finnish, this happens non-adjacentlly.

$$\begin{array}{c} e^- \\ | \\ c \quad v \end{array}$$

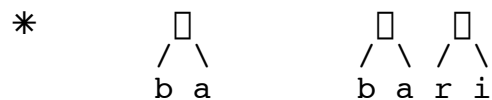
Like Finnish, part of the affix information is pre-specified but some must come from the stem to which it attaches. Unlike Finnish, this happens non-adjacently.



- (25) Those of you who have taken phonology (or who are taking it now) will be familiar with more detailed groupings of Cs and Vs into syllables, feet, etc. and you will have seen reduplication patterns where the reduplicant affix is a syllable or a foot. Consider the following pairs of nouns in Agta.

‘body’	bari	barbari	‘whole body’
‘leak’	saddu	sadsaddu	‘leak in many places’
‘lost’	wakay	wakwakay	‘many things lost’
‘leg’	takki	taktakki	‘legs’
‘head’	ulu	ululu	‘heads’

This language allows syllables of the form CVC. (Closed syllables). Clearly, we would like to say that the reduplicant is a one-syllable prefix. Nevertheless, we would get the wrong results if we just said that the reduplicant copied the first syllable of the base:



Even though the prefix is a syllable, it is important to note that it does not correspond to any actual syllable in the base. Rather, the copying procedure starts at the left edge of the base and copies as much as it can until it fills a maximal syllable, but without regard to the syllable groupings in the base.

(Note that, as with (20) above, it is important that the association be from the segmental tier to the skeletal tier).

7.2.2 *Arabic*.

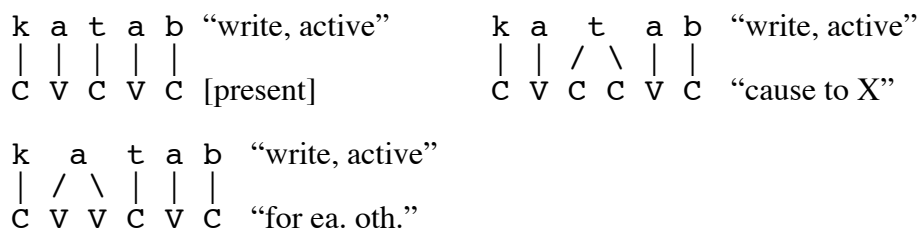
- (26) Classical Arabic (all forms 3sg masc subject)

	“write”		“earn”	
	ACTIVE	PASSIVE	ACTIVE	PASSIVE
present (X-s)	katab	kutib	daraj	durij
cause to X	kattab	kuttib	darraj	durrij
X for ea. other	kaatab	kuutib	daaraj	duurij
make X	?aktab	?uktib	?adraj	?udrij

- (27) Consider first the active forms of “write” and “earn”.
We may note the following:

- All of the present tense verbs have the shape CVCVC.
- All of the cause to X forms have the shape CVCCVC
- All of the X for each other forms have the shape CVVCVC
- All of the make X forms have the shape ?VCCVC

- (28) We have seen that there are morphemes consisting simply of Cs and Vs (as in the reduplicative prefix in Tagalog) or of CV sequences only partially linked (the Finnish Illative), we could say that the CV patterns just noted ARE the morphemes. This would leave us with stems that simply do not come with their own skeletal units:

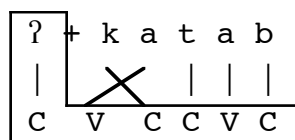


- (29) Two questions:

What about the active vs. passive distinction?

Active always /a/, passive always /u...i/

How can we get the “make X” forms without crossing lines?



- (30) The ROOT is just the consonants:

k-t-b is the only common element in ‘write’

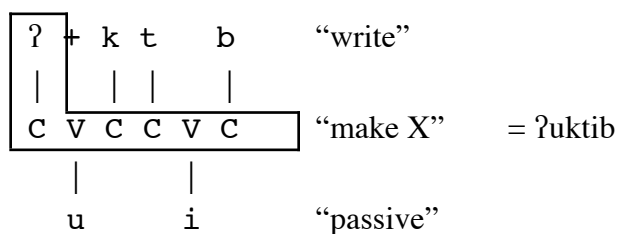
d-r-j is the only common element in ‘earn’

The vowels are separate morphemes:

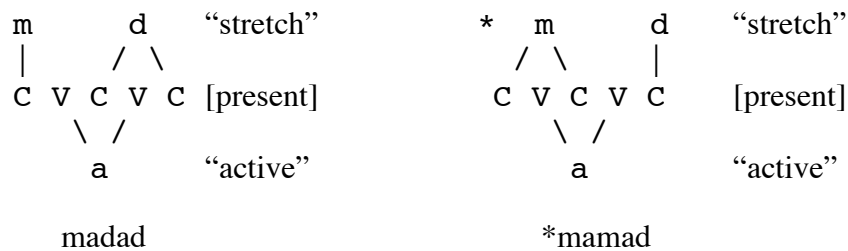
a is the morpheme for “active”

u-i is the morpheme for “passive”

- | | | | | | | | | |
|------|---|---|---|-----------|---|---|---|--------------|
| (31) | k | t | b | “write” | k | t | b | “write” |
| | | | | | | / | \ | |
| | C | V | C | V | C | C | V | C |
| | | | | [present] | | | | “cause to X” |
| | | \ | / | | | | | |
| | | a | | “active” | | u | i | “passive” |



- (32) There is an independent reason to want to keep the consonants separate in this way. Some roots have only two different consonants, these always show up as $C_1C_2C_2$ and never as $*C_1C_1C_2$:



- (33) We can derive this result on two assumptions. First, linking is left to right (at least in Arabic) and second, sequences of two identical adjacent elements are illicit in underlying representations (this is called the Obligatory Contour Principle—there must be some difference between adjacent segments).
- (34) With the mechanism of Autosegmental Tiers (i.e., keeping the various pieces on different planes), we are able to account for the range of phenomena attested in languages like Arabic (many Semitic languages, including Hebrew, work this way to a greater or lesser degree). At the same time, we can predict things that are not attested in any known natural language (note that such permutations do exist in invented languages / word-games, such as Pig-Latin, and the “speech disguise” of the Hijaazi of Saudi Arabia [Kenstowicz 1992, p408]).
- (35) For example: Cibara, a non-existent language:

*k-t-b ‘write’	→	katab [present]	batak [past]
*k-t-b ‘write’	→	uktib [passive]	utbik [active]

We can describe what’s going on here: the tense alternation involves reversing the order of the consonants in the root and the passive active alternation involves postposing the first consonant. Even with the rich power of a skeletal tier, we cannot capture such forms without crossing lines. We predict, correctly, that languages such as Cibara, despite their similarity to Arabic, cannot exist.

7.3 Floating Features

- (36) Our types of affixes in (23) above involved a range from fully specified to almost entirely unspecified (just “copy”), with, in the middle some affixes that were partially specified, having skeletal information (Cs and Vs) but little or no segmental or featural information. Now consider the following:

7.3.1 Consonant Changes: (MUTATION)

- (37) Nivkh verbs (spoken on Sakhalin Island, Russian Far East) ɖ = palatalized /d/

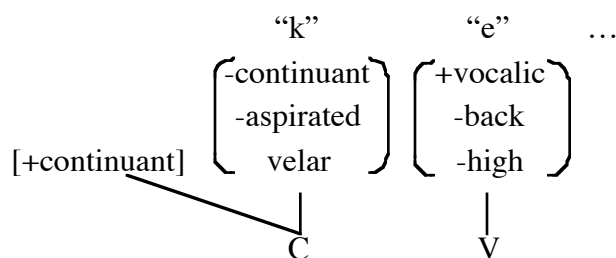
Intransitive	Transitive	
tʌŋzʌlʌɖ	rʌŋzʌlʌɖ	‘weigh (s.th.)’
qʰavud	χavud	‘warm (s.th.) up’
kesqod	yesqod	‘burn something/oneself’
pʌkzɖ	vʌkzɖ	‘lose something / get lost’

OBSERVATION: Place of articulation the same for both members of each pair, what changes is manner – stop vs. fricative.

Assume the intransitive forms are basic and that there is a morpheme marking transitivity (this is not uncommon in less exotic forms).

- (38) $\begin{bmatrix} \text{[+continuant]} \end{bmatrix}$ “transitive” prefix The phonological form of this prefix is just the feature [+continuant].

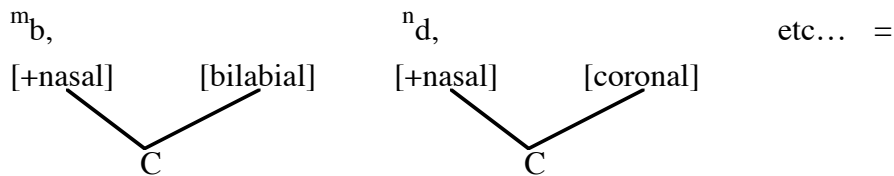
NOTE that the feature is not linked to any skeletal information.



A phonological note: One should think of the voicing alternations in terms of a feature like “aspiration” (really, Voice Onset Time). For stops, +aspirated and –aspirated have their respective values, while for fricatives, +aspirated means voiceless and –aspirated means voiced.

(39) Another Example of a Floating Feature:

Many languages have “prenasalized” stops. These are, in effect, affricates – two sets of features trying to share a single “C” on the skeletal tier (see Unit 6, ex. (11)). (There are no pre-nasalized fricatives in this data).

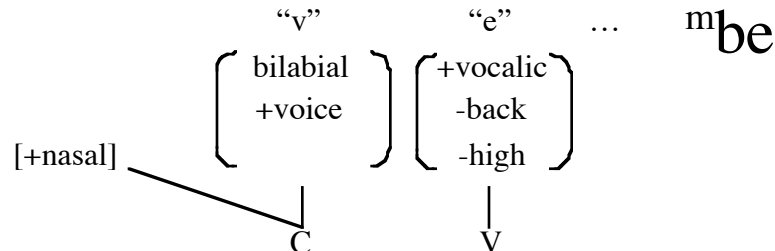


(40) Kihehe (Bantu, Africa) – 1st singular agreement on verbs

ku-vee ^m bela	‘to mourn’	kuu- ^m bee ^m bela	‘to mourn me’
ku-tesa	‘to hurt’	kuu-nesa	‘to hurt me’
ku-lima	‘to cultivate’	ⁿ dima	‘I’ll cultivate’
ku-teleka	‘to cook’	neleke	‘May I cook?’

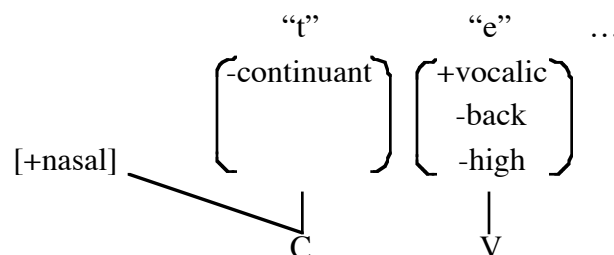
(41) There are two things going on here.

- a. When the first consonant of the root is not /t/, the result is a prenasalized stop with the place of the consonant but the nasal feature of the morpheme.



The [+nasal] feature is itself the morpheme for [1sg] (or more precisely, the allomorph of the abstract agreement morpheme, in the context of [1sg]). Crucially, it has no C slot of its own, and shares one with the following consonant, creating a prenasalized stop.

- b. When the first consonant is a /t/, the result is an /n/.



For our purposes, what matters is that the features of “t” don’t use up the C slot, and the “mixture” of the features defining [t] and the [+nasal] feature share the slot to define [n].

<i>aside</i>	This can be analysed as above on the assumption that [t] is a default stop consonant in this language (there is evidence for this cross-linguistically). So in this case, the [n] surfaces linked to a C on the skeletal tier, but it does so parasitically.
--------------	--

7.3.2 Vowel change/ ABLAUT revisited

Do we want to treat English (or German) as having non-concatenative morphology, floating features, etc...?

(42) English plurals (MUTATION)

leaf	[lijf]	leaves	[liv-z]
shelf	[...f]	shelves	[...vz]
house	[hʌws]	houses	[hæwzəz]

(43) English verbs (ABLAUT)

<u>pres.</u>	<u>past.</u>	<u>part.</u>
sing	sang	sung
bind	bound	bound
leave	left	left
buy	bought	bought
tell	told	told
flee	fled	fled

- There are two important reasons NOT to treat English as involving floating features or a separate vowel tier.

(44) Compare the English case to the Nivkh case.

In Nivkh, we could say that the feature [+continuant] IS the phonological shape of the morpheme meaning “transitive”. If we tried the same in English, we would have to say that the feature [+voice] IS the morpheme meaning plural (or an allomorph of it). But what would that predict?:

(46) Detectable Morphemes (Voice):

Active:	a	The vocalization (or melody) is meaningful.
Passive	u i	
...		

(47) English:

present	past	present	past	present	past
sing	sang	hang	hung	run	ran
tell	told	hold	held		

(48) If we wanted to treat the vowels as morphemes, as we do in Arabic, what would they be?

Present:	i, a, u, e, o
Past:	a, u, e, o

The choice of vowel in Classical Arabic is systematic. It correlates with “voice”.
The choice of vowel in English is arbitrary. There are no correlations to be made.

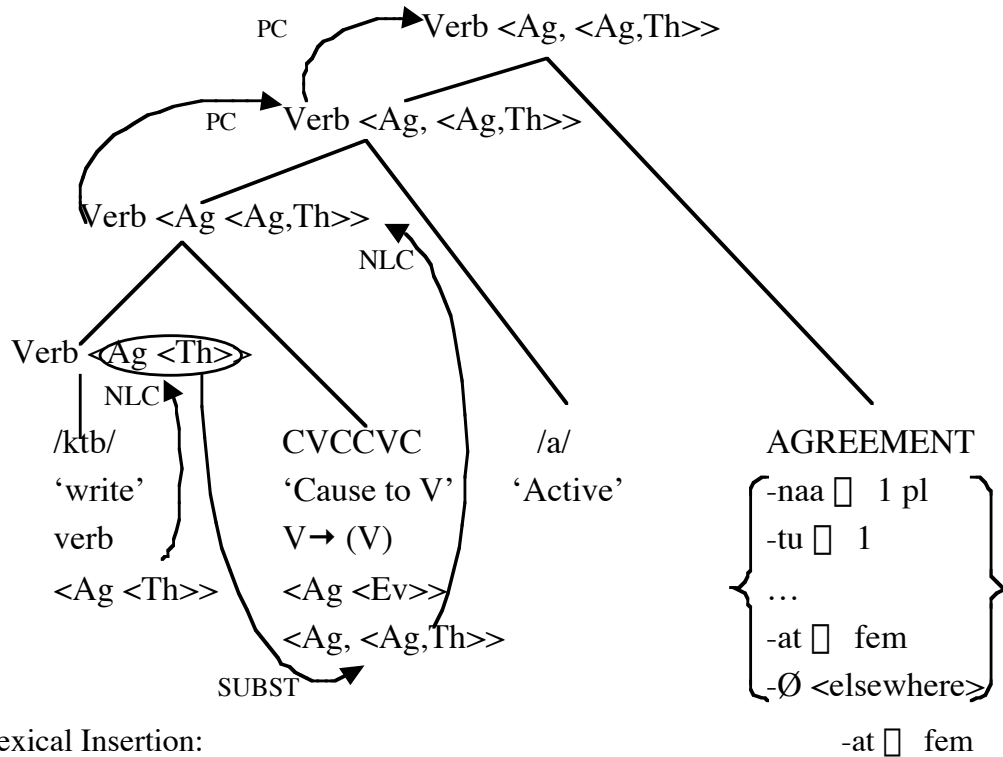
8. PUTTING IT ALL TOGETHER

(49) Classical Arabic verbs in (partial) context:

	Active (past)	Passive (Past)	Cause to X (active)
1sg	katab-tu	kutib-tu	kattab-tu
1pl	katab-naa	kutib-naa	kattab-naa
2sg(masc)	katab-ta	kutib-ta	kattab-ta
2pl (masc)	katab-tum	kutib-tum	kattab-tum
3sg (fem)	katab-at	kutib-at	kattab-at
3pl (fem)	katab-na	kutib-na	kattab-na
3sg (masc)	katab	kutib	kattab
3pl (masc)	katab-uu	kutib-uu	kattab-uu

(50) Can we draw word structure trees for these?

e.g.: kattabat ‘She caused someone to write something’



(51) A complex English example:

Consider the following word: parks managers [párks.mænəʒərz]

What do we know about how this word gets its properties?

Can we draw a tree for it?