CHAPTER 11
Suprasegmental Structure
PREVIEW

Phonological rules and representations also include structures above the level of the segment. Such structures include syllables, sub-constituents of syllables, and the organization of syllables into higher rhythmic units. This chapter reviews aspects of suprasegmental phonology, and some of the evidence for such structures. Whereas segmental properties can be largely "heard", suprasegmentals can only be inferred based on the effect that these structures have on segmental aspects of phonology, which has significant implications for the theory of human reasoning abilities.

KEY TERMS
syllable
foot
timing
mora
skeleton

Another aspect of nonlinear representational theory is the idea that there are phonologically significant structures above the level of the segment, units which encompass multiple segments. Such structures are often referred to as being "prosodic", a term refering to poetic meter, rhythm and singing, which are aspects of language use that involve "how strings of segments are performed".

11.1 Syllables

The best-known unit of prosody, one know to most speakers of English from primary school, is the traditional concept of the **syllable**. The term itself is one of the oldest in linguistics, originating from Ancient Greek *sullabe*, but the nature of the syllable and arguments for it have long been elusive. At various points in contemporary linguistics, scholars have rejected or embraced the syllable, and the syllable was not part of standard generative phonological theory, until 1976 when Kahn produced strong arguments for it within autosegmental theory.

The intuitive concept of "syllable" is not particularly difficult to grasp: it is a string of segments centering around one or more vowels, including some consonants to the left and to the right. The challenge resides in justifying the addition of this concept to our arsenal of analytical devices. In segmental representations, there are audible consequences of features, for example you can hear voicing, nasality and glottalization on segments, even though relating features to specific phonetic properties is difficult. Voicing may be manifested as positive presence of vocal fold vibrations, but also as a difference in the timing of vocal fold vibrations with respect to vowels – it is still about vocal fold vibrations. The best evidence for features is how they explain what phonological rules do, but we tend to be more comfortable with entities that we can directly perceive.

The problem of the syllable is that it has *no* audible defining property, no organ of articulation being activated, thus it cannot be justified as a prima facie transcriptional and physical fact. No amount of ear training will enable you to "hear" how many syllables there are in a word of the form [CVVVVC] in an unfamiliar language, and in [VCCCCV] you cannot "hear" where one syllable begins and the other ends. The evidence for the syllable is indirect, in that grouping sequences of segments into a unit can lead to simpler rules for phonological rules in numerous languages.

Possible consonant clusters. One of the most widely-invoked arguments of this nature regards the rules for possible consonant clusters. This reflects the fact that sequences of segments have to be organized into definable syllables, and languages impose various restrictions on how syllables can be formed. No language allows words to be composed of completely arbitrary permutations of the consonants and vowels in its segmental inventories, all languages have some kind of restrictions of possible sequences.

We start with possible word-beginnings and word-ends in English, seeing how they relate to syllables. Initial clusters can have the form sC (C=consonant), as in *stick, spit, skunk*, also *snow, smite, slay*, or they may be of the type OR (O=obstruent, R=glide or liquid) as in *fray, through, fly, bleed, breed, pray, clue, cube*. The longest possible initial cluster has the shape sCR (*sprint, sklerotic, strip, splice, squint*), which reflects the interaction of these two rules pertaining to possible initial consonant clusters. No words begin with *smpr…*, or *lk…* or any of an innumerable set of similar consonant sequences not allowed by these two rules.

Words violating these rules cannot be words of English, thus consonant plus stop clusters other than sC are both non-existent and judged by native speakers as being impossible (**bnick, *pnort, *ptack, *dbonk, *fnilge*). Likewise there are no stop+fricative clusters (**kfimp, *ksunk*, **pthing*). Sonorants as the first member of a cluster are also excluded: **mbop, *rtot, *lfay, *yluck, *wnurge*. People often think that it is "too hard" to probuce those sequences, but the difficulty is not with the labor of moving articulators, it is with violating the rules of the language. People speaking languages where *mbop, rtot* are possible words don't have bettwe vocal tract anatomy, they have different grammatical systems. There are additional, more specific restrictions on the pattern of allowed initial clusters. For example, coronal plus *l* is excluded (**tluth, *dlifficult, *thlash, *chlort*), except for [sl] (*sleep*) thanks to the special rule allowing sC clusters. Sequences of labial+w are also disallowed (**pwang, *bwint, *mwerge*, **fwet*).

Clusters of consonants at the end of English words are also subject to restrictions. Any consonant except *h* can stand at the end. Final consonant clusters can be of the form sonorant+consonant. Thus, words can end with glide+consonant (*height, clown, mouse, leaf*), liquid plus consonant (*halt, harp, hart, bilk, false, film, born, farm, carl*), or nasal+consonant (*dance, runt, punk, brand, lamp, lymph, lense*). There are certain restrictions on such final clusters. One is that in a nasal plus voiced stop cluster, the stop must be coronal, therefore *fringe, hand* are allowed and *[læmb], *[hæŋg] with pronounced final [b], [g] are disallowed. The consonants [r j w] cannot be the second consonant in a cluster; [l] can follow [r j w] but not a nasal, and nasals can only follow [r j w l].

Certain sequences of voiceless obstruents are also allowed, as long as either the second consonant is [+anterior,+coronal] (*apt, act, depth, apse, raft*), or else the first consonant is /s/ (*cast, cask, clasp*). Obstruent sequences ending in a non-coronal or non-anterior consonant are excluded (**atp, *atc, *lupsh, *ratf*), as are clusters of fricative+obstruent where the fricative is not *s* (**cashk, *lithp,*rafk*). Clusters ending with voiced obstruents are also disallowed (**abd, *abz*). Notice that all of these rules involve allowed or disallowed sequences of two consonants. There are no special rules of combination which specifically apply to just three-member clusters or four-member clusters, and observed limits on initial and final clusters all reduce to a chain of limits on two-consonant sequences. It is also important to note that certain otherwise excluded clusters do arise when inflectional affixes are added, for example the final cluster [bz] exists in the plural *cabs* and [gd] exists in past tense *flagged*, but such clusters only exist as combinations of root plus suffix. These sequence restrictions are limits on the lexicon, not the output of the grammar.

The significance of the syllable for understanding these restrictions comes from the fact that these are not just restrictions on how words begin or end, they are restrictions on how *syllables* begin and end. Taken together, the preceding rules for syllable beginnings and endings also define possible word-medial clusters. Examples of allowed word-medial clusters are [tm] in *atmosphere*, [mb] in *camber,* [ʃl] in *Ashley*, [rt] in *barter,* [sb] in *asbestos*, [bn] in *Abney,* [md] in *Camden*, [db] in *Ledbetter*, [ʃk] in *ashcan* and [kf] in *breakfast*. These are not possible initial or final clusters, except that [rt] is a possible final cluster. The reason for this difference in "allowed word-initially", "allowed word-finally" versus "allowed word-medially" is that "word medially" concatenates two syllables, allowing two syllable locations for certain of the consonants. In such cases, the first consonant is the final consonant of one syllable, and the second is the initial consonant of the next syllable – [kæm.bṛ], [bar.tṛ], [æb.nij], [lɛd.bɛ.tṛ], [brɛk.fʌst]. Three-consonant clusters are also possible, for example *bolster, Andrew, hamster, translate, electron, costly*, which can be phonologically understood as a possible syllable-final sequence followed by a possible sylable-initial sequence, viz. [bol.stṛ], [æn.druw], [hæm.stṛ], [træn.slejt], [ʌ.lɛk.tran], [kast.lij].

Now consider some illicit three-consonant medial clusters, exemplified by **catmbop* (*[tmb]), **fishrtot* *[ʃrt], **gasbnick* (*[sbn]), **lamdbonk* (*[mdb]), **gushkfimp* (*[ʃkf]). The individual consonant pairs entering into these triplets are possible. The sequences [tm], [mb], [ʃr], [rt], [sb], [bn], [md], [db], [ʃk], and [kf] are all possible, but only when the first member is a syllable-final consonant and the second is syllable-initial. The three-consonant cluster *[tmb] is ruled out because *tm* is not a possible syllable-final cluster and *mb* is not a possible syllable-initial cluster, thus *m* cannot by assigned to any syllable. Neither *cat.mbop* nor *catm.bop* follow the English rules for syllabification of consonants. Similarly, *sb* is not a possible syllable-final cluster and *bn* is not a possible initial cluster, therefore the cluster in **gasbnick* cannot be syllabified. A syllable-based analysis of possible clusters automatically predicts the existing restrictions on word-medial 3-consonant clusters. Without the syllable as an organizing unit over segments, a very complex set of additional rules would be required to account for the restrictions on medial clusters.

Syllable-building rules. It is usually assumed that underlying representations do not contain syllable structure, instead syllables are always built by rules. There are many theories of such rules, ranging from most-explicit to most-reduced. A most-explicit system of syllable-construction rules states in the rules themselves exactly how segments are assigned to syllables, and the most-reduced system of syllable construction assumes a rich set of universal grammatical conventions that underly syllable rules, which automatically guide the syllabification rules. We will start with a most-explicit set of rules, deriving English syllables.

The first step is to create a syllable for each vowel and associate the syllable to the vowel. "Syllable" is a representational node notated as σ. Rule (1) inserts σ and associates it to a vowel which is not in a syllable: this is notated as V′, meaning "vowel which is not associated", implicitly meaning "to a syllable".

(1) $\begin{array}{c} \emptyset \rightarrow \sigma \\ \vdots \\ V' \end{array}$ *Syllable creation*

Subsequently, other rules associate unsyllabified consonants with some position in the syllable. To start, *almost* any consonant which comes before a syllabified vowel can be in that vowel's syllable.

(2) $\begin{array}{cc} & \sigma \\ \diagup & | \\ C' & V \end{array}$ *Onset association 1*

The exception is /ŋ/, which can only be associated to the right of the vowel in its syllable. We will temporarily ignore this problem, hoping to find a solution eventually.

Now we divide syllable-initial clusters into two subtypes (governed by two rules): C + sonorant sequences, and [s] + consonant sequences. C + sonorant sequences are composed of clusters like *kw, tw, kj, dr, gl, fl*: a non-sonorant, followed by an oral sonorant. This precludes initial clusters like **θm* (C_2 is not oral), **θk* (C_2 is not sonorant), **ml* (C_1 is a sonorant). It will allow *vl, dl, pw, ðr* which do not seem to exist so we will have to come back to this. We posit (3) to account for these clusters.

(3) *Onset association 2*

$$\begin{array}{cc} & \sigma \\ C' & C \\ [-\text{son}] & [+\text{son},-\text{nasal}] \end{array}$$

Clusters of s+C such as *sp, st, sm* can be syllabified by (4), which allows *s* before any non-continuant.

(4) *Onset association 3*

$$\begin{array}{cc} & \sigma \\ C' & C \\ \begin{bmatrix} +\text{cor} \\ +\text{ant} \\ -\text{voice} \\ +\text{cont} \end{bmatrix} & \begin{bmatrix} +\text{son} \\ -\text{nasal} \end{bmatrix} \end{array}$$

For every vowel in a word, a syllable is created with that vowel in it (we will set aside the question of diphthongs for the moment). The word *pity*, with the vowels /ɪ…i/ therefore has 2 syllables, *spit* and *sprint* have one syllable. Applying Onset Assocation 1 to those words gives $[_\sigma$pɪ] $[_\sigma$ti], s$[_\sigma$pɪ]t and sp$[_\sigma$rɪ]nt – nothing else applies to *pity*. Next, Onset Association 2 puts *p* in the syllable with *r*, thus s$[_\sigma$prɪ]nt. Finally, Onset Association 3 associates *s* to the syllable that follows it: $[_\sigma$spɪ]t and $[_\sigma$sprɪ]nt. We can now syllabify onsets (we get to syllable-final consonants in a moment).

Clusters not following these rules cannot be syllabified, meaning that an imaginary word like **vmig* will not exist, because no rule combines a voiced fricative plus nasal in the onset, likewise **lbock* will not exist because *l* cannot precede any consonant given rules (3) or (4). If such an underlying form as /lbɔk/ were to somehow be surgically implanted in a grammar, the rules would leave *l* unsyllabified, and it is a fundamental assumption that all segments must be associated to some syllable, in order to be phonetically realized, therefore the word would be pronounced as [bɔk], and nobody would know that there is supposed to be an underlying /l/ at the beginning.

Stepping outside of English for a moment, there is variation across languages in how unsyllabifiable sequences are dealt with. Chukchi, Yawelmani and many dialects of spoken Arabic have underlying unsyllabifiable clusters. Chukchi, illustrated in (5), does not allow syllable-initial or final consonant clusters, but underlyingly three- and four-consonant clusters arise within and across morphemes. These clusters are then resolved by inserting a vowel or deleting a consonant, depending on the context. In these examples, epenthetic schwa is indicated by italization and underlying roots are in the first column. In (a) we see schwa inserted between a consonant-final root and a cluster-initial suffix. In (b) we find root-initial clusters which are split up when the root is unprefixed, but not when a vowel-final prefix is added (also there is epenthesis into final clusters like /tŋiwk/ → [t*ə*ŋiw*ə*k]). Finally, word-initial CCC clusters are split up by an epenthetic vowel in (c) when the verb is unprefixed, but root-initial word-medial clusters undergo deletion of the first consonant, and not epenthesis.

> You may rightly wonder why schwa appears between C_1 and C_2 in *wʔenətʔul* but between C_2 and C_3 in *tumγəret* – in Chukchi, vowels are preferentially inserted at morpheme breaks rather that within morphemes.

(5)	a.	/milute/	milutetʔul		'hare meat'
		/wʔen/	wʔen*ə*tʔul		'navaga meat'
		/qejŋ/	qejŋ*ə*tʔul		'brown bear meat'
		/lili/	liliret		'pair of sleeves'
		/qonaγ/	qonaγrat		'set of pants'
		/tumγ/	tumγ*ə*ret		'group of comrades'
			infinitive	*past*	
	b.	/pne/	p*ə*nek	γemnelin	grind
		/tŋiw/	t*ə*ŋiw*ə*k	γenŋiwlin	send
		/tnut/	t*ə*nut*ə*k	γennutlin	swell
		/tlγ/	t*ə*lγ*ə*k	γelγ*ə*lin	melt
	c.	/rtril/	r*ə*tril*ə*k	γetril*ə*k	supply
		/rnr/	r*ə*nr*ə*k	γenr*ə*lin	hold

Languages will then vary as to how they resolve unsyllabifiable clusters. In English, there are no abstract clusters as we find in Chukchi.

Returning to the English pattern, two lines of evidence are frequently called on to support these non-existence predictions that certain sequences are disallowed. The first is plain factual observation: no syllables begin with /rp, mk, ps, kn/ and so on, ergo we may want to encode this pattern in the grammar.

> There *is* an archaic word *knap*, pronounced [næp], which most English speakers do not know.

It is generally conceded that a grammar encodes the linguistically significant knowledge of the language which a speaker-hearer has that enables them to understand and produce an infinity of utterances, but that learned knowledge does not automatically extend to "all current facts about the language" including statistical distribution of segments. We can still scientifically explain the facts outside of the grammar, by appeal to prior historical change, that, had there been a word /knæp/, it was historically restructured to /næp/ and nothing (linguistically cogent) in Modern English would show that original /k/ is still present. Some time over the past 300 or so years, the articulation of initial *kn* in English changed so that *k* was no longer audible, but it remains in the spelling of words like *knife* (and *knap*) These orthographic remnants do not constitute valid evidence for claiming hidden unsyllabifiable consonants in English, and positing such consonants leads to a major increase in analytic indeterminacy – is the underlying cluster /kn, tn, pn, tʃn/ or something else? A second, more powerful argument remains, that English speakers have behavioral problems dealing with such clusters, that they *cannot* pronounce words with such clusters (except for linguists, many of whom can pronounce the otherwise-deleted consonants of *knight, knife* and *psychology*). Even linguists struggle with initial clusters like *lb, rt*, typically inserting schwa somewhere to make the sequence syllabifiable. What makes this behavioral problem linguistically interesting is that speakers don't just have problems pronouncing or hearing [rtuθ], they may have an active judgment that [rtuθ] cannot exist as a word, just as they have an active judgment that *[nozɨs] is incorrect as the plural of *nose*. Though the argument for requiring a grammatical statement of what is non-existent is not overwhelming at least for English syllables, we will assume that some account needs to be included in the grammar.

This brings us to the problem that the rules above make some incorrect predictions, as noted. There are some 2-consonant sequences incorrectly predicted to exist, such an *pw, bw, mw, tl, dl, θl, vw, vr, vr, zr, sr, gj*. We also predict the possibility of [ŋ] before a vowel within the syllable. How then are we to encode these facts in the grammar? In the first generative treatment of syllable-building rules, Kahn (1976) simply states that the rule associates "the longest possible" sequence, not formalizing what is possible. Later approaches have focused on specifying the possible versus impossible, usually invoking some separate condition specific to syllables, that "X is possible", or "X is impossible". Rules generally encode what is possible directly, as in our above rules. We might then add a condition, which is an over-arching guidance saying "but Y is *not* possible", limiting what rules do, without directly saying "but not if" in the rule itself. This approach is followed in Clements & Keyser (1983), which incorporates negative conditions limiting syllabification. Thus conditions such as (6) could preclude sequences like *pw, bw, mw, tl, dl, θl, vw, vr, vr, zr, sr*. They even generalize the notion of condition so that construction of sC initial sequences at the beginning of a syllable is via a positive condition saying that "sC is allowed", and syllable-structure construction then reduces to the universal mandate to construct syllables, under the guidance of language-specific conditions of admissability including a blocking condition against *tl, dl* onsets.

(6) [$_{\sigma}$ [-son] [+son,-nas] is admissible

[$_{\sigma}$ [+cor,-cont] [+lat] is inadmissible

This type of condition-based thinking about grammars gained popularity in phonology under the impetus of Autosegmental Phonology, eventually leading to a theory of language – Optimality Theory – which has no rules at all, just conditions on a universal output-generating device, as we discuss in the final chapter.

Another possibility within an explicit rule-based approach would be syllable-expulsion, that is, allow the maximally general rules to construct syllables, but then repair the results by detaching consonants which stand in forbidden positions. In this approach, hypothetical [$_{\sigma}$dla] is converted to d[$_{\sigma}$la] with *d* detached from the syllable so that we no longer have an illicit sequence at the beginning of the syllable.

(7) *Lateral expulsion*

σ

x

[+cor,-cont] [+lat]

Moving to syllable codas, we next need a set of rules that syllabify consonants standing to the right of the vowel, given that so far we can only syllabify [$_{\sigma}$trɪ]k, [$_{\sigma}$spæ]ŋk, [$_{\sigma}$bɛ]lt, [$_{\sigma}$æ]ks. We allow any single consonant immediately after the syllabified vowel, except for [h] (which may famously lead some to posit that [h] and [ŋ] are positionally-determined allophones, solving two problems, but we will not pursue that idea here). A coda-association rule like (8) may be posited, which links a free post-vocalic consonant to the syllable of the preceding vowel.

(8) *Coda association 1*

σ

V C′

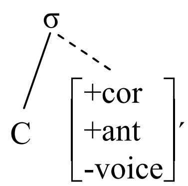
Final sequences of two consonants are possible when the first consonant is a sonorant, for example *shark, belt, camp, false, twelve, band, lunge, prank*, pointing to another rule of syllable-building.

(9) *Coda association 2*

σ

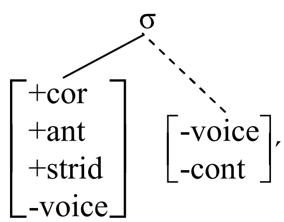
[+son] C′

Another class of two-consonant sequences is exemplified by the words *tract, axe, craft, depth* which have an obstruent as the first consonant and an alveolar obstruent (*t, s, θ*) as the second consonant. In light of (9) which has already created sonorant-plus-C clusters, no condition needs to be put on the first consonant in the sequence, we just put conditions on the second consonant.

(10) 

Coda association 3

We must also allow clusters of *s* plus stop, such as *mask, past, grasp*.

(11) 

Coda association 4

The operation of the two-consonant rules overlap in words like *preempt, glimpse, twelfth* giving us a certain range of three-consonant sequences. The sequences *mp, lf* are syllabified by (8) and (9), and *pt, ps, ft* are syllabified by (10).

These conditions also automatically account for a dependency between long vowels or diphthongs and following consonant sequences. Note that *glɪmps, bɪldʒ, pajnt* are words of English but **glajmps, *bajldʒ* are not. This is as expected if we treat "long vowels" and "diphthongs" of English as sequences with a glide (a consonant) as the second member. That consonant consumes the first coda position, what can follow must observe the restrictions of rules (9) (any consonant after a sonorant), allowing syllabification of [bajl]. But by (10), – only an alveolar obstruent can be further added. Would-be **bajldʒ* is therefore blocked at the step of associating *dʒ*.

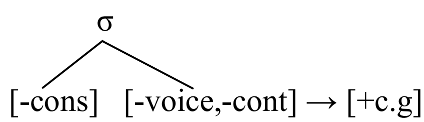
Just as onsets can be constructed either by explicit rules or by universal structure-building conventions controlled by language-specific conditions on association, final consonant sequences may be built either by explicit rule or by general convention guided by language-particular constraints. There are many more details of English syllable structure deserving deeper investigation, the preceding presents a simple overview of the systematicity of syllables in English along with some facts that suggest the need for rule-external machinery.

Phonological rules. Rules of English consonant allophony discussed in Chapter 3 also supports the postulation of the syllable, insofar as those rules are best stated with reference to the syllable. The best-known such rule is the aspiration rule. As is commonly recognised and explicitly assumed in our previous discussion of the aspiration rule, voiceless stops are aspirated at the beginning of a syllable, explaining the aspiration in [p^{h}ɪt, p^{h}lat, ə.ˈp^{h}ɪr, ʌ.ˈp^{h}laj] but not in [spɪt, splɪt, ʌ.ˈspɛ.rə.gəs, slæp, æpt]. Stating this rule becomes much more difficult if there is no such thing as a syllable.

Another rule of American English which refers to the syllable glottalizes syllable-final voiceless stops, whereby /p t k/ become unreleased glottalized [p˺ t˺ k˺] after a vocoid in the same syllable. There is dialectal variation in the extent to which all voiceless consonants undergo this rule, but examples involving *t* (the most susceptible to glottalization) include *hi**t**, hear**t**, ca**t**kin, ligh**t**, clou**t**, heigh**t**s, hear**t**s, a**t**las, a**t**lantic* and *Wa**t**ney's* are valid in most American dialects. By contrast, there is no glottalization of *t* in *s**t**em, ap**t**, bel**t**, ma**tt**ress, a**t**rocious*. In the word *stem*, *t* is clearly not preceded by a vocoid at all, so the conditions of the rule are not satisfied: likewise in *apt* and *belt*. In *mattress, atrocious*, the cluster *tr* is a cluster at the beginning of the second syllable, so while *t* is preceded by a vocoid, it is not in the same syllable as the vocoid. Consequently, there is no glottalization in these examples. On the other hand, there is glottalization in *atlas, atlantic* since **tl* is not a permitted initial cluster in English, so these words are syllabified as *at.las, at.lan.tic*. Likewise *tn* is not an allowed cluster at the beginning of the syllable, so *Watney's* is syllabified *Wat.ney's*. Since *t* is in the same syllable as the preceding vocoid, the consonant becomes glottalized.

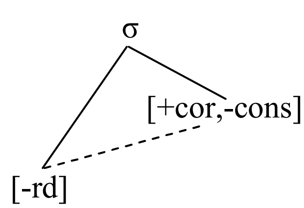
The rule deriving glottalized consonants can accordingly be formulated as (12), the relevance of the syllabis is that the triggering non-consonantal preceding must be in the same syllable as the affected consonant.

(12)



r-unrounding. A third rule of English phonology providing evidence for the syllable is the one which pertains to rounding of *r*. In some dialects, *r* is realized both as a rounded and an unrounded rhotic approximant, [ɹ] and [ɹ^{w}], following the rule that /ɹ^{w}/ unrounds after a nonround vowel in the same syllable. Thus *r* is round in [ɹ^{w}ejɲd^{ʒ}] *range*, [t^{h}ɹ^{w}ej] *tray,* [stɹ^{w}ej] *stray,* [fɹ^{w}ej] *fray*, also in [kɔɹ^{w}] *core,* [tʊɹ^{w}] *tour* where the vowel preceding *r* in the syllable is round, and in [ʌˈɹ^{w}ej] *array* where the preceding vowel is in a separate syllable; but, *r* is unrounded in [kaɹ] *car,* [kaɹt] *cart,* [ˈbɪɹ] *beer,* [hɛɹld] *Harold*. The following rule unrounds /ɹ^{w}/ after a tautosyllabic non-round vowel.

(13)



Vowel reduction. Vowel reduction provides another argument for the syllable in English. The data below show, as we observed in chapter 5, that unstressed vowels reduce to schwa.

(14)

	Reduced		*Unreduced*	
	[əˈlæw]	‘allow’	[ˈælow]	‘aloe’
	[əˈnɔj]	‘annoy’	[ˈænəlɪst]	‘analyst’
	[təˈlɛgrʌfij]	‘telegraphy’	[ˌtɛləˈgræfɪk]	‘telegraphic’

A simple statement like “an unstressed vowel becomes schwa” forms the core of the correct generalization, but the following data show that the matter is more complex, since the nature of the following consonants matters. In some cases, a CC cluster can stand between the target of reduction and the next vowel, but in other cases, a CC cluster blocks reduction.

(15)

	Reduced		*Unreduced*	
	[əˈbrʌpt]	‘abrupt’	[ædˈmanɪʃ]	‘admonish’
	[əˈtrowʃəs]	‘atrocious’	[ætˈlæntɪk]	‘atlantic’
	[əˈstranəmij]	‘astronomy’	[ænˈdijən]	‘Andean
	[əˈfrejd]	‘afraid’	[arˈtɪstɪk]	‘artistic’
			[ælˈpækə]	‘alpaca’

If we take cognizance of syllable boundaries, especially the ends of consonant clusters that are allowed in the beginning of the syllable, then the generalization becomes much clearer: unstressed vowels reduce to schwa in English when they are at the end of the syllable.

(16)

	Reduced		*Unreduced*	
	[*ə*.ˈbrʌpt]	‘abrupt’	[*æ*d.ˈma.nɪʃ]	‘admonish’
	[*ə*.ˈtrow.ʃəs]	‘atrocious’	[*æ*t.ˈlæn.tɪk]	‘atlantic’
	[*ə*.ˈstra.nə.mij]	‘astronomy’	[*æ*n.ˈdij.ən]	‘Andean
	[*ə*.ˈfrejd]	‘afraid’	[*a*r.ˈtɪ.stɪk]	‘artistic’
			[*æ*l.ˈpæ.kə]	‘alpaca’

Other phenomena referring to the syllable. Across languages, there has been a recurring puzzle regarding the expression of natural classes via features, and the role of word boundaries. The problem is that there exist many rules which treat a consonant and a word boundary alike, but only for a specific set of rules. Many dialects of Arabic have such a rule, one of vowel epenthesis which inserts [i] after a consonant which is followed by either two consonants or one consonant and a word boundary. Thus in many dialects of Eastern Arabic, underlying /katab-t/ becomes [katabit] ‘I wrote’ and /katab-l-kum/ becomes [katabilkum] ‘he wrote to you pl.’. The following rule seems to be required, in a theory which does not have recourse to the syllable.

(17) $$\varnothing \rightarrow [\text{i}] \;/\; \text{C} \_\_ \text{C} \begin{Bmatrix} \text{C} \\ \# \end{Bmatrix}$$

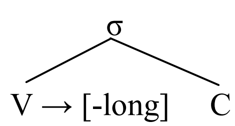
Similarly, a number of languages, such as Yawelmani (chapter 7), have rules shortening long vowels when followed by two consonants or by a word-final consonant (thus

/taxa:k’a/ → *taxa:k* → [taxak] ‘bring!’, /do:s-hin/ [doshin] ‘report (nonfuture)’), which would be formalized as follows.

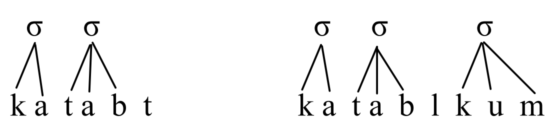
(18) $[+\text{syl}] \rightarrow [-\text{long}] \;/\; \_\_\; C \begin{Bmatrix} C \\ \# \end{Bmatrix}$

The problem is that these rules crucially depend on the brace notation (“{..., …}”) which joins together sets of elements having nothing in common, a notation which has generally been viewed with extreme skepticism. But what alternative is there, since we cannot deny the existence of the phenomena?

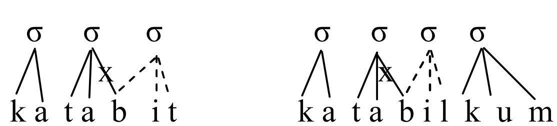
The concept of syllable provides an alternative way to account for such facts. The thing that clusters of consonants and word-final consonants have in common is that in many languages syllables have the maximal structure CVX, therefore in /ta.xa:k/ and /do:s.hin/ where there is shortening, the long vowels have in common the fact that the long vowel is followed by a consonant – the syllable is “closed”. In contrast, in [do:.sol] ‘report (dubitative)’, no consonant follows the long vowel. Expressed in terms of syllable structure, the vowel shortening rule of Yawelmani (and many other languages) can be expressed quite simply without requiring recourse to brace notation.

(19)


The case of of a rule like (17), vowel epenthesis, seems to have a functional motivation of avoiding otherwise unsyllabifiable clusters like /katabt, katablkum/ by inserting a vowel which creates a syllable to which the problematic consonant can associate. We would otherwise predict the following structures.

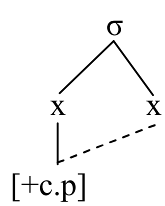
(20)


One may posit a rule inserting [i] before an unsyllabified consonant: the utility of syllable structure is that it simplifies and unifies the description of when the vowel is inserted. This does come at a cost, however, because syllable-structure then has to be re-arranged to accommodate these new vowels and their syllables.

(21)


If syllable structure is the result of language-specific rules, this is a significant problem since it requires “both before and after” ordering of rules, contrary to the fundamental claim of rule ordering. But if syllable structure is an automatic, everywhere response to a universal mandate to syllabify and the language-specific content of syllabification is only the template of well-formedness regulating which structure is assigned, this ordering problem goes away, since universal conventions are not rules.

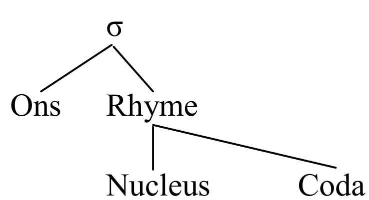
Another type of argument for the syllable is the domain argument, examples being the arguments from English glottalization and r-unrounding where the fact of being in the same syllable is a crucial condition on the rule. One example comes from Cairene Arabic, where pharyngealization spreads to all segments in the syllable (originating from some coronal sonsonant – *t* and *tˤ* are contrastive phonemes in Arabic, idem *d* and *dˤ*, *s* and *sˤ* and in some dialects *r* and *rˤ*). Pharyngealization also affects vowels via this pharyngealization-spreading rule. Examples of this distribution are [rˤaˤbˤ] ‘Lord’ from /rˤab/vs. [rab] ‘it sprouted’; [tˤiˤ:nˤ] ‘mud’ from /tˤi:n]/ vs. [ti:n] ‘figs’; see especially the alternation [lˤaˤtˤiˤ:fˤ] ‘pleasant (m)’ ~ [lˤaˤtˤiˤ:fa] ‘pleasant (f)’ from /lˤatˤi:f/. The addition of the feminine affix /-a/ has the consequence that the root-final consonant is syllable final in the masculine, but initial in the following syllable in the feminine. The rule of pharyngealization is formalized in (22).

(22) 

(mirror-image)

Because of the syllabification differences between /lˤa.tˤi:f/ and /lˤa.tˤi:.fa/, *f* is subject to the rule only in the masculine, despite the fact that the conditioning factor, a vowel with the pharyngealization feature (derived by spreading pharyngealization from the syllable-initial consonant) is immediately adjacent to the consonant in both cases.

Internal structure to syllables. The traditional vocabulary for talking about syllables includes terms for sub-units of the syllable, some of which were invoked in the preceding account. It is often said that the consonants at the beginning of the syllable are in a sub-unit of the syllable, the “onset”. Consonants after the vowel may be organized into a “coda”. The vowel which forms the core of the syllable may be in a structure “nucleus”, and the nucleus plus coda may form a syllable “rhyme”.

(23) 

It is far from settled how much structure exists within the syllable: these are matters for more-focused investigation.

11.2 The Skeleton

At the same time as the syllable was accepted into generative phonology, arguments were made in support of a notion of "timing slot". The classic phonological puzzle which led to this concept was segmental length. Vowels and consonants can come in long and short varieties. The standard theory of length was that there is a feature "long" available to any segment. The problem is that long vowels and consonants typically behave as though they are two segments. A classic example is Lithuanian as discussed in Kenstowicz (1972). Rising and falling tones only appear on long vowels and diphthongs or else syllables ending with a liquid or nasal, and not on short vowels, see for example /kâimas/ 'village', /vǎikas/ 'child' with diphthongs, /kârtis/ 'pole', /kǎrtis/ 'bitterness' with V + resonant sequences, and /matî:ti/ 'to see', /matǐ:s/ 'he will see' with long vowels. There are no level-H-toned syllables with a diphthong, VC or long vowel, and only level H or L (unmarked) appear on short vowels (/mésti/ 'to throw', /dúri:s/ 'door'). This is explained if the language allows just one H on one segment within the word, where coda sonorants can bear tone, and crucially, long vowels are two-segment sequences, thus /matíiti/, /matiís/, /káimas/, /vaíkas/, /kártis/, /kaŕtis/. The featural representation of length makes the explanation of the distribution of H in Lithuanian very complex.

Shortening rules of Lithuanian treat long vowels and diphthongs analogously. A word-final falling-toned long vowel and diphthong will shorten, resulting in alternations in adjectives in their indefinite and definite forms. Examples are /ger-*úo*/ → [ger-ú] 'good (instr.sg.masc.indef.)', /ger-úo-j-*úo*/ → [ger-úo-j-u] 'good (instr-sg-masc-def)', /ger-*áa*/ → [ger-á] 'good (instr.sg.fem.indef.)', /ger-áa-j-*áa*/ → [ger-áa-j-a] 'good (instr.sg.fem.def)'. Long vowels and diphthongs also shorten before syllable-final liquids, cf. [gîr-ti] 'to praise', [gíir-ee] 'he praised' and [pûl-ti] 'to tumble', [púol-ee] 'he tumbled', again pointing to the functional identity of monosegmental long vowels and bisegmental diphthongs. If long vowels *are* two vowels, shortening of vowels and diphthongs is easy to account for.

The online article on Matumbi phonology (Chapter 7) provides numerous lines of evidence that a long vowel is functionally equivalent to two short vowels, and there are various demonstrations of this analogy for consonants. This generalization has been recognized for a long time, the problem was finding a theoretically-sensible account of how that might be so, especially since for some phonological rules, long vowels act like a single segment. For example, a rule of vowel rounding will always apply to both halves of a long vowel, nasalization will apply to both halves, as we expect if a long vowel is one segment.

11.2.1 CV slots

A resolution to this paradox – two, yet one – is available in autosegmental theory. A long vowel or consonant *is* a single segment, but it has two suprasegmental positions. Following the theory of Clements & Keyser (1983) and applying this theory to Logoori [ɪd:í:d^{3}i] ‘wall’ (using colon transcription of length) we have single vowel and consonant segments, but the long segments branch to two C-positions or V-positions.

```
(24)   V  C  C  V  V  C  V
       |   \/     \/   |  |
       ɪ   d      i   d³  i
```

There have been a number of variants of this position, for example positing two metrical positions (see below in §11.3 for metrical structure) or generic X slots not distinguishing V from C, or two segmental root-nodes, which converge at the idea that long segments have two of what short segments have one of.

The phenomenon of compensatory lengthening is cited in support of this two-unit theory of length, drawing an analogy to tone preservation and contour creation. An example is that in Klamath, glottal stop deletes syllable-finally, which causes the preceding vowel to lengthen, e.g. /sleʔ-a/ → [sleʔa] ‘sees’ vs. /sleʔ-ca-a/ → [sle:ca] ‘goes to see’. The problem for SPE-style rules is capturing the connection between vowel lengthening and glottal stop deletion. Only by writing a transformational rule is it possible to connect the two processes.

(25)	<V>	ʔ	C		<1>	Ø	3
	1	2	3	→	+long		

In CV theory, this is another example of preservation of representational structure, with reorganization of existing substance, not addition of new substance, just as the creation of contour tones under vowel merger is the rearrangement of individual input tones.

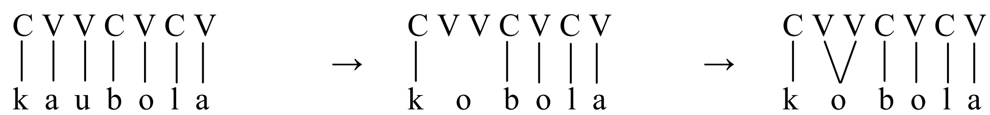
```
(26)   C C V C C V        C C V C C V        C C V C C V
       | | | | | |    →   | | |   | |    →   | |  |/  | |
       s l e ʔ c a        s l e   c a        s l e    c a
```

Matumbi has a rule of vowel contraction whereby the sequence /ai, au/ in prefixes (not in the stem) optionally become [e:, o:] respectively.

(27)	*uncontracted*	*contracted*		
	ba-i-káatite	beekáatite	(ba- ‘they’, -i- ‘it’)	‘they cut it (cl. 9)’
	ba-u-káat-ite	bookáatite	(-u- ‘you’)	‘they cut you’
	ka-i-bolá	keebolá	(ka- ‘when’)	‘when it (cl. 9) rotted’
	ka-u-bolá	koobolá	(-u- ‘it (cl. 3)’)	‘when it (cl. 3) rotted’

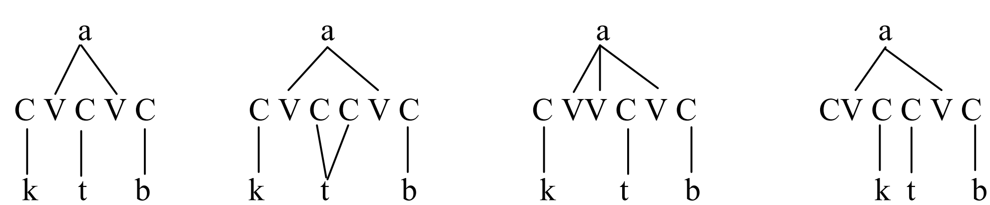
When the high vowel is root-initial, as in [a-ínite] ‘he danced’, no contraction is possible so *[eénite] is ungrammatical. Long [aa] also does not contract, for example /ka-a-i-kaáta/ → [kaaikaáta] ‘when he was cutting it (cl. 9), not *[keekaáta] (which would be ‘when it (cl. 9) was cutting’ ← /ka-i-kaáta/). These data, the restrictions, and the fact of optionality show that vowel coalescence is a single, optional rule. If the process were sub-divided into vowel lowering, lengthening, and low-vowel deletion, we would have to guarantee that the vowel lowering rule is not triggered by a long vowel (not [*kaaekaáta], the same problem with the vowel lengthening rule ([*kaaiikaáta]), as well as low-vowel deletion ([*keekaáta] etc.). Since the process is optional, we would be unable to tie the three rules together so that they all apply, or none apply, and we would predict ungrammatical outputs from /ka-u-bolá/ such as [*kaobolá, *kobolá, *kaoobolá, *kauubolá] where only one or two of the rules applies. This all-or-nothing behavior is precisely what is predicted under an autosegmental theory of length and compensatory lengthening, as exemplified in CV theory. When the segmental sequences /ai, au/ are reduced to monosegmental [e, o], the original V structure is automatically redistributed over the resulting vowel segment, yielding a long vowel.

(28)



Templatic morphology provides another line of evidence for abstract skeletal positions. A number of languages exhibit word-formation processes whereby roots are modified to follow prosodic templates which are easily characterized in CV terms. These skeletal templates are analogous to floating tone melodies as encountered in numerous tone languages, which were the initial impetus for Autosegmental Phonology. Semitic languages especially Classical Arabic provide examples. Verbs in Arabic are described as containing just consonant segments in the lexicon, which are then aligned with general shape templates having derivational meaning. The roots /ktb/ ‘write’ and /ʕml/ ‘work’ combine with the derivational morphemes CVCVC, CVCCVC, CVVCVC and ʔa-CCVC to give *katab-a* ‘he wrote’, *kattab-a* ‘he made to write’, *kaatab-a* ‘he corresponded’, *ʔaktab-a* ‘he dictated’; *ʕamal-a* ‘he worked’, *ʕammal-a* ‘he appointed as governor’, *ʕaamal-a* ‘he dealt with’, *ʔaʕmal-a* ‘he put to work’. Representations of the forms of /ktb/ are as follows.

(29)



One other function of the CV tier is to replace the segmental feature *syllabic* which distinguishes vowels from glides and figures into “syllabic sonorants” as exist in a number of languages. This is the essential function of having V. Syllables are initially created to house V segments, C segments then get added to those syllables.

11.2.2 Kamba Empty C

Some languages provide evidence for a skeletal concept analogous to floating tones, namely an empty skeletal position. An example is the evidence for a “floating C” in Kamba. The data in (30) motivate the underlying forms of various verb prefixes as they appear before a consonant. Kamba has 4 tone levels, extra-high [a̋] (XH), high [á] (H), low [a] (L) and extra-low [à] (XL).

(30)	ko-ta̋l-a̋	‘to count’	ko-kon-à	‘to hit’
	ko-ké-ta̋l-a̋	‘to count it (cl. 7)’	ko-ké-kon-à	‘to hit it (cl. 7)’
	ko-má-ta̋l-a̋	‘to count them (cl. 2)’	ko-má-kon-à	‘to hit them (cl. 2)’
	ko-mo-ta̋l-a̋	‘to count him (cl. 1)’	ko-mo-kon-à	‘to hit him (cl. 1)’
	ko-kɛ̋ɛ̀ðj-à	‘to greet’	ko-n̪ɛɛ̀n̪-à	‘to lick’
	ko-má-kɛ̋ɛ̀ðj-à	‘to greet them (cl. 2)’	ko-má-n̪ɛɛ̀n̪-à	‘to lick them (cl. 2)’
	ko-mo-kɛ̋ɛ̀ðj-à	‘to greet her (cl. 1)’	ko-mo-n̪ɛɛ̀n̪-à	‘to lick her (cl. 1)’

In the infinitive, the final suffix /-à/ has XL tone, which spreads to a long penultimate vowel in [kon̪ɛɛ̀n̪à] and [kokɛ̋ɛ̀ðyà] from /kon̪ɛɛn̪à, kokɛ́ɛðyà/. Then H in the penultimate syllable becomes XH before XL thus [kokɛ̋ɛ̀ðyà], also /kotálà/ → *kota̋là*. Finally, word-final XL assimilates to XH when the preceding vowel is XH, so /kotálà/ → *kota̋là* → [kota̋la̋]. The XL tone of /-à/ is deleted phrase-medially, resulting in the following forms which show that the tone complications in (30) are brought about by final extra-low.

(31)	kotála…	‘to count…’	kokona…	‘to hit…’
	kokétála…	‘to count it (cl. 7)…’	kokékona…	‘to hit it (cl. 7)…’
	kokɛ́ɛðja…	‘to greet…’	kon̪ɛɛn̪a…	‘to lick…’

These tone patterns are crucial to identifying the surface distinctions that argue for empty C-slots.

When a verb stem begins with a vowel, rules modify the resulting V+V sequence. The imperative reveals the underlying initial vowel (not necessarily its underlying length). If a vowel-initial stem is preceded by /a/, the V+V sequence merges into a long lowest-height version of the second vowel, as long as the second vowel is not a high vowel. In these examples, the vowel /a/ is from the plural object prefix /má/. H verbs are marked below by underlining the root-initial vowel.

(32)	asjâ	‘tell!’	koma̋àsjà	‘to tell them’
	<u>á</u>ka	‘build!’	koma̋a̋ka̋	‘to build them’
	etâ	‘spill!’	komɛ̋ɛ̀tà	‘to spill them’
	<u>é</u>ta	‘respond!’	komɛ̋ɛ̋ta̋	‘to respond to them’
	<u>ɛ́</u>nza	‘shave!’	komɛ̋ɛ̋nza̋	‘to shave them’
	ɛndâ	‘like!’	komɛ̋ɛ̀ndà	‘to like them’
	<u>ó</u>mba	‘mold!’	komɔ̋ɔ̋mba̋	‘to mold them’
	okíta	‘fight!’	komɔ́ɔkità	‘to fight them’
	<u>ɔ́</u>na	‘see!’	komɔ̋ɔ̋na̋	‘to see them’
	ɔngɛ́la	‘increase!’	komɔ́ɔngɛlà	‘to increase them’
	<u>ú</u>na	‘fetch!’	komáűna̋	‘to fetch them’

í̲ta	‘strangle!’	komái̋ta̋	‘to strangle them’
iâ	‘put!’	komáià	‘to put them’

The imperative has no XL tone at the end, therefore the complex tonal changes of (30) are lacking. Instead H is assigned to the second vowel of the verb, with the complication that H verbs such as /áka, ɛ́nza/ having only two stem vowels do not add that H. Underlying /ko-má-età/ first undergoes vowel fusion to give *komɛ́ɛtà*, then XL spreads to the long penult (*komɛ́ɛ̀tà*), after which H becomes XH when on the vowel immediately before XL, giving surface [komɛ̋ɛ̀tà] ‘to spill them’. In minimally-contrastive /ko-má-étà/, fusion results in *komɛ́ɛ́tà*, and with the assumption that the two Hs on the resulting long syllable are a single tone linked to two positions, we instead derive *komɛ̋ɛ̋tà* by H-raising and [komɛ̋ɛ̋ta̋] by rightward XH spreading, analogous to /kotálà/ → *kota̋là* → [kota̋la̋].

If the vowel preceding the root is /o, e/ the first vowel becomes the corresponding glide [w, j], illustrated with the infinitive prefix /ko/, /mo/ ‘him, her’, and the class 7 object prefix /ké/. Resulting *kj* then becomes *tʃ*, though derived *j* first deletes before /i,e/, just as *w* deletes before /u,o/.

(33)	kwaàsjà	‘to tell’	/ko-asjà/
	kwa̋a̋ka̋	‘to build’	/ko-ákà/
	kotʃa̋a̋ka̋	‘to build it’	/ko-ké-ákà/
	kwɛ̋ɛ̋nza̋	‘to shave’	/ko-ɛ́nzà/
	kwe̋e̋ta̋	‘to respond’	/ko-étà/
	kotʃe̋ètà	‘to spill it’	/ko-ké-età/
	kőőmba̋	‘to mold’	/ko-ómbà/
	kotʃóokità	‘to fight it’	/ko-ké-okità/
	kotʃɔ̋ɔ̋na̋	‘to see it’	/ko-ké-ɔ́nà/
	kwɔɔngɛlà	‘to increase’	/ko-ɔngɛlà/
	kotʃɔ́ɔngɛlà	‘to increase it’	/ko-ké-ɔngɛlà/
	kűűna̋	‘to fetch’	/ko-únà/
	komúúneðjà	‘to make fetch her’	/ko-mo-ún-eðjà
	kotʃúúneðjà	‘to make fetch it’	/ko-ké-ún-eðjà
	kwi̋i̋ta̋	‘to strangle’	/ko-ítà/
	koki̋i̋ta̋	‘to strangle it’	/ko-ké-ítà/

In these examples, we observe a rise-simplification tone rule, for example in /ko-ɛ́nzà/ → [kwɛ̋ɛ̋nza̋] ‘to shave’. By analogy to /kotálà/ → [kota̋la̋] ‘to count’ and /ko-mo-súkumà/ ‘to push him’, we expect *[kwɛɛ̋nza̋] ‘to shave’ and *[komuúneðjà]. Kamba does not have rising tones, so whenever the combination L+H or L+XH would arise by tone-assignment or syllable-fusing rules, we instead find level H or XH, by spreading the syllable-final tone to the left within a long syllable. For an example of the rise-to-H rule as a biproduct of direct tone assignment, we saw in (32) that H is assigned to the second stem vowel in the imperative. If the first syllable of a L verb has a short vowel, that means the H is on the second syllable. When the first syllable is long, the entire syllable has a level H tone.

(34)	kokaaàðà	'to praise'	kááða	'praise!'
	koliìndà	'to cover'	líínda	'cover!'
	konuungeà	'to smell'	núúngea	'smell!'
	kotaandeðjà	'to trade''	táándeðja	'trade!'

A number of roots which seem to begin with a vowel flout these generalizations, in that vowel coalescence and glide formation do not apply before / to those vowels. The infinitives in (35) should have changed the prefix /ko/ to [kw] (with lengthening of the following vowel), but surprisingly that does not happen.

(35)	koaljoòlà	'to translate'	koaàmbà	'to make noise'
	koasjà	'to help divide'	koebà	'to pay'
	koɛkà	'to leave'	koɛɛ̀ndà	'to go'
	koiìndà	'to submerge'	koĭ̋ingà	'to drive cattle'
	koɔà	'to bewitch'	koɔ̋ɔ̀tà	'to dream'
	koɔɔ̀kà	'to go early'	koű̃mã̋	'to curse'

A further problem seen in (36) is that there are verbs with apparently initial /o/ which should merge with the preceding prefixal /o/ to give a long vowel. Instead, we find phonetically extra-long and extra-extra long vowels, as transcribed below with a continuum of raised plus-marks. The same surface extra-length problem arises with combinations of certain a-initial verb and a prefix ending in /a/. Up to this point, it has been sufficient to indicate short versus long vowels with standard single versus double vowel writing, however at the phonetic level Kamba has well more than two distinctive degrees of vowel duration a.k.a. "length", it has perhaps 6 degrees of "length".

(36)	koő$^{+}$mã̋	to bite	koo^{+}sjà	to help buy
	kotóo^{+}sjà	to help us buy	koo^{+}lokà	to fly
	koo^{++}ngamà	to stop	komáà$^{++}$mbjà	to help them make noise

Not only are these forms exceptional in presenting extra degrees of vowel length, they run counter to tonal generalizations of the language. We observe a rising tone in [koő$^{+}$mã̋], and final XL which does not spread to a long penult in [koo^{+}sjà]. This only happens in connection with extra-long vowels. In [kotóo^{+}sjà] final XL does not spread to the preceding (extra)-long vowel, but we also see in [komáà$^{++}$mbjà] that *extra*-extra long vowels (with two overlength marks) show the effect of the spreading rule, yet still do not show the raising of H to XH as we did observe in /koké-età/ → [kot^{ʃ}ẽ̋ètà].

This same phonetic over-length anomaly exists for the non-identical vowel clusters in (35) – in terms of phonetic duration facts, more accurate transcriptions would be [koa^{+}ljoòlà, koaà$^{++}$mbà, koe^{+}bà, koɛɛ̀$^{++}$ndà] but this over-length can be read off the fact of there being a sequence of phonetically distinct vowels, so it need not be marked in a sequence like [ea], it only needs to stand in for transcriptions like [eee, ooo, aaaa].

> In some languages with heterorganic vowel sequences there is an amplitude drop between the individual vowels, suggesting re-articulation of individual vowel segments. This does not happen in Kamba: the differences between [kőőma̋, koő$^{+}$ma̋, koo^{++}ngamà] are only in duration.

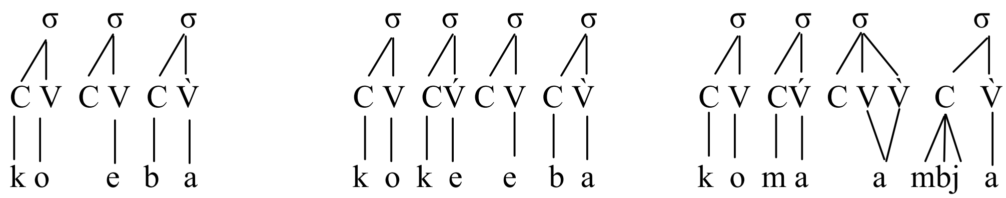
Prefixation of [má, ké] in (37) shows that these verbs systematically resist glide formation and vowel fusion, and pose the same over-length problems as (36).

(37)	komáa^{+}ljoòlà	to translate them	kokéaljoòlà	to translate it
	komáa^{+}sjà	to help them divide	kokéasjà	to help divide it
	komáebà	to pay them	kokée^{+}bà	to pay it
	komáɛɛndɛà	to go for them	kokéɛɛndɛà	to go for it
	komáı̋ìngà	to drive them	kokéı̋ìngà	to drive it
	komáőma̋	to bite them	kokéőma̋	to bite it
	komáɔà	to bewitch them	kokéɔà	to bewitch it
	komáűma̋	to curse them	kokéűma̋	to curse it

Tone patterns point to another anomaly. Normally when a long vowel stands in the penultimate syllable, in the infinitive we find that the final extra-low tone spreads to the second half of the long vowel, so /kon̪ɛɛn̪à/ → [kon̪ɛɛ̀n̪à] 'to lick', /kokɛ́ɛðjà/ → [kokɛ̋ɛ̀ðjà] 'to greet', also /ko-má-asjà/ → ko-máasjà → [koma̋àsjà] 'to tell them', /ko-me-enzà/ → komeenzà → [komeènzà] 'to dig them'. An apparent extra-long syllable in the penult is not the target of this rule, as we see in /ko-má-asjà/ → [komáa^{+}sjà], not *[koma̋à$^{+}$sjà], likewise [kokée^{+}bà] and not *[koke̋è$^{+}$bà]. We do still have to allow the rule to apply to extra-extra-long vowel as in [komáà$^{++}$mbjà].

A single generalization explains the anomalies of this class of verb stems, centering around the fact that the vowel sequences do not collapse into single long vowels. Instead, the component vowels remain in separate syllables, and the reason why they do not collapse into one syllable is that there is a C-slot at the beginning of the verb root, one with no segmental content and no direct phonetic implementation. The syllable-collapsing rules depend on there being a sequence of vowels in separate syllables, but in the examples (35) to (38), there is no vowel sequence at the skeletal level, there is VCV.

(38)



‘to pay’ ‘to pay it’ ‘to pay (dowry) them’

The overlength-marking question can then be resolved by noting that there is a surface difference (in duration) between a disyllabic vowel sequence and a long vowel, even when the vowels are the same quality. We can transcribe these as [ko.ebà, koké.ebà, komá.aàmba]. And we can explain the source of this phonetically and phonologically distinctive difference in syllabification by positing a bare C slot as the blocker of otherwise general vowel-fusion rules.

In (39) we observe infinitives and imperatives of L-toned verbs which appear to have the shape CVVCV, but the group in (39a) still differs from (39b). The verbs in (39a) have ordinary long vowels, but those in (39b) have extra-long vowels; final XL spreads to the long penultimate vowel in (39a) but not in (39b). The group in (39a) has a level H because of the rule changing rising tone to level H, but the verbs in (39b) retain the LH pattern expected in L verbs which do not have a long first syllable.

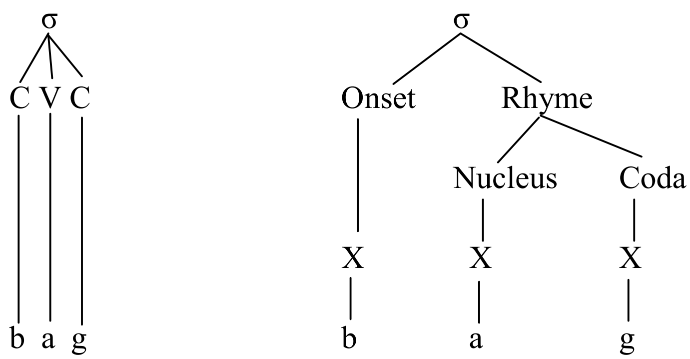
(39)	a.	kobaàndà	‘to plant’	báánda	‘plant!’
		komaànthà	‘to search’	máántha	‘search!’
		koliìndà	‘to cover a hole’	líínda	‘cover!’
		kokaàthà	‘to praise’	káátha	‘praise!’
	b.	koto.otà	‘to weed’	to.óta	‘weed!’
		kou.umà	‘to rumble’	u.úma	‘rumble!’
		koka.atà	‘to go sour’	ka.áta	‘go sour!’
		koko.okà	‘to act wild’	ko.óka	‘act wild!’
		kolu.utà	‘to push off’	lu.úta	‘push off!’

The explanation is straightward, that the root in [kobaàndà] has two syllables, the first containing a long vowel, and in [koto.otà] the root has three syllables but no long vowel: the underlying form is /kotoCotà/.

11.2.4 X Slots

A competing theory simplifies the inventory of skeletal objects to a single object X, where the function served by the distinction C vs. V is replaced with the distinction X versus X marked as the nucleus of a syllable. To the extent that “vowel” status cannot be predicted from segmental features (e.g when there is a contrast between glides and vowels or when the syllabicity of sonorants is not predictable), one must provide some partial syllable subconstituency in lexical representations, for example by specifying that a vowel is pre-specified as a syllable nucleus.

(40)

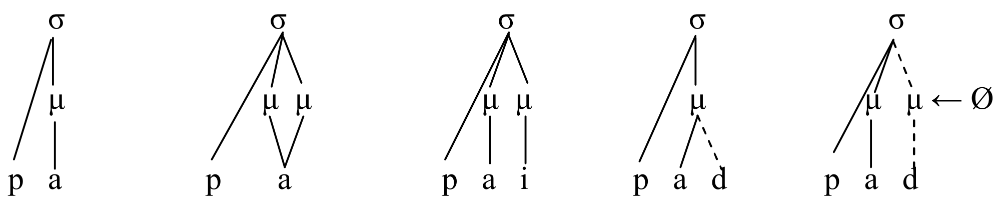


The difference between the CV theory and the generic X theory is minimal. CV theory posits two skeletal units (C and V) and there is an implied relationship between CV slots and syllabicity (syllabic segments are dominated by V). X theory only has X at the skeletal level: it does unquestionably require a nucleus node to reconstruct the notion "is a vowel" as opposed to "is a glide". Although CV theory did not posit Onset and Rhyme constituents, it does also admit a Nucleus node. The arguments for that model largely reduce to the desideratum that a theory should not provide two (or more) tools to do one thing, so as long as the structures needed in X-theory are independently necessary, X-theory may be judged to be simpler. The most compelling argument for X theory is the case that certain suffixes in Mokilese have the property of lengthening whatever segment they follow. Examples are [wɔl] 'man', [wɔlle] 'this man', [wɔllo] 'that man', [pwo] 'pole', [pwo:j] 'this pole', [pwo:w] 'that pole'. The demonstrative suffixes have the form /-Xe/, /-Xo/ where X links to the final segment of the preceding root. To the extent that the distinction C versus V encodes the property "is syllabic", we would not expect both a vowel and a consonant to be able to link to a single skeletal slot.

11.2.5 Non-segmental timing: the mora

Yet another prosodic unit has been posited to represent segment length, namely the **mora**, symbolized as μ. The mora was primarily a device for expressing the difference between a light versus heavy syllable in accent systems, where a heavy syllable has a long vowel, a diphthong, or a final consonant, but a light syllable has the structure C_0V (C_0VC sometimes, if consonants don't "count"). There is also an implicational rule that if a language has a concept of heavy versus light syllables, long vowels and diphthongs define heavy syllables, and coda consonants may, but need not (and sometimes, only certain consonants define heavy syllables). It is a puzzle how we can accommodate that variation. It is proposed in Hayes (1989) that long vowels and diphthongs have two underlying morae therefore their syllables are necessarily heavy, but coda consonants receive a mora by language-specific rule. Onset consonants are associated directly to the syllable node, and coda consonants associate automatically under the mora of the preceding consonant as in (41d) unless a language-specific rule creates a mora for the coda consonant to associate to, as in (41e). Because (41b,c) involve underlyingly moraic segments (vowels) and since vowels can lexically specify two morae, bimoraic syllables can emerge without any language-specific rule, in case the syllable has a long vowel or diphthong.

(41) a. light syllable b. heavy syllable c. heavy syllable d. light syllable e. heavy syllable



This analysis solves a puzzle about the notion of a heavy syllable in the theory of syllable structure, that "heaviness" exists always when the nucleus branches (long vowels or diphthongs) or when the rhyme branches (a language specific option). Instead, as indicated in (41), the difference lies in when a syllable is bimoraic, and not in a more complex definition of "heavy" that phonological rules might refer to. Since onsets never have a mora, branching at the level of the syllable (presence of an onset) never figures into the definition of syllable weight.

Each of these models of length – two moras, two X-positions, or two CV positions, has certain advantages and disadvantages and no consensus has been reached as to which model should be considered "standard".

11.3 The Foot

The third main change in the theory of representations arising from the nonlinear line of research in the mid 70's is a reconceptualization of "stress" as being not a feature of the segment, but as a prominence relation between syllables. Saying that a syllable (not a vowel) is stressed means that is it "prominent" compared to other syllables. The construct "foot" was a crucial tool for accounting for stress assignment, naturally other uses, separate from stress assignment, were sought to support the existence of the foot.

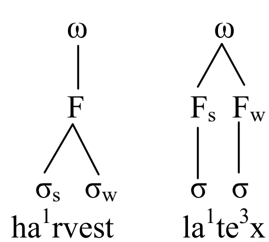
11.3.1 Foot groupings: stress

Stress posed a significant problem for the theory of phonological rules and representations. The original standard theory of generative phonology, SPE, Chomsky & Halle (1968), went to great lengths to account for the complexity of English stress. English has not just a two-way distinction between stressed and unstressed, it has numerous degrees of stress, notated with superscripted numbers after the vowel (no number means no stress). In simple words we find not only location differences like *permi^{1}t, pe^{1}rson*, we also have degree differences like *ha^{1}rvest, la^{1}te^{3}x; ba^{3}nda^{1}nna, bana^{1}na; Pa^{1}nama^3, Pa^{1}mela*. There are alternations in location and degree of stress related to derivational morphology as examplified by *co^{3}mpensa^{1}tion, co^{1}mpensa^{3}te, co^{3}nde^{4}nsa^{1}tion, conde^{1}nse, ari^{1}stocra^{3}t, a^{3}ri^{4}sto^{1}cracy*. In addition, there are stress changes at the phrasal level when words combine into different kinds of phrases, for instance *che^{1}mistry la^{2}borato^{4}ry, a^{2}bsolu^{4}te e^{4}quali^{1}ty*, or the expanding sequence of modifiers in *la^{1}w degree^2, la^{1}w degree^3 requi^{2}rement , la^{1}w degree^4 requi^{3}rement cha^{2}nges*.

To account for these very complex facts, the SPE analysis required otherwise unnecessary expansions to the theory of rules and representations: the feature [stress] has numeric coefficients rather than just plus and minus, and rules perform global mathematical operations of subtraction (so that combining *law^1* and *degree^1* results in *la^{1}w degree^2*, by assigning 1 stress to *law* and subtracting 1 from all other stresses). A significant mis-prediction of this theory is that some rule of phonology might care whether a particular vowel has 3 stress versus 4 stress, and nothing in principle would prevent a language from distinguishing lexical items like [see^1] versus [see^2], but this doesn't happen. For phonological purposes, rules only distinguish stressed from unstressed, numeric degrees are only necessary to get the pronunciations right. There may be better tools for getting correct pronounciation without a numeric stress continuum.

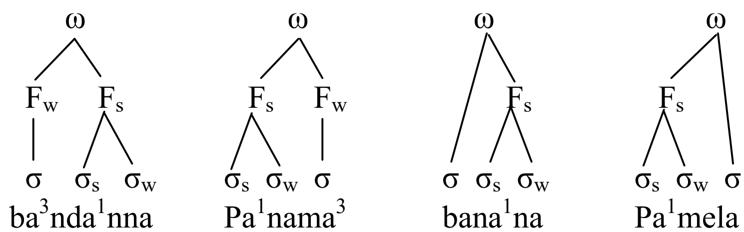
An alternative representation has been proposed, namely the **metrical tree** which groups syllables into higher-level constituents, called **feet** (usually containing only two syllables). When metrical structures branch, one branch is labelled *s* (strong: prominent, stressed) and the other is labeled *w* (weak, not prominent, unstressed). Not only are syllables grouped together into higher units, those feet are gathered together into a labeled word-structure. The numbers of SPE are then eliminated in favor of this relational and recursive understanding of binary prominence relations. For example, the difference between *ha^{1}rvest* and *la^{1}te^{3}x* is that *ha^{1}rvest* has one stress therefore one foot, which includes two syllables, but *la^{1}te^{3}x* has two stresses therefore two feet, and the first of those feet is the strongest within the word. In these metrical structures, F stands for "Foot" and ω stands for "Word".

(42)



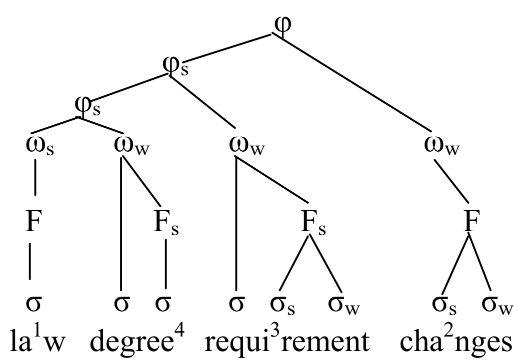
The differences between *ba^{3}nda^{1}nna* and *Pa^{1}nama^3* are treated analogously as having two feet, in both cases a single-syllable foot along with a two-syllable foot, differing only in where the single-syllable foot resides. The exact disposition of the unstressed initial or final syllables of *bana*^{1}*na, Pa*^{1}*mela* is not obvious, but it is proposed that these syllables are simply added directly to the word tree, and do not fall into any foot

(43)



Such labeled trees can also be applied to phrases to derive the relationship between depth of syntactic embedding and degree of stress subordination (φ stands for a recursive phrasal constituent **phonological phrase**, discussed in §11.4.

(44)



The transcriptional stress numbers can be mimicked by deeming that the syllable dominated only by *s* has the highest stress, and the relative stress of other stressed syllables is a function of the number of *s* or *w* nodes dominating in the tree. Numeric manipulation is then eliminated from grammatical theory, which is a desirable result.

English has a rather complex and arbitrary stress system, the application of metrical structure to other languages has led to numerous insights into stress rules, via a research program initiated in Hayes (1980). Stress systems of the world can be largely reduced to a small number of variables regarding the parsing of syllables into stress feet. The primary controlling properties are in (45).

(45) Iteration: Does the rule apply only once or does it apply to successive substrings?
Starting point: Does the (iterative) rule apply left-to-right or right-to left?
Labeling: Are units labeled sw or ws (specified both for groups of syllables within the foot and groups of feet within the word
Quantity-sensitivity: Are heavy syllables treated differently from light syllables in foot-construction?
Minimality: Can a foot contain only one syllable (or mora)?

Extrametricality: Is some segment or syllable at the margin of a word ignored in foot construction (marked with parentheses)?

Foot size: Can a foot have three syllables; are there unbounded feet?

Here are some exemplifications of the metrical analysis. Non-iterative foot construction means that there is only one foot therefore one stress per word, examples being Czech (first syllable), Polish (second to last syllable), Persian (final). These may involve differences in *sw* versus *ws* labeling in the case of final versus penultimate stress, or penult stress could be analysed as *ws* with extrametricality of the final syllable. Such indeterminacy of analysis could motivate re-thinking of the principles of metrical parsing. Quantity-sensitive single stress systems (such as Latin) are of significant interest and will be considered later.

In the realm of iterative quantity-insensitive stress systems, direction of footing and labeling are the main parameters relevant. In Maranungku (46), feet are constructed iteratively starting at the left, and are labeled *sw*. Stress is not sensitive to syllable quantity. Feet (at the end of the word) are allowed to contain only one syllable, and feet are gathered into a word-tree where the first foot is strongest, therefore is labeled *s*. In the following displays, the strong member of a constituent is indicated in bold, feet are contained in parentheses, syllables are in braces, and extrametrical items are in angled brackets.

(46)	$(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{ti}\}\{_{\sigma}\text{ralk}\})$	‘saliva’
	$(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{me}\}\{_{\sigma}\text{re}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{pet}\})$	‘beard’
	$(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{ja}\}\{_{\sigma}\text{ŋar}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{ma}\}\{_{\sigma}\text{ta}\})$	‘the Pleaiades’
	$(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{laŋ}\}\{_{\sigma}\text{ka}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{ra}\}\{_{\sigma}\text{te}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{i}\})$	‘prawn’
	$(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{we}\}\{_{\sigma}\text{le}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{pe}\}\{_{\sigma}\text{ne}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{man}\}\{_{\sigma}\text{ta}\})$	‘duck (sp)’

Mapuche (Araucanian) (47) presents a variant pattern where the first syllable of the word is ignored. The main stress is on the second syllable, and secondary stresses appear on every even-numbered syllable thereafter. This can be generated by marking the first syllable of the word as extrametrical (in angled brackets), then constructing bisyllabic *sw* feet and a word tree as in Maranungku. Alternatively, the first syllable could be assigned to the first foot (not extrametrical), and the labelling could be *ws*.

(47)	$<\{_{\sigma}\text{wu}\}>(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{le}\})$	‘tomorrow’
	$<\{_{\sigma}\text{t̪i}\}>(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{pan}\}\{_{\sigma}\text{to}\})$	‘year’
	$<\{_{\sigma}\text{e}\}>(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{lu}\}\{_{\sigma}\text{mu}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{ju}\})$	‘give us’
	$<\{_{\sigma}\text{e}\}>(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{lu}\}\{_{\sigma}\text{a}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{e}\}\{_{\sigma}\text{new}\})$	‘he will give me’
	$<\{_{\sigma}\text{ki}\}>(_{\mathbf{F}\,ˈ}\{_{\boldsymbol{\sigma}}\text{mu}\}\{_{\sigma}\text{ba}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{lu}\}\{_{\sigma}\text{wu}\})(_{\mathrm{F}\,ˌ}\{_{\boldsymbol{\sigma}}\text{laj}\})$	‘he pretended not to know’

The Weri pattern in (48) is the opposite of that of Maranungku. Feet are constructed iteratively from right to left, labeled *ws*. The word tree is labeled so that the last foot is strongest, therefore main stress is on the right and preceding stressed are weaker.

(48) ($_{\mathbf{F}}${$_\sigma$ŋin}{$_{\boldsymbol\sigma}$ˈtip}) ‘bee’
($_{\mathrm{F}}$ˌ{$_{\boldsymbol\sigma}$ku})($_{\mathbf{F}}${$_\sigma$li}{$_{\boldsymbol\sigma}$ˈpu}) ‘hair of arm’
($_{\mathrm{F}}${$_\sigma$u}{$_{\boldsymbol\sigma}$ˌlu})($_{\mathrm{F}}${$_\sigma$a}ˈ{$_{\boldsymbol\sigma}$mit}) ‘mist’
($_{\mathrm{F}}$ˌ{$_{\boldsymbol\sigma}$a})($_{\mathrm{F}}${$_\sigma$ku}{$_{\boldsymbol\sigma}$ˌne})($_{\mathbf{F}}${$_\sigma$te}{$_{\boldsymbol\sigma}$ˈpal}) ‘times’

In Warao, stress is computed starting at the end of the word like in Weri, but syllables are labeled *sw* rather than *ws*.

(49) {$_\sigma$ji}($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌwa}{$_\sigma$ra})($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈna}{$_\sigma$e}) ‘he finished it’
($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌja}{$_\sigma$pu})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌru}{$_\sigma$ki})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌta}{$_\sigma$ne})($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈha}{$_\sigma$se}) ‘verily to climb’
{$_\sigma$e}($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌna}{$_\sigma$ho})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌro}{$_\sigma$a})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌha}{$_\sigma$ku})($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈta}{$_\sigma$i}) ‘the one who caused him to eat’

Warao poses a minor problem that the initial syllable in words with an odd number of syllables such as [jiˌwaranae] is predicted to have a stress – *[ˌjiˌwaranae] – but they do not. Any odd-syllabled word has one word-initial syllable which should form its own foot and have its own stress (just as the final odd syllable does in Weri). This problem can be handled by requiring in Warao that every foot *must* have two syllables, thus foot construction is subject to a minimal size requirement. That requirement means that an initial syllable in an odd-syllabled word will be not in any foot, since there are not enough syllables to form a foot. Instead it just directly adjoins to the word tree, just like the initial syllable of English *bana*^{1}*na*.

Pintupi then exemplifies what Maranungku would look like if the language had the same two-syllable minimum requirement on foot construction. Word-final odd-numbered syllables in Maranungku receive stress because, even though they are the only syllable in the final foot, that is allowed in Maranungku. That is not allowed in Pintupi: odd-numbered final syllables do not have stress, because every foot must contain two syllables.

(50) ($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈpa}{$_\sigma$ɳa}) ‘earth’
($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈtju}{$_\sigma$ʈa}){$_\sigma$ja} ‘many’
($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈma}{$_\sigma$ɭa})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌwa}{$_\sigma$na}) ‘through from behind’
($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈpu}{$_\sigma$ɭiŋ})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌka}{$_\sigma$la}){$_\sigma$tju} ‘we (sat) on the hill’
($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈʈi}{$_\sigma$li})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌri}{$_\sigma$ŋu})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌlam}{$_\sigma$pa}){$_\sigma$tju} ‘the fire for our benefit flared up’
($_{\mathbf{F}}${$_{\boldsymbol\sigma}$ˈtja}{$_\sigma$mu})($_{\mathrm{F}}${ˌ$_{\boldsymbol\sigma}$lim}{$_\sigma$pa})($_{\mathrm{F}}${$_{\boldsymbol\sigma}$ˌtju}{$_\sigma$ŋka}) ‘our relation’

It is important to note that there are often multiple analyses of a particular stress pattern. Analytic ambiguity cannot always be eradicated, but perhaps some independent modification to the theory will limit the indeterminacy.

Another property exhibited by many stress systems is quantity-sensitivity, where stress is assigned based on the weight of a syllable. This property is exhibited in the stress system of Muskogee (Creek), where syllables are assigned to feet going from left to right. Foot construction is sensitive to syllable weight, so that a syllable with a long vowel or a coda consonant can form a foot, but a light syllable cannot be a foot on its own. Unlike previous examples, only the last footing in the word translates into phonetic stress (or tone). The result is that the surface stress is either on the last or the penultimate syllable, depending on the number of syllables preceding, and whether they are heavy or light. The metrical analysis is that syllables group into binary feet starting at the left, labeled *ws*, but never putting a heavy syllable under the label *w*. Since every foot in Muskogee must have two moras, leftover final light syllables are not in any foot, therefore they have no stress – which generates the penultimate stress pattern.

(51)	(ifá)	'dog'	(amí)fa	'my dog'
	(osá)na	'otter'	(pomo)(saná)	'my otter'
	(apa)(taká)	'pancake'	(ama)(patá)ka	'my pancake'
	(fó:)	'bee'	(nihá:)	'lard'
	(hok)(tí:)	'woman'	(hito)(tí:)	'snow'
	(ít^{ʃ})ki	'mother'	(kofót^{ʃ})ka	'mint'
	(ak)(t^{ʃ}óh)ka	'stork'	(ta:s)(kitá)	'to jump (sg. subj)'
	(ati)(lo:)(jitá)	'to gather (sg. obj)'	(naf)(kiti)(ka:)(jitá)	'to hit (pl. obj)'
	(tokoł)(hokí)ta	'to run (pl. subj)'		

Since there is only a single stress (or tone) in a word, the motivation for building multiple feet containing one or two syllables is indirect: it provides a theoretical rationalization of how penult versus final stress depends on a count of preceding syllables and their weight.

Another quantity-sensitive stress system is Palestinian Arabic, which has a stress system similar to Classical Latin, where one stress is assigned on one of the last three syllables of the word, as determined by syllable weight. Stress appears on the final syllable if that syllable is heavy, it is on the penult if the penult is heavy and the final syllable is light, and is on the antepenult otherwise. The typical definition of a heavy syllable is one with either a long vowel or a final consonant, but in Arabic, final syllables have a special definition for "heavy," which is that a single consonant at the end of the word does not make the syllable heavy, while two consonants do. This is accounted for metrically by a rule making a single word-final consonant extrametrical, hence /qara/ and /qarat/ are treated identically because for stress purposes, final *t* is invisible. But only a single final consonant can be ignored.

(52)	radˈjoo	'radio'	qaˈree<t>	'I read'
	kaˈtab<t>	'I wrote'	ˈqara	'he read'
	ˈqara<t>	'she read'	kaˈtabna	'we wrote'
	qaˈreethu<m>	'I read them'	ˈkatabu	'they wrote'
	ˈkataba<t>	'she wrote'	ma kataˈbat<ʃ>	'she didn't write'

In Arabic, "tʃ" is a cluster made up of t+ʃ, not a single-segment affricate t^{ʃ}

The foot is also useful in solving problems beyond stress assignment. One common use is to account for minimal word-size requirements in languages. In the Zezuru dialect of Shona, words must contain at least two syllables. We can reduce this to a requirement in the language for words to be at least a foot in size, and requiring that the foot have at least two syllables (there are no long vowels or closed syllables). This requirement can be violated in underlying forms where a verb root in the imperative may have only one syllable. When that happens, [i] is inserted at the beginning of the word to make the word long enough.

(53)	ku-($_F$ rima)	'to plow'	($_F$ rimá)	'plow!'
	ku-vere($_F$ keta)	'to read'	veré($_F$ kétá)	'read!'
	($_F$ ku-pá)	'to give'	($_F$ i-pá)	'give!'

Damascus Arabic also has a requirement that a word must contain at least a foot, and a foot must contain at least two moras (with the proviso that, as is typical of Arabic dialects, a single word-final coda consonant does not have a mora). In (54) are examples of verb imperatives for the three inflected subjects masculine (no suffix), feminine (*-i*) and plural (*-u*). The examples in (a) show the subject agreement suffixes for feminine and plural. In (b) we see the same suffixes in monosyllabic roots with a medial long vowel, and in (c) a rule of apocope deletes certain short vowels in an open syllable. The roots in (d) also have a long vowel in the first syllable, but the root ends in a vowel, either /a/ or /i/. The root vowel is deleted when the suffixes /-i, -u/ are added.

The verbs in (e) have the underlying stem shape /CCaC/ or /CCoC/, evident from the suffixed forms, but in the unsuffixed forms, the vowel is lengthened and we have [sʔa:l, kto:b, wsˤa:l] rather than *[sʔal, ktob, wsˤal]. In (f), we have roots with the shape CCV, where the root vowel or the suffix vowel is lengthened.

(54)		*masc.*	*fem.*	*pl.*	
	a.	tamm	tamm-i	tamm-u	'remain'
		ħəss	ħəss-i	ħəss-u	'feel'
	b.	ba:n	ba:n-i	ba:n-u	'build'
		fi:ʔ	fi:ʔ-i	fi:ʔ-u	'wake up'
	c.	sa:ʕed	sa:ʕd-i	sa:ʕd-u	'help'
		ʒarreb	ʒarrb-i	ʒarrb-u	'try'
	d.	ħa:ki	ħa:k-i	ħa:k-u	'talk to'
		tħa:ʃa	tħa:ʃ-i	tħa:ʃ-u	'avoid'
	e.	sʔa:l	sʔal-i	sʔal-u	'ask'
		kto:b	ktəb-i	ktəb-i	'write'
		wsˤa:l	wsˤal-i	wsˤal-u	'arrive'
	f.	ʔra:	ʔr-i:	ʔr-u:	'read'
		ʔwa:	ʔw-i:	ʔw-u:	'become strong'
		mʃi:	mʃ-i:	mʃ-u:	'walk'
		bʔi:	bʔ-i:	bʔ-u:	'stay'

Since a foot in Arabic must have at least two moras, forms like /ʔra, sʔal/ cannot form a foot, so such words are disallowed as such. The resolution of this problem is that the existing vowel of the word is lengthened. In principle, a vowel could be inserted as it is in Zezuru (*iʔra) or in the case of consonant-final roots the consonant could be geminated (*ktobb), but this does not happen. Although we can sensibly relate vowel lengthening to requirements of the foot, such requirements do not eliminate the need for a mechanism saying what does happen to generate a satisfactory output. We will touch on this problem again in the final chapter.

11.3.2 Foot groupings: Estonian

Estonian is famous for having three degrees of length for both consonants and vowels, referred to as Q1, Q2 and Q3. Ordinary long segments are Q2, Q3 (marked with ":" in the data) is an extra-long version of a Q2 segment, and is determined by the syllable being "extra-heavy", it is not an additional segmental length distinction. Some examples are in (55). In the Estonian examples, grammatical inflections are abbreviate "s." = singular, "p." = partitive, "g." = genitive, "i." = illative.

(55)	[vilu]	'cool'	[lina]	'flax'	Q1
	[viilu]	'slice g.s..'	[linna]	'town g.s.'	Q2
	[vii:lu]	'slice p.s.'	[lin:na]	'town i.s.'	Q3

This posed a problem since there are only two feature values: [–long] and [+long]. With an autosegmental representation of length, one could in principle represent any number of degrees of segment length by adding more skeletal slots or moras. But that tack would obscure a number of interesting phonological facts about these apparent degrees of length, which are set forth in the metrical analysis of Prince (1980).

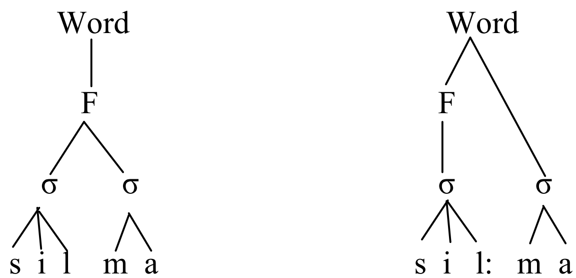
One such fact is that the different degrees of length do not freely co-mingle in the language. The possible syllables of Estonian are schematized in (56) with a set of two-syllable words having a final CV syllable preceded by one of the possible first syllables, in the three quantities. The examples also show that alternations in quantity plays a role in Estonian morphology.

(56)	Q1	Q2	Q3
	CV-CV	**CVV-CV**	**CVV:-CV**
	lina 'flax n.s.'	viilu 'slice g.s.'	vii:lu 'slice p.s.'
		CVVC-CV	**CVV:C:-CV**
		kaarti 'card g.s.'	kaa:r:ti 'card p.s.'
		CVC-CV	**CVC:-CV**
		silma 'eye g.s.'	sil:ma 'eye p.s.'

An important generalization about the language is that a syllable cannot mix Q2 and Q3, thus a Q3 vowel cannot be followed by a Q2 consonant and a Q2 vowel cannot be followed by a Q3 consonant. This immediately suggests that Q2 and Q3 are not different points on a duration continuum of segments, rather, the language simply has long and short vowels and consonants as many languages to, but in addition syllables come in two varieties, ordinary (which can be light or heavy) and “extra-heavy”, a special type of heavy syllable. A long vowel or consonant in an extra-heavy syllable is Q3, a long vowel or consonant in a regular syllable is Q2. This accounts for the lopsided distribution of Q2 and Q3 within the syllable, but now we have to understand what causes the difference between these two syllable types.

The distinction proposed by Prince is that a Q3 syllable is the only syllable in a foot, whereas Q2 is followed by another syllable within the foot.

(57)



Apart from providing a representation of the difference between these quantities and accounting for the non-mixing of Q2 and Q3, this theory also accounts for other facts. First, Q3 only appears in a stressed syllable – necessarily, the sole syllable of a foot *is* stressed. Second, every monosyllable in the language is in Q3, e.g. [kæt:t] ‘hand p.s.’, [sil:m] ‘eye n.s.’, [haa:v] ‘scar’, which follows from the fact that there is only one syllable in the foot (and every monosyllable has a diphthong, long vowel or long consonant, i.e. *[pil] cannot arise). Third, exceptional main stress on the final syllable in words like [aveˈnyy:] ‘avenue’ is always in Q3, because final stress means that there is no other syllable in that foot, thus it is a monosyllabic foot – Q3. Fourth, Estonian has alternating stress following the pattern [ˈσσˌσσˌσσ] or optionally [ˈσσσˌσσσ] e.g. [ˈhiliˌsematˌtele, ˈhiliseˌmattele] ‘later allative pl.’. As expected, there will not be two adjacent stressed syllables because Estonian feet normally contain two syllables, and optionally can contain three syllables. But adjacent stresses *are* possible if the first syllable is Q3, for example [ˈkau:ˌkele] ‘far away’, [ˈtøø:s:tˌtustesse] ‘industry ill. pl.’, [ˈjul:ˌkesse] ‘bold ill. sg’. This is as predicted because a Q3 syllable is the only syllable in the foot, and a new foot therefore can start right after that syllable.

While this gives us a means of representing the quantity distinction and explaining the connection between quantity and stress, by positing contrastive foot structure (similar to English *ˈlatest* vs. *ˈlaˌtex*), there remains the puzzle of how such contrasts arise. This brings us to the quantity alternations in [kaarti] ‘card g.s.’ and [kaa:r:ti] ‘card p.s.’. In Estonian, foot structure is contrastively assigned in part on the basis of grammatical category, it is not a fundamental lexical difference. The trio of nominative, genitive and partitive singular noun forms in (58) reveals some basic prosodic and segmental alternations of the language.

(58)	n.s.	g.s.	p.s	
	haa:v	haava	haa:va	'scar'
	lin:n	linna	lin:na	'town'
	sil:m	silma	sil:ma	'eye'
	kap:p	kappi	kap:pi	'closet'
	pata	paja	pata	'kettle'
	təpi	təve	təpe	'disease'

The underlying roots of these words are /haava, linna, silma, kappi, pata, təpe/. The case-marking difference was originally signalled by segmental affixation, /Ø/ = n.s., /-n/ = g.s., /-ta/ = p.s, but these suffixes have substantially (though not completely) eroded. The underlying final vowel deletes at the end of the word provided that the word has more than two moras (which is why there is no deletion in [pata, təpi]). In nominatives like /haava/, this results in a monosyllabic foot, by definition Q3. Vowel deletion does not happen in the genitive singular which had a suffix -*n* protecting the final vowel from deleting. That consonant still plays a role in the phonology of Estonian, because just like in Finnish it also triggers a rule of "gradation" where /p,t,k/ weaken – in the above examples we see /d/→ [j], /p/ → [v]. In addition, there is a rule raising and fronting word-final /e/ to [pi], observed in n.s. [təpi, təve, təpe] and similar words. Genitive singular [təve] does not undergo e-raising because *e* is not final, it is followed by the g.s. suffix -C (which generally has no phonetic realization). The reason (at least historically) for weakening of /t,p/ in the genitive singular is a rule weakening a consonant on the onset of an unstressed heavy syllable, where /təpeC/ → [təve(C)].

Weakening in the genitive singular can result in a Q3 syllable, even though generally the genitive singular does not intrinsically cause a quantity change (whereas the partitive singular regularly does). That happens just in case a consonant is weakened to Ø, as is the case with /k/.

(59)	n.s.	g.s.	p.s	
	vika	vea:	vika	'error'
	reki	ree:	reke	'sled'
	jəki	jəe:	jəke	'river'

The partitive singular suffix has two main effects, one being that like the genitive singular the root-final vowel is neither deleted nor raised ([haa:va, kap:pi, təpe]), which follows from assuming a suffix CV. Other facts suggest that the "ghost" consonant is coronal, given the appearance of *t* in the partitive singular in the genitive ~ partitive pairs [maa:, maa:t] 'land', [tee:, tee:t] 'road', [tule, tul:t] 'fire', [lume, lun:t] 'snow', [kuse, kus:t] 'urine'. The partitive suffix will be symbolized here as /CV/, leaving the question of its exact content to further investigation. The partitive is the only CV inflectional suffix.

The other special property of the partitive is that if the root-final syllable can be excluded from the preceding foot and still have a legal foot (two moras), the final syllable *is* omitted from that foot. In /pata+CV/ the final syllable *ta* cannot be omitted from the first foot since every foot must have two moras, and *$(_F\{_\sigma pa\})\{_\sigma ta\}$+CV does not, therefore this word must be parsed as $(_F\{_\sigma pa\}\{_\sigma ta\})$+CV. Since $(_F\{_\sigma haa\})\{_\sigma va\}$ and $(_F\{_\sigma sil\})\{_\sigma ma\}$ both have an initial bimoraic syllable, the special foot-construction rule for partitives can force a monosyllabic foot, therefore those syllables will be Q3.

11.4 Even higher units

It should not be surprising to learn that with the skeleton, syllable and foot motivated as hierarchical representations above the segment, other constituents have been posited, collectively known as the **Prosodic Hierarchy**. These include:

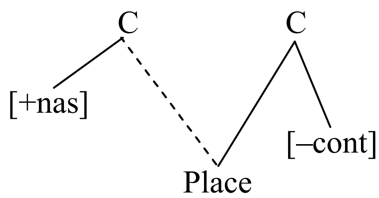
(60) Phonological Word
Clitic Group
Minor Phrase
Phonological Phrase
Intonational Phrase
Utterance

The rationale for these entities is fairly straightforward: some rule in a language needs to refer to either a relationship between two lower units which stand in the hypothesized constituent (but not in separate higher constituents), or a rule refers to the fact of being at the edge of one of these constituents. Numerous rules apply to a segment at the end of the word, for example vowels delete at the end of the word in Lardil. The various stress rules presented above are computed based on the position of a syllable within the word (and not some higher unit like the sentence). Vowel harmony rules typically apply to a sequence of vowels within the word. In traditional rule theory, an object "#" is placed in the representation to mark "end of word", and rules refer to word-ends by mentioning "#" in the rule statement. If "phonological word" is a representational constituent, symbolized as ω, the same effect is achieved by referring to the label "$_\omega]$" in the rule. In the case of a rule that applies when α, β are "in the same word", the rule must be notated as "having the word as its domain", meaning that elements matched by the rule must be "in the same word". In traditional rule theory, all rules are automatically subject to this limitation unless "#" is explicitly included in the rule statement. For example, in Sanskrit /n/ assimilates to a following stop both within the word and across words. The traditional rule statement is (61), which says that word boundary is admitted but not required.

(61) $$[+\text{nas}] \rightarrow \begin{bmatrix} \alpha\text{ant} \\ \beta\text{cor} \\ \gamma\text{back} \\ \delta\text{hi} \end{bmatrix} / \_\_ (\#) \begin{bmatrix} -\text{cont} \\ \alpha\text{ant} \\ \beta\text{cor} \\ \gamma\text{back} \\ \delta\text{hi} \end{bmatrix}$$

Naturally, the autosegmental account does not express this rule as a collection of 4 feature changes, it frames the process as a single change in a higher node "place". In this case, where the nasal and the stop can be within a word or in separate words, the rule simply does not mention a particular domain, therefore any nasal-plus-stop sequence undergoes the rule.

(62)



If the rule were to *only* apply within the word and not between words, the condition "Domain: ω" would be added to the rule.

Sometimes a rule applies only if the word is absolute-final in the utterance, and is not followed by anything else – a rule can then refer to the constituent label $_{U}]$, where in traditional rule theory this is notated as "##". An example of such a rule is the one assigning H tone to a toneless word at the end of the utterance in Matumbi discussed in §5.5.2.

Between ω and U, there are a number of other posited units. The most widely-required is φ, the "Phonological Phrase", which attempts to encode syntactic relations as they figure into phonology. An example of such a relation is Phrasal Shortening in Matumbi (discussed in the online sketch for Chapter 7). An underlying long vowel is shortened in a word when that word is the head of a phrase which the following word is a modifier of – when the two words are in the syntactic phrase headed by the first word. Various tone alternations follow from this change in vowel length. We see NP, VP and AP structures in (63a-c).

(63)	a.	mboópo	'machete'
		mbopó jaáŋgu	'my machete'
		mbopó ŋgʊ́lʊ́	'big machete'
	b.	u-paánde	'you should plant'
		[u-pandé maloómbe]	'you should plant corn'
		[u-pandé paapasá]	'you should plant the day after tomorrow'
		[u-pandé lí]	'you should not plant'
	c.	mikaáte	'loaves'
		[mikaté mikeéle]	'red loaves'
		[mikaté [mikelé mikeéle]]	'red loaves'

In (64) are examples of word pairs where the first word does not undergo Phrasal Shortening, because the first word is not the head of the syntactic phrase dominating the two words. In (64a), the head of the VP is the verb, not the specifier *keénda*; in (b) *Mandoondo* is not the head of the VP which contains *Mamboondo* and *eéla* (the verb /naampéei/, on the other hand, is the head of the VP and therefore it undergoes Phrasal Shortening). The analogous situation holds for /naakibwéeni + kikóloombe/ where Phrasal Shortening applies, and /kikóloombe + líilí/ where the rule does not apply. Finally in (d) the word pair /mboopó + ndaasó/ undergoes Phrasal Shortening because ndaasó is a modifier within mboopó's phrase, but not in /ndaasó + nkʊ́lʊ́/ since 'big' is not a modifer of 'long', they are in separate syntactic phrases.

(64) a. $[_{vp}[_{spec}$ keénda] $[_{v}$akákaláanga] 'if he had fried'
if he had fried

b. $[_{vp}$naampéi $[_{np}$Mamboondo] $[_{np}$eéla]] 'I gave Mambondo money'
I gave him Mamboondo money

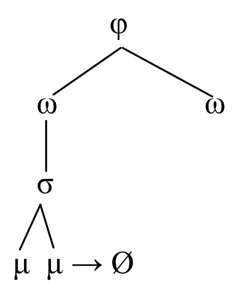
c. $[_{vp}$naakibwéni $[_{np}$kikóloombe] [líilí]] 'I didn't see a shell'
I saw it shell NEG

d. $[_{np}$mpopó $[_{ap}$ndaáso] $[_{ap}$ngʊlʊ́]] 'long big machete'
machete long big

Verb prefixes do not shorten their long vowels, because those vowels are underlyingly the result of a later rule mergine two vowels into a single long segment.

The theory of the Prosodic Hierarchy holds that it is the grouping of these words into a phonological unit φ which governs application of Phrasal Shortening. The rule would be stated as (65).

(65) *Phrasal Shortening*



The required groupings are those in (66).

(66) ($_{\varphi}$mbopó jaáŋgu) 'my machete'
($_{\varphi}$upandé maloómbe) 'you should plant corn'
($_{\varphi}$mikaté ($_{\varphi}$mikelé mikeéle)) 'red loaves'
($^{\varphi}$naabwéni ($_{\varphi}$mikaté ($_{\varphi}$mikelé mikeéle))) 'I saw red loaves'
($_{\varphi}$keénda) ($_{\varphi}$akákaláanga) 'if he had fried'
($_{\varphi}$naampéi Mamboondo) ($_{\varphi}$eéla) 'I gave Mambondo money'
($_{\varphi}$naakibwéni kikóloombe) ($_{\varphi}$líilí) 'I didn't see a shell'
($_{\varphi}$mbopó ndaáso) ($_{\varphi}$ngʊlʊ́) 'long big machete'

The main puzzle regarding how words group into φ is what phonological and syntactic factors are relevant to constructing φ and perhaps other word-groups. For the most part, the dominant structural relationship relevant to such constituency is the head-modifier relation, at least for solidly phonological groupings like those in Matumbi. More fluid and information-structure dependent parsing holds sway as we make the transition from abstract phonology to physical realization especially in the realm of intonation. Phenomena such as Matumbi Phrasal Shortening are not well attested, therefore it is difficult to form cross-linguistically robust phonological generalizations about how many higher-level units there are, and what principles govern their construction.

Summary

Prosodic constituency is particularly controversial in phonology because unlike segmental properties which have relatively well-understood audible definitions and means of verification, prosodic properties are intrinsically relational and lacking in fixed phonetic content. Because we cannot directly point to an absolute physical definition of a syllable, timing unit, foot, word or phonological phrase, we expect there to be controversy and theoretical debate over the precise nature of these objects. Solid results in the phonological analysis of these constituents will, for this reason, have a significant impact on our understanding of the human capacity to reason beyond immediate physical impressions.

Exercises

1. Turkish. This exercise repeats some earlier data from ch. 7, supplemented with additional examples. The point of this exercise is to recast the prior analysis (if you were did that problem) plus the new data in terms of the prosodic concepts presented in this chapter. Those data are augmented with additional stem types that pose a significant challenge to pre-autosegmental representations.

It is important to bear in mind that long vowels like [a:] are phonetically distinct from and transcribed differently from identical vowel clusters like [aa].

Nom	*Poss*	*Dat*	*Abl*	*Nom pl*	
oda	odasɨ	odaja	odadan	odalar	‘room’
dere	deresi	dereje	dereden	dereler	‘river’
balo	balosu	baloja	balodan	balolar	‘ball’
arɨ	arɨsɨ	arɨja	arɨdan	arɨlar	‘bee’
la:	la:sɨ	la:ja	la:dan	la:lar	‘la (note)’
bina:	bina:sɨ	bina:ja	bina:dan	bina:lar	‘building’
imla:	imla:sɨ	imla:ja	imla:dan	imla:lar	‘spelling’
kep	kepi	kepe	kepten	kepler	‘cap’
at	atɨ	ata	attan	atlar	‘horse’
ahmet	ahmedi	ahmede	ahmetten	ahmetler	‘Ahmed’
kurt	kurdu	kurda	kurttan	kurtlar	‘worm’
tyrk	tyrky	tyrke	tyrkten	tyrkler	‘Turk’
gent^{ʃ}	gent^{ʃ}i	gent^{ʃ}e	gent^{ʃ}ten	gent^{ʃ}ler	‘young’

halk	halkɨ	halka	halktan	halklar	‘folk’
yst	ysty	yste	ystten	ystler	‘upper plane’
sarp	sarpɨ	sarpa	sarptan	sarplar	‘steep’
harp	harbɨ	harba	harptan	harplar	‘war’
alt	altɨ	alta	alttan	altlar	‘bottom’
renk	rengi	renge	renkten	renkler	‘color’
his	hissi	hisse	histen	hisler	‘feeling’
hyr	hyrry	hyrre	hyrden	hyrler	‘free’
mahal	mahallɨ	mahalla	mahaldan	mahallar	‘place’
hak	hakkɨ	hakka	haktan	haklar	‘right’
zam	zammɨ	zamma	zamdan	zamlar	‘inflation’
af	affɨ	affa	aftan	aflar	‘excuse’
arap	arabɨ	araba	araptan	araplar	‘Arab’
kojun	kojunu	kojuna	kojundan	kojunlar	‘sheep’
kitap	kitabɨ	kitaba	kitaptan	kitaplar	‘book’
domuz	domuzu	domuza	domuzdan	domuzlar	‘pig’
davul	davulu	davula	davuldan	davullar	‘drum’
bajɨr	bajɨrɨ	bajɨra	bajɨrdan	bajɨrlar	‘slope’
somun	somunu	somuna	somundan	somunlar	‘loaf’
fikir	fikri	fikre	fikirden	fikirler	‘idea’
isim	ismi	isme	isimden	isimler	‘name’
bojun	bojnu	bojna	bojundan	bojunlar	‘neck’
tʃevir	tʃevri	tʃevre	tʃevirden	tʃevirler	‘injustice’
devir	devri	devre	devirden	devirler	‘transfer’
kojun	kojnu	kojna	kojundan	kojunlar	‘bosom’
karɨn	karnɨ	karna	karɨndan	karɨnlar	‘thorax’
burun	burnu	burna	burundan	burunlar	‘nose’
akɨl	aklɨ	akla	akɨldan	akɨllar	‘intelligence’
ʃehir	ʃehri	ʃehre	ʃehirden	ʃehirler	‘city’
namaz	namazɨ	namaza	namazdan	namazlar	‘worship’
zaman	zama:nɨ	zama:na	zamandan	zamanlar	‘time’
harap	hara:bɨ	hara:ba	haraptan	haraplar	‘ruined’
i:kaz	i:ka:zɨ	i:ka:za	i:kazdan	i:kazlar	‘warning’
hajat	haja:tɨ	haja:ta	hajattan	hajatlar	‘life’
ispat	ispa:tɨ	ispa:ta	ispattan	ispatlar	‘proof’

inek	inei	inee	inekten	inekler	‘cow’
mantɨk	mantɨɨ	mantɨa	mantɨktan	mantɨklar	‘logic’
ajak	ajaɨ	ajaa	ajaktan	ajaklar	‘foot’
t^{ʃ}abuk	t^{ʃ}abuu	t^{ʃ}abua	t^{ʃ}abuktan	t^{ʃ}abuklar	‘quick’
dakik	dakii	dakie	dakikten	dakikler	‘punctual’
merak	mera:kɨ	mera:ka	meraktan	meraklar	‘curiosity’
tebrik	tebri:ki	tebri:ke	tebrikten	tebrikler	‘greetings’
hukuk	huku:ku	huku:ka	hukuktan	hukuklar	‘law’
t^{ʃ}ɨɨr	t^{ʃ}ɨ:rɨ	t^{ʃ}ɨ:ra	t^{ʃ}ɨɨrdan	t^{ʃ}ɨɨrlar	‘era’
baɨr	ba:rɨ	ba:ra	baɨrdan	baɨrlar	‘rump’
oul	o:lu	o:la	ouldan	oullar	‘son’
uur	u:ru	u:ra	uurdan	uurlar	‘good luck’
bøyr	bø:ry	bø:re	bøyrden	bøyrler	‘side’
gøys	gø:sy	gø:se	gøysten	gøysler	‘breast’
aɨz	a:zɨ	a:za	aɨstan	aɨzlar	‘mouth’
da:	daɨ	daa	da:dan	da:lar	‘mountain’
mevzu:	mevzuu	mevzua	mevzu:dan	mevzu:lar	‘topic’
tɨ:	tɨɨ	tɨa	tɨ:dan	tɨ:lar	‘crochet needle’

2. North Saami

This exercise presents alternations in suffix vowels and prosodic changes to consonants triggered by suffixes (“gradation”). The simple vowels in this dialect are [i u e o a ɑ]. The front vowel /a/ is always long or part of a diphthong (ea, oa), and /ɑ/ is short. This distribution is relevant to one of the vowel alternations. It is accurate (as is evident in the examples) that apart from [aa] vs. [ɑ], there is not a rampant vowel length contrast. Long [ii, uu, ee, oo] usually derive from /ie, uo, ea, oo/, but this is not a problem to be solved here. Vowel+glide+C sequences in the transcriptions, for example [mujhtɑlusɑjt] ‘stories (acc.pl)’, might instead be transcribed as [mŭihtɑlusɑit]. There is no contrast between [uiC, ɑiC] and [ujC, ɑjC], and you are encouraged to briefly consider whether anything in the data favors a two-vowel analysis versus a vowel plus glide analysis.

There is a clear contrast between geminate voiced and voiceless stops (*loddi* ‘bird n.s.’, *lotti* ‘bird a.s.’). Singleton intervocalic stops are transcribed as [b ð d^z d^ʒ g], the alveolar non-affricate being a clear voiced fricative, the others being phonetically ambiguous as variable weakly voiced stops. In some dialects they are more-clearly voiceless unaspirated stops [p t^s t^ʃ k]. Two suffixes in the data below, the nominative singular and essive markers, cause a prosodic “strengthening” of the previous consonant. Most relevant for this problem set is determining what one phonological fact causes the change, and the exact segmental differences (e.g. ð vs. b) can be made to follow fairly simply from that fact.

Cases are abbreviated as n.s. = nominative singular, a.s. = accusative singular, l.s. = locative singular, n.p., a.p. = nominative / accusative plural, ess = essive. The diminutive derivational form is marked as ‘dim.’. The following examples of disyllabic vowel-final noun stems lay out some of the gradation alternation, and 6 of the case affixes (or lack).

staallu	nierrɑ	gaaffe	dievvaa	baadd^{z}i	n.s.
staalu	nierɑ	gaafe	dievaa	baatt^{s}i	a.s.
staalus	nierɑs	gaafes	dievaas	baatt^{s}is	l.s.
staalut	nierɑt	gaafet	dievaat	baatt^{s}it	n.p.
staalujt	nierɑjt	gaafit	dievaajt	baatt^{s}it	a.p.
staallun	nierrɑn	gaaffen	dievvaan	baadd^{z}in	ess.
‘troll’	‘cheekbone’	‘coffee’	‘mound’	‘monument’	

The following n.s. vs. a.s. disyllabic vowel-final noun pairs expand the range of specific consonant alternations. The corresponding essives are exactly the n.s. plus final -*n*, the n.p. and l.s. are exactly the a.s. form plus respectively -*t* and -*s*, and the above a.p. examples suffice to allow you to fill in the a.p. form of the following nouns.

n.s.	a.s	
fuoððu	fuoðu	‘wild animal’
maannaa	maana	‘child’
t^{ʃ}iehkɑ	t^{ʃ}iegɑ	‘corner’
neahpi	neabi	‘nephew/niece of man’
geaht^{ʃ}i	gead^{ʒ}i	‘end’
goahti	goaði	‘big tent’
deaddu	deattu	‘weight’
reabbaa	reappaa	‘crab’
aadd^{ʒ}ɑ	aatt^{ʃ}ɑ	‘grandfather’
loddi	lotti	‘bird’
muggɑ	mukkɑ	‘jug’

Consonant-final stems behave differently. The following examples give more information about suffix form and some important points about consonant gradation. In understanding these data, it is useful to know that words end with at most a single consonant, which must be one of [s ʃ t n r l].

aallɑt	jaandor	rɑhpɑn	ustit	eajkkaat	n.s.
aallɑhɑ	jaandorɑ	rɑhpɑmɑ	ustibɑ	eajkkaaðɑ	a.s.
aallɑhis	jaandoris	rɑhpɑmis	ustibis	eajkkaaðis	l.s.
aallɑhin	jaandorin	rɑhpɑmin	ustibin	eajkkaaðin	ess
aallɑhɑt	jaandorɑt	rɑhpɑmɑt	ustibɑt	eajkkaaðɑt	n.p.
aallɑhit	jaandorit	rɑhpɑmit	ustibit	eajkkaaðit	a.p.
‘snow sparrow’	‘day’	‘opening ceremony’	‘friend’	‘owner’	

mujhtɑlus	bunjostɑt	saaʔmelɑʃ	t^{s}oohkolɑt	n.s.
mujhtɑlusaa	bunjostɑgaa	saaʔmelɑtt^{ʃ}ɑ	t^{s}oohkolɑgaa	a.s.
mujhtɑlusaas	bunjostɑgaas	saaʔmelɑtt^{ʃ}ɑs	t^{s}oohkolɑgaas	l.s.
mujhtɑlussɑn	bunjostɑhkɑn	saaʔmelɑdd^{ʒ}ɑn	t^{s}oohkolɑhkɑn	ess
mujhtɑlusaat	bunjostɑgaat	saaʔmelɑtt^{ʃ}ɑt	t^{s}oohkolɑgaat	n.p.
mujhtɑlusaajt	bunjostɑgaajt	saaʔmelɑtt^{ʃ}ɑjt	t^{s}oohkolɑgaajt	a.p.
'story'	'dip'	'Saami person'	'reason'	

baaʔneluʃ	oajvaamuʃ	borrɑmuʃ	bɑhaanaalɑt	n.s.
baaʔneluʃaa	oajvaamutt^{ʃ}ɑ	borrɑmuʃaa	bɑhaanaalɑgɑ	a.s.
baaʔneluʃaas	oajvaamutt^{ʃ}ɑs	borrɑmuʃɑs	bɑhaanaalɑgis	l.s.
baaʔneluʃʃɑn	oajvaamudd^{ʒ}ɑn	borrɑmuʃʃɑn	bɑhaanaalɑgin	ess
baaʔneluʃaat	oajvaamutt^{ʃ}ɑt	borrɑmuʃaat	bɑhaanaalɑgɑt	n.p.
baaʔneluʃaajt	oajvaamutt^{ʃ}ɑjt	borrɑmuʃaajt	bɑhaanaalɑgit	a.p.
'animal jaw'	'leader'	'meal'	'badly behaved reindeer'	

Diminutives are formed with a suffix, seen below in vowel-final stems. The lowering of final /u/ to [o] is due to a rule which is outside the focus of this problem, and can be handled by saying "before certain suffixes".

nierrɑ	gaaffe	dievvaa	staallu	N n.s.
nierɑʃ	gaafeʃ	dievaaʃ	staaloʃ	N+dim n.s.
nierɑd^{ʒ}ɑ	gaafed^{ʒ}ɑ	dievaad^{ʒ}ɑ	staalod^{ʒ}ɑ	N+dim a.s.
nierɑd^{ʒ}is	gaafed^{ʒ}is	dievaad^{ʒ}is	staalod^{ʒ}is	N+dim l.s.
nierɑd^{ʒ}in	gaafed^{ʒ}in	dievaad^{ʒ}in	staalod^{ʒ}in	N+dim ess
nierɑd^{ʒ}ɑt	gaafed^{ʒ}ɑt	dievaad^{ʒ}ɑt	staalod^{ʒ}ɑt	N+dim n.p.
nierɑd^{ʒ}it	gaafed^{ʒ}it	dievaad^{ʒ}it	staalod^{ʒ}it	N+dim a.p.
'cheekbone'	'coffee'	'mound'	'troll'	

Diminutives of C-final stems follow. A crucial question to consider is when and why the consonant of the diminutive suffix strengthens, given preceding facts about consonant-strengthening.

aallɑt	jaandor	mujhtɑlus	coohkolɑt	N n.s.
aallɑhɑʃ	jaandorɑʃ	mujhtɑlusaaʃ	coohkolɑgaaʃ	N+dim n.s.
aallɑhɑtt^{ʃ}ɑ	jaandorɑtt^{ʃ}ɑ	mujhtɑlusaad^{ʒ}ɑ	coohkolɑgaad^{ʒ}ɑ	N+dim a.s.
aallɑhɑtt^{ʃ}ɑs	jaandorɑtt^{ʃ}ɑs	mujhtɑlusaad^{ʒ}is	coohkolɑgaad^{ʒ}is	N+dim l.s.
aallɑhɑdd^{ʒ}ɑn	jaandorɑdd^{ʒ}ɑ	mujhtɑlusaad^{ʒ}in	coohkolɑgaad^{ʒ}in	N+dim ess
aallɑhɑtt^{ʃ}ɑt	jaandorɑtt^{ʃ}ɑt	mujhtɑlusaad^{ʒ}ɑt	coohkolɑgaad^{ʒ}ɑt	N+dim n.p.
aallɑhɑtt^{ʃ}ɑjt	jaandorɑtt^{ʃ}ɑjt	mujhtɑlusaad^{ʒ}it	coohkolɑgaad^{ʒ}it	N+dim a.p.
'sparrow'	'day'	'story'	'reason'	

These further examples of basic and diminutive nouns in the l.s. and ess. should reinforce your understanding of these alternations.

gahpir	gahpiris	gahpirin	gahpirattʃas	gahpiraddʒan	‘cap’
bajaan	bajaanis	bajaanin	bajaanattʃas	bajaanaddʒan	‘thunder’
beejot	beejohis	beejohin	beejohattʃas	beejohaddʒan	‘white reindeer’
logaahat	logaahagaas	logaahahkan	logaahagaadʒis	logaahagaadʒin	‘high school’
gonaagas	gonaagasaas	gonaagassan	gonaagasaadʒis	gonaagasaadʒin	‘king’
sarvaaskat	sarvaaskagaas	sarvaaskahkan	sarvaaskagaadʒis	sarvaaskagaadʒin	‘hide of buck’
N n.s	N l.s.	N ess	N+dim. l.s.	N+dim ess	

The final set of examples presents one last suffix, the 2nd person plural possessor, which follows the case marker. The first set of examples gives further evidence regarding the correct underlying forms of case suffixes. The surface forms of the possessive pronoun all reduce to a single underlying form which change because of certain rules, most of which are already in your analysis. For example the rule pertaining to the alternating consonant of the suffix should already be accounted for in the analysis of the preceding data (the change u→o in ‘troll’ can be disregarded).

staallo-ðet	nierra-ðet	gaaffe-ðet	dievvaa-ðet	baadd^{z}aa-ðet	ns
staalo-ðet	niera-ðet	gaafe-ðet	dievaa-ðet	baatt^{s}aa-ðet	as
staalust-eattet	nierast-eattet	gaafest-eattet	dievaast-eattet	baatt^{s}ist-eattet	ls
staalujd-eattet	nierajd-eattet	gaafid-eattet	dievaajd-eattet	baatt^{s}id-eattet	ap
staallun-eattet	nierran-eattet	gaaffen-eattet	dievvaan-eattet	baadd^{z}in-eattet	e
‘troll’	‘cheekbone’	‘coffee’	‘mound’	‘monument’	

In the following examples, “*” means that the particular expected form is not allowed.

jaandor-eattet	jaandoradd^{ʒ}-aðet	ns
jaandor-eattet	jaandoratt^{ʃ}-aðet	as
jaandorist-aaðet	jaandoratt^{ʃ}ast-eattet	ls
jaandoritt-aðet	jaandoratt^{ʃ}ajd-eattet	ap
*	jaandoradd^{ʒ}an-eattet	ess
‘day’	‘day-DIM’	

mujhtaluss-aðet	mujhtalusaad^{ʒ}-eattet	ns
mujhtalus-aaðet	mujhtalusaad^{ʒ}-eattet	as
mujhtalusast-eattet	mujhtalusaad^{ʒ}ist-aaðet	ls
mujhtalusajd-eattet	mujhtalusaad^{ʒ}itt-aðet	ap
mujhtalussan-eattet	*	ess
‘story’	‘story-DIM’	