

Phonology Ex Nihilo aka Radical Substance-Free Phonology and why I might recant

Wednesday?  Phonology Get-together

by

David Odden

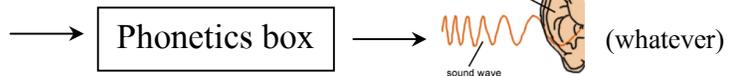
(Some newly discovered passages from *The Theaetetus*, probably penned by Aristotle)

Question 1: How substance-free can one (you, me, he, she) be in the theory of grammar?

Answer 1: My good Theaetetus, completely. No principle of grammar cares whether the feature that spreads or deletes is [nasal], [round] or [voice]. The features may be A, B, C, D, E, F.

Question 2: But don't you have to be able to pronounce phonological outputs?

Answer 2: Yes, but that's handled by the phonetic interpretation component. Phonology doesn't cover what's on the other side of the door.



Question 3: But without substance, how could a child learn that [β] is A, B, E and not C, F, G?

Answer 3: Why do you think [β] *isn't* C, F, G?

Question 4: Well, because [β] is a natural class with [γ, ð] which are all A, and with [m p] which are B.

Answer 4: No, I would say that [γ, ð] are G and [m, p] are F.

Question 5: But, but, but... Socrates, you're just playing with labels.

Answer 5: Moi? *I'm* not the one reifying the labels. Here, let me show you how you could arrive at feature classification inductively, importing **no** assumptions about phonetics into grammar.

Thus was born a new, minimally-nativistic linguistic theory that works both for human phonology and, with a tiny bit of jiggering, for the language of the famous Martian linguist Foley Ilähsihram.



Note the lack of ears or tongue. It turns out Illha¹ shares only 1% of the human genome but 98% of the human phonological faculty. That's why Illha invented MFP (Martian-Friendly Phonology). Martian cognition is primarily based on abstracting categories grounded in observing groups of objects that must, for a specific reason, be distinguished from another group of objects. I.e. for them, "because I can" is not enough of a reason. Martians can also hold 100^{googol^2} bits of information in their brains (a.k.a. veldig mykje), so "to economize" is not a reason.

¹ This is the appropriate 6th person masco-ptääh? quadratic pronoun for the referent: Martian has 16 genders and 7 numbers, as well as 8 persons.

The 8-Fold Assumption path

Assumption 1. The idea of “feature(¬node)”, “hierarchy”, and the theory of computation (the theory of grammar — what rules are, do, and how they can be arranged) are *all* that is in UG.

Corollary: Specific features are not in UG. The maximum number of features in a language is not in UG. Whether features are binary or monovalent *is* in UG (and since there is no universal content, either *all* features are binary, or *none* are).

Assumption 2. Feature assignment is arrived at inductively by reference to how sounds in the language behave phonologically — and not by authoritarian reference to physical properties of the sounds.

Assumption 3. Children acquire the segments of the language, learn underlying and surface forms of morphemes, identify alternations in the language and hold them in the form of inchoate (or final forms of) rules, and learn the contents of their language’s phonetics box which tells them to pronounce segment #41 as [u] (versus what some distant linguistic cousins do with segment #27 namely utter [u]).

Assumption 4. Features are induced on the basis of the need to “identify”, specifically to “group” and “distinguish” segments, both lexically and with respect to rules.

Corollary: Rules are written with features, but features have to be induced partially on the basis of what segments are in what classes, w.r.t. rules, so the rule-under-construction context “__ {t,s,n,d,l}”, which states what the child is aware of factually, maps to the rule object “__[F₃]”. Furthermore, if two sounds α, β are grammatically distinct sounds of the language, even in lieu of natural class behavior and rule-related considerations, some feature assignment must make the sounds formally different.

Assumption 5. Segments are composed of features (i.e. features are not just set-membership labels).

Corollary: Partial matrices (floating elements) are also well-formed: to take a typical example, a floating tone has some tone features, but no (segmental) laryngeal, vowel place or even skeletal features, as opposed to a H-toned vowel which does.

Assumption 6. Rules operate in terms of the “is a” relationship (i.e. can mention A but not the lack of A), no complements. This plus the theory of privative features gives privative behavior.

Assumption 7: Structure preservation is a real requirement on phonological rules: “structure” = “registered segment” and rule outputs must produce registered segments.

Corollary: rule outputs may have to be fiddles with to preserve structure.

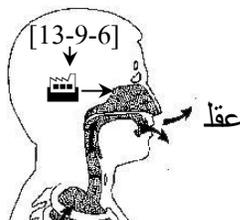
Assumption 8: Rules may apply vacuously (e.g. when /n/ → {m;n;ŋ}/ __ {p;t;k}, either /t/ is not in the conditioning class A so change is *precluded* before /t/; or, /t/ **is** in the class A and /n,t/ have the same value for the changing feature (vacuous application)).

A model of non-grammar (short-course in psycholinguistics)

Symbolization: acquire a representation of the acoustic waveform (left) as a mental symbol (right)

i.e. learn to ignore unimportant differences, pay attention to important ones. Find the units. Grammar-construction operates on [1] and not on د. Prerequisite: Field Methods 550 or permission.

Speech Production:



Puzzle for the child (not my concern): what is in ?

Puzzle for the phonetician (not my concern): what is the nature of ?

Kikerewe: a Bantu language of Tanzania spoken here:

(on the dry land parts)



The noticeable (acquired) phones of Kikerewe (omitting tone) —

p t č k f s β v z b d ĵ g w l y h m n ɲ ŋ m: n: ɲ: i u e o a i: u: e: o: a:

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34

Please, no peeking

[v] [omudeléeva] “driver” — exhaustive list AFAIK
([βéeβá] “their husbands”, [kunéeβa] “to be flexible”, [ileeβa] “dance contest”).

[ŋV] [kuɲámúla] “to discover trickery”, [kuɲoloota] “to groan” — exhaustive list AFAIK
(but robust before [k,g])

[-nas]+ĵ [eĵaaházi] “jahazi dhow” [joomba] “maternal uncle”, [ekĵíko] “spoon”
([kuβúúɲja] “to peddle”, [embúúɲjá] “jigger” etc.)

[-nas]+b [ebalúúwa] “letter”, [kubómóka] “to fall apart”, [kuhútúbiya] “to make a speech”
([mbálá] “I kick”, [kuβáámba] “to tied down”, [mbe] “ja vel”)

The phonological rules of Kikerewe. The visible changes, and guidance from the invisible hand.

{a,b,c} is the unordered set {a,b,c}. {a;b;c} is “respectively”-ordering of sets. a~x indicates that the rule could be stated in terms of /a/ or /x/, given rule ordering: “before m~n” would mean “before m, or you could state it as before n if this rule precedes the change of /n/ to [m]”

Nasal Assimilation

n → {m;ɲ;ŋ} / __ {p,b,f,m; č,(ĵ),ɲ; k,g}

Input /m,ɲ,ŋ/ don't appear / __ C (except homorganic C) so this could refer to /n/ or /m,n,ɲ,ŋ/, depending on what is simpler. /n+ĵ/ doesn't arise but [ɲĵ] is plentiful. /n+ŋ/ doesn't arise. /v w/ don't appear in the context, and /y/ does something else, so these can be included or excluded *ad libitum*. /β l h/ → [b d p] so could be included or excluded, depending on what is advantageous. /nn/=n:/. /a,e,o,i,u,a:,e:,o:,i:,u:/ and /t d l s z/ could be directly excluded, or vacuously paired with output [n].

Postnasal Fortition

{β;l,h} → {b;d;p} / {m~n;n;m~n} ____

/w y v/ don't appear here or else they do something else. /f s z/, also /p t č k b d ĵ g/ do appear here, and are unchanged.

Vowel Harmony

{i;u} → {e;o} / {e;o; o} {p,t,č,k,f,s,β,z,b,d,ĵ,g,w,l,y,h,m,n,ɲ,ŋ}₀ ____

Rare /v/ is never attested in X₀. Note the famous Bantu asymmetry: /u/ → [o] only after /o/.

Fusion

a + {u,o; i,e; a} → {e;; o;; a;} word-internally only

Phrasal V-whack

{a,e}# {a;e;i;o} → {a;;e;;i;;o;}

/u/ does not appear word-initially; there are no long vowels word-initially or word finally.

Degemination

{a,a;; e,e;; i,i;; o,o;; u,u;}+{m;; n;; j;; ŋ;} → {a;; e;; i;; o;; u;}+{m; n; j; ŋ}

Pre-NC lengthening

{a;e;o;i;u} → {a;;e;;o;;i;;u;} / ___ {m;n;j;ŋ} {p,b,f; t,d,s,z; č,ǰ; k,g}

/mm~m:/ etc, so /mm, nn/ etc don't arise. /β l h/ go away, /v/ doesn't appear in this context.

Final Shortening

{a;;e;;o;;i;;u;} → {a;e;o;i;u} / __ #

Glide Formation

{i;u;o} {+,#} {a;e;i;o;u} → {y;w;w} + {a;;e;;i;;o;;u;}

o#V only arises phrasally; e#V similarly, but can be excluded by prior Phrasal V-whack

Coronal mutation

{t;l;d} → {s;z;z} / ___ i_x (a class of 3 morphemes beginning with /i/)

all consonants except [v] can appear before those morphemes

Rules that refer just to single segments

ky → č (kw is not changed)

y → ∅ / __ i (/wu/ does not arise, other glide+V are retained except that...)

y → ∅ /V [___ (optional: w does not occur)

y → ∅ /w ___ (yw does arise and isn't changed)

Now induce a grammar

Scanning for sets of segments with overlap in membership

Phrasal V-whack: {a;e;i;o}, → {a;;e;;i;;o;}

Degemination: {a,a;; e,e;; i,i;; o,o;; u,u;}, → {a;;e;;i;;o;;u;}

Glide Formation: {a;e;i;o;u}, → {a;;e;;i;;o;;u;}

Pre-NC length: {a;e;o;i;u}, → {a;;e;;o;;i;;u;}

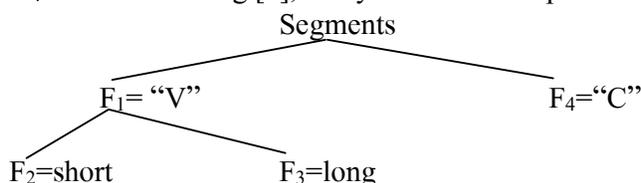
Final shortening: {a;;e;;o;;i;;u;}, → {a;e;o;i;u}

The preceding are included in {a,a;;e,e;;i,i;;o,o;;u,u;} (trigger for degemination), establishing F₁. F₁ is exhaustively partitioned into {a;;e;;i;;o;;u;} (output of degemination, pre-NC length) = F₂ vs. {a;e;o;i;u} = F₃ (output of shortening). This fully distinguishes long and short versions of vowels.

“Consonants”

{p,t,č,k,f,s,β,z,b,d,ǰ,g,w,l,y,h,m,n,j,ŋ} vowel harmony
{p,t,č,k,f,s,(β),z;b,d,ǰ;g,(w,l,y,h,m,n,j,ŋ)} pre-NC length

This motivates $F_4 = \text{“C”}$ excluding [v], and yields the strict partitioning:



Vowel height

Fusion: $\{u,o;i,e\} \rightarrow \{e; o\}$ (caused by /a/)

Vowel Harmony: $\{i(e);u(o)\} \rightarrow \{e;o\}$ (trigger /e,o/ for /i/; relation of /i/ to /e/, /u/ to /o/)

$\{u,o\}$ and $\{i,e\}$ observed in fusion: shared feature(s) define these sets. Harmony (of *i*) indicates that /e,o/ have F_5 which spreads. /i/, /e/ differ only in F_5 . Thus $F_5 \equiv \{e,o\}$ (for now).

What is retained when /a/ fuses? F_5 ? If /a/ were F_5 , it would trigger Harmony, *contra naturam*. So either (A) /a/ is not F_5 , or (B)[Ⓢ] it *is* and Harmony also refers to F_n . Since fusion only results in change to F_5 by combining with /a/, /a/ must have *some* property (cannot be fattiglæm).

☞ Can A be true? If /a/ is not F_5 , it is F_6 . Anticipating a feature for /i,e/ vs. /u,o/ where /i,e/ are F_7 , then: **A?**

i	e	u	o	a
	5		5	
				6
7	7			

a+i i.e. 6+7 gives an illegal vowel, repaired by $6 \rightarrow 5$ / [__7]. This fails to explain [a+u] \rightarrow [o]. Therefore, /u,(o)/ have F_8 , 6+8 is also illegal, and $6 \rightarrow 5$ / [__8]. These two rules merge with a generalized version of the structure-preserving rule if both 7 and 8 are grouped as F_9 . Thus:

i	e	u	o	a
	5		5	
				6
7	7			
		8	8	
9	9	9	9	

But $F_8 \neq F_9$ was superfluous: we only needed a way of referring to $\{i,(e),u,(o)\}$. So finally:

i	e	u	o	a
	5		5	
				6
7	7			
8	8	8	8	

☞ /a/ is F_5 but lacks a feature of /e,o/, and Harmony refers to *that* feature. F_7 is the difference /i,e/ vs. /u,o/; then /e,o/ (and not /a/) are F_6 . The high vowels could be F_6 , but we don't yet see any direct evidence for that. Thus: **B?**

i	e	u	o	a
	5		5	5
	6		6	
7	7			

Harmony spreads F_5 from F_6 . Fusion combines F_5 and the features of the second vowel, so /i/ (F_7) + /a/ (F_5) gives $[F_7, F_5]$, which does not exist but can be created by structure-preservingly supplying F_6 to any $[F_7, F_5]$. But again this fails for /u/ which is featureless (so far). Conclusion: /u/ is F_6 as well. /i/ can be underlyingly F_6 , eliminating the requirement for a F_6 -supplying mechanism. So:

i	e	u	o	a
	5		5	5
6	6	6	6	
7	7			

Notice that A and B are essentially identical, except whether a feature identifies /a/. The B solution posits fewer feature entities, so is assumed, following Ockham (1321) and forthcoming.

Vowel color

GF indicates a relation between /i;u/ and /y;w/: /i,y/ is F_7 and /u,w/ not; or vice versa (i.e. $F_1 \rightarrow F_4$ and /i,y/, /u,w/ are the same except for F_1/F_4). Since $ky \rightarrow \check{c}$ and w does not trigger, /y/ must have a mark. Thus /i,y/ are F_7 and /u,w/ are not. Additionally, mutation is triggered by /i/ (lexically specified), and /y/ deletes after a glide but /w/ actively doesn't. This gives a complete differentiation of vowels (above), based exclusively on the metatheoretical principle that a segment has a feature F_n only if it is required to distinguish the segment from other segments not so referred to in a rule.

Special problems (cracks to not trip on)

Harmony: /i/ lowers after {e,o}; /u/ lowers after /o/, not /e/. Maybe /o/ has a property which /e/ lacks (a new feature). If u -lowering is a separate rule, /o/ must be distinguished from /e/ and /u/ (non-triggers): /o/ is F_8 , /e/ is not.

/a/ fuses word-internally, deletes phrasally. /e,o/ not prevocalic in words; phrasally, all V-combinations are possible except $V\#u$. So regarding deletion (a,e) and glide formation (i,o,u): (A) 2 rules or 3? and (B) GF before deletion, or the opposite?

1. GF(i,u), GF(o) > Deletion(V)
2. Deletion(a,e) > GF(V)
3. GF(i,u,o) > Deletion(V)

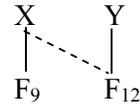
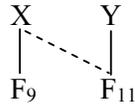
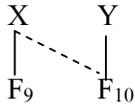
1 requires a feature for /i,u/. 2 requires a feature for /a,e/. 3 requires a feature for /i,u,o/. /o/ triggering u -lowering shows that /o/ has a feature distinguishing it from /e/. Simplest solution: the feature for u -harmony is the feature for phrasal vowel patterns, ruling out 2. 1 differs from 3 only in positing 2 rules rather than 1 for GF, and requiring a feature for /i,u/. Thus F_8 for /i,u,o/. GF applies to F_8 vowels, and deletion applies to any remaining V-V. Final answer:

i	e	u	o	a	
	5		5	5	
6	6	6	6		
7	7				
8		8	8		Vowel length and C/Vstuff is by magic.

Consonants

GF reveals the features of /y/ and /w/. The main C rules are Nasal Assimilation, Postnasal Fortition, Degemination, Pre-NC lengthening, and Coronal mutation. Nasals are actively referred to by Degemination and Pre-NC lengthening (context of N: and NC). Thus /m n ɲ ŋ/ are F_9 .

Nasal Assimilation motivates 4 features: F_{10} shared by /h/, b, (β), f, p, m/; F_{11} shared by /č, (j), ɲ/; and F_{12} shared by /k, g/. A first approximation:



“Feature” can be feature-sets, so $F_{13} = \{F_{10}, F_{11}, F_{12}\}$. Spreading F_{13} carries the subordinate features, as usual. Features are not recursive (features do not dominate themselves). Structure is “flat” in lieu of evidence for grouping (simultaneous operation on multiple features).

F_{14} refers to /β l h/ (fortition): {β,b}, {l,d} and {h,p} are identical except that the first member is identified by F_{14} and the second is not. Then $F_{14} \rightarrow \emptyset / F_9$ ____. F_{15} identifies the targets of Coronal Mutation (/l d t/). Unless the change is deletion of F_{15} , {s,z} are F_{15} since they derive from segments that are F_{15} . The sets {t,s} and {l,d,z} are featurally identical except for one new feature (plus F_{14} which distinguishes {l,d}). This gives the partial solution:

f	p	t	č	k	β	v	b	d	ǰ	g	l	s	z	h	m	n	ɲ	ŋ
															9	9	9	9
10	10				10		10							10	10			
			11						11								11	
				12						12								12
13	13		13	13	13		13		13	13				13	13		13	13
					14						14			14				
		15						15			15	(15)	(15)					

Not yet distinguished: {f,p,b}, {t,d,(s,z)}, {č,ǰ}, {k,g}, entailing at least F_{16} . If {s,z} are not F_{15} , we must distinguish among {v, s, z}. Assume that {s,z} are F_{15} . Because /k,č/ have common feature(s) not found in all consonants. Either /k,č/ are F_{16} and /g,ǰ/ are not, or vice versa.

This leaves just {f,p,b}, {t,d,s,z}, entailing F_{16} or another feature). *Ex hypothesii* /β,l,h/ and /b,d,p/ differ only in the former having F_{14} . So either /l,d/ are F_{16} and /t/ is not, or vice versa. Given the coronal-mutation analogy, we would also put /z/ in a set with /l,d/ and /s/ in with /t/. No other rules refer specifically to the undistinguished segments (we cannot decide between marking {p,t,č,k} and leaving {b,d,ǰ,g} unmarked or the converse based on rules). Apart from {k,č} relationship, nothing indicates that {p,t,č,k} that are a class — perhaps {b,d,č,k} vs. {p,t,ǰ,g}?

There are at most 4 segments in an undifferentiated set, so two features suffice: F_{16} and F_{17} . Values can be assigned at random (keeping {k,č} together, also {β,b} and {h,p}). In the row for F_{16} , indices on marked values and the corresponding non-subscripted letter indicate which segment the *i*th phoneme needs to have an opposite value from. Voiceless consonants and fricatives arbitrarily marked for a feature.

f	p	t	č	k	β	v	b	d	ǰ	g	l	s	z	h	m	n	ɲ	ŋ
															9	9	9	9
10	10				10		10							10	10			
			11						11								11	
				12						12								12
13	13		13	13	13		13		13	13				13	13		13	13
					14						14			14				
		15						15			15	15	15					
	16 _i	16 _j	16 _k	16 _k	i		i	j	k	k	j	16 _j	j	16 _i				
17 _i		j					i	k			k	17 _j	17 _k					

Feature Recycling

Vowels use F_{5-8} , and are distinguished from consonants with F_1 and F_4 , so these can be used for consonants. Features can be reduced by using F_{5-8} on consonants (deleting 4 consonant features). It does not initially seem to matter which 4 are “recycled”. If features are recycled, rules may be restated, e.g. vowel harmony only applies to an F_1 segment that is F_6 . The cost of feature reduction is rule-complication; this brings up the classic *praeter necessitatem* question. Also be sure that the glides can be dealt with. But: $ky \rightarrow \check{c}$ indicates that $F_7 = F_{11}$, hence *necessitatem*.

Concretely, assume a language with $[i,e]$ and $[k,q]$; also assume sufficient active evidence for partitioning segments into “consonants” ($=F_1$) and “vowels” ($=F_2$). Further assume no rule natural class evidence showing that $[i:e::k:q]$. Finally assume a rule that refers to $[e]$ to the exclusion of i and also q .

<p>More features, simpler rules</p> <p>$/[F_1],[F_2],[F_1,F_3],[F_2,F_4]/$</p> <p>$/e/ = [F_4]$</p>	<p>$/k q i e/$</p>	<p>Fewer features, more complex rules</p> <p>$/[F_1],[F_2],[F_1,F_3],[F_2,F_3]/$</p> <p>$/e/ = [F_2,F_3]$</p>
---	-------------------------------	---

Does the number of rules picking $/e/$ matter, i.e. 2 such rules favors “more features, simpler rules”? I say yes, and that’s due to the nature of the concept “generalization”. Generalizing meaning “abstracting over two or more similar things, focusing on the common properties and setting aside the differences”. To disregard the common characteristic is “missing a generalization” about a common cause, in violation of the Obligatory Generalization Principle. When you have a single instance (rule, for example), there is no “generalization”, just an ungeneralizable observation. Then the tradeoff is to posit a new feature, vs. add mention of a feature in a rule. Here, I argue that the evidence for a new feature isn’t compelling.

Since F_9 includes both F_7 and non- F_7 members (given $F_7=F_{11}$), then F_9 cannot stand for a feature for so-called “place of articulation”, idem F_{13} . Plausible candidates for feature-collapsing include (consonant) F_{10}, F_{12}, F_{15} , which could map to (vowel) F_5, F_6, F_8 . No grammatical facts force a particular mapping, so randomly map $F_{10}=F_5, F_{12}=F_6, F_{15}=F_8$ (divide by 2), and renumber:

f	p	t	č	k	β	v	b	d	ǰ	g	l	s	z	h	m	n	ɲ	ŋ
															9	9	9	9
5	5				5		5							5	5			
			7						7								7	
				6						6								6
10	10		10	10	10		10		10	10				10	10		10	10
					11						11			11				
		8						8			8	8	8					
	12 _i	12 _i	12 _k	12 _k	i		i	j	k	k	j	12 _j	j	12 _i				
13 _i		j					i	k			k	13 _j	13 _k					

See Odden (1992) for discussion of vast-search problems in deriving maximally-cheap URs, both w.r.t. count of used features and the total volume of ink in writing out the inventory (no consideration of lexical or speech-token frequency). I assume only type-minimalist assumptions about features, viz. posit no feature that lacks data-based justification, and posit no limits on the use of justified features which are not themselves justified by facts, so smoke ‘em if you’ve got ‘em.

Empirical Underdetermination Justified Escape Clause

Suppose Language₁₉₉₉ presents no grammatical data to indicate whether $\{p,b,m\}$ have the feature that groups $\{u,o,\ddot{o}\}=F_\alpha$, or the feature that groups $\{i,e,\ddot{a}\}=F_\beta$; but $\{p,b,m\}$ are a phonological group, and under the terms of the feature-nonproliferation treaty, these consonants must have one of these two existing features, rather than causing creation of a new feature. In lieu of definitive

evidence, speakers would be expected to map at random, so that $\approx 50\%$ of the population conclude $\{p,b,m\}=F_\alpha$ and $\approx 50\%$ conclude that $\{p,b,m\}=F_\beta$. Now assume language change over 1 year due to social disruption from the movie Borat, where Language_{1999} maps to Language_{2007} , and Language_{2007} introduces data which crucially establish that $\{p,b,m\}=F_\alpha$. Such scenarios must be studied, to see if they falsify the random-mapping hypothesis.

Suppose that Language_{2007} establishes $\{p,b,m\}=F_\alpha$ behavior by all speakers — does this falsify formally contentless phonology (which, in lieu of nativist slavish obedience to phonetics, says that speakers make random assignments)? No, because grammars do not change in an uncaused fashion: $\forall i \forall j (L_{t_i} \neq L_{t_j} \text{ (} t_i < t_j \text{)})$ iff some new fact forces non-identity. I.e.: established conclusions are not rejected when they *can* be, they are rejected when they *must* be. Observe the longitudinal acquisition of an adult grammar.

Time	<i>Ambient data</i>	<i>Child conceptual linguistic form</i>	<i>Analysed as</i>
T ₀	[ʃɔni kəm hɪr] (Kid gets a clue that people sometimes mean something by using [a] versus [i]. He guesses that you have to pay attention to that acoustic scrunched-together pattern versus the spread out pattern).	[b ¹ ab ¹ ab ¹ ab ¹ a]	[F ₁][F ₂][F ₁][F ₂][F ₁][F ₂][F ₁][F ₂]
T ₁	[ʃɔni kəm hɪr] (Kid gets a clue that people sometimes mean something by using different consonants, so the buzzy sound is important).	[b ¹ ab ¹ ib ¹ ab ¹ i]	[F ₁][F _{2,F3}][F ₁][F ₂][F ₁][F _{2,F3}][F ₁][F ₂]
T ₂	[ʃɔni kəm hɪr] (Kid gets a clue that the bottle-sounding consonant is actually different from the hissy-sounding one)	[b ¹ ab ¹ ip ¹ ab ¹ i]	[F ₁][F _{2,F3}][F ₁][F ₂][F _{1,F4}][F _{2,F3}][F ₁][F ₂]
T ₃	[ʃɔni kəm hɪr] etc.	[ʃab ¹ ip ¹ ab ¹ i]	[F _{1,F5}][F _{2,F3}][F ₁][F ₂][F _{1,F4}][F _{2,F3}][F ₁][F ₂]

A child does not initially know that A and B are distinct phonemes of the language, and only learns this by inducing consistent form/reference relations: i.e. [ba] “sheep” at first sounds linguistically the same as [pa] “dad” and the child thinks the language has a lot of homophones,² since the sensorily-available difference doesn’t seem to be important for the symbolic system at that moment.

Early phonological error correction reduces to seeing that certain acoustic facts may be important. Just as adult fieldworkers have to learn to distinguish between [h] and [h], [ʌ] and [≠] or [u] and [u], children have to learn to attend to classes of acoustic cues (≠“attend classes on acoustic cues”). This can give rise to largish and apparently phonetically-based new classes, assuming that the Saami child doesn’t learn the surface distinction [hp ht hc hč hk] as 5 separate and unrelated things. Once [F_n] is induced as a means of distinguishing a bunch of sounds, [F_n] will persist until overridden by direct evidence (i.e. the rules of grammar require different feature assignments).

Final Note

The child induces a hierarchy of phonological ideas by generalizing from concretes to higher abstractions that subsume the concretes:

$$n \rightarrow m / _ p; n \rightarrow m / _ b; n \rightarrow m / _ f; l \rightarrow d / n _ ; h \rightarrow p / m _$$

I pre-unified the rules at this level into obvious generalizations. Well, my knowledge of “single rule” is *highly* influenced by nativist feature theory. Syntactically, the phoneme-specific generalizations “ $n \rightarrow m / _ f$ ” and “ $l \rightarrow d / n _$ ” aren’t a rule (because of focus/determinant order). Simi-

² Sometimes this is due to the child thinking that his dad *is* a sheep.

larly, Vowel Harmony and Fortition don't unify into one large "after" rule, since harmony has a variable C_0 . It seems to turn out that all rules of Kerewe can be distinguished on such formal grounds.

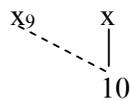
This leaves open the possibility of unifying the following sets into one rule which adds the feature F_3 to final segments in some language:

$$b \rightarrow p / _ \#; g \rightarrow k / _ \#; i \rightarrow e / _ \#; u \rightarrow o / _ \#; n \rightarrow \eta / _ \#$$

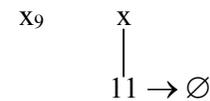
Fortunately, I don't have this problem in Kerewe. I might, postnasally, in Bukusu. The question of what a formally possible rule is becomes excruciatingly important, so if it turns out that in reality Rule Theory must allow "immediately after A or before the sequence BC" as a single rule, substance-free phonology might be nigh impossible.

Some Rules. X_n means "a segment which is F_n ", and is equivalent to autosegmental $X-F_n$.

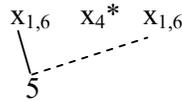
Nasal Assimilation



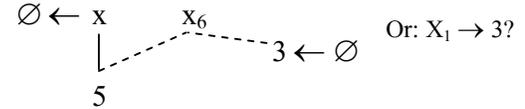
Postnasal Fortition



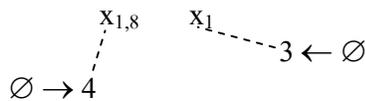
Vowel Harmony



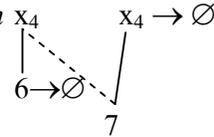
Fusion



Glide Formation



Palatalization



You might imagine inductively extracting CV/ μ phonology to handle compensatory lengthening.

The syntax of rules can't be "anything you want" in MFP: otherwise, Kikerewe would have a single horking rule:

$$\left\{ \begin{array}{l} h \\ t \\ l \\ n \\ i \\ e \\ \beta \end{array} \right\} \rightarrow \left\{ \begin{array}{l} p \\ s \\ \langle d, z \rangle_i \\ \langle m, n, \eta, \emptyset \rangle_j \\ \langle e, y \rangle_k \\ \emptyset \\ b \end{array} \right\} / \text{etc} \quad \left\{ \begin{array}{l} n _ \\ _ i \\ \langle n \rangle_{i\alpha} _ \langle i \rangle_{i\beta} \end{array} \right\}$$

But imagine a language with just these rules:

$$t \rightarrow s / _ i \quad t \rightarrow n / _ m \quad t \rightarrow r / _ a \quad t \rightarrow l / _ l$$

Rules must be collapsed if they can be (see SPE conventions); this must be one rule; therefore the features that distinguish t from s , n , r , l must be spread as a group, so F_5 dominates F_1 (s, i), F_2 (n, m), F_3 (r, a), F_4 (l), and the rule is "Spread F_5 ". If this good? bad? Only brain surgery will tell us for sure.