Andrew Nevins
anevins@mit.edu
Harvard, 24 Feb 2004

Phonological Ambiguity
(What UG can and can't do to help the reduplication learner)

1. Underdetermination. As C.S. Peirce has discussed, the problem of *underdetermination* in the pursuit of knowledge about the world and generalization over phenomena is, perhaps a logical necessity: "Observed facts relate exclusively to the particular circumstances that happened to exist when they were observed. They do not relate to any future occasions upon which we may be in doubt how to act. They therefore do not, in themselves, contain any practical knowledge. Such knowledge must include additions to the facts observed. The making of those additions is an operation which we can control. How is it that Man ever came by any correct theories about Nature? We know by Induction that man has correct theories, for they produce predictions that are fulfilled. But by what process of thought were they ever brought to his mind? A chemist notices a surprising phenomenon. He will remember that Mill tells him that he must work on the principle that, under precisely the same circumstances, like phenomena are produced. Why then does he not note that this phenomenon was produced on a certain day of the week, the planets presenting a certain configuration, his daughter having on a blue dress, he having dreamed of a white horse the night before, the milkman having been late that morning, and so on? How was it that man was ever led to entertain that true theory? You cannot say that it happened by chance, because the possible theories, if not strictly innumerable, at any rate exceeded a trillion. " (p.237-238).

Underdetermination, then, simply refers to the fact that a phenomenon of interested is always accompanied by hundreds of irrelevant details. Consider the Chinese youth trying to learn why her parent repeated the word *tian*, meaning "day", when using universal quantification over days: saying *tiantian*, meaning "every day". She must ignore not only the happenstance of location, time, deixis, and clothing: paying attention to strictly linguistic factors, if she is to attempt to generalize at all (and apparently every Chinese youth does!), she might posit that *tiantian* has the form it does because universal quantification is reflected by repetition for all words that begin with the sound *t*. This would be the wrong hypothesis for Chinese. It would also be the wrong hypothesis for Temne, a West African language that turns out to have an identical means for quantifying over nouns: reduplicating the whole noun, regardless of the properties of any of the individual phonemes that compose the word. And as Peirce's quote illustrates, there are countless other possibilities: *tian* happens to be a word with high-tone, but no reduplication process in Chinese, or in fact, in the world (at least that I have found) is sensitive to the tonal properties of the input.

Let's continue where we left off with Peirce: "Besides, you cannot seriously think that every little chicken that is hatched, has to rummage through all possible theories until it lights upon the good idea of picking something up and eating it. On the contrary, you think the chicken has an innate idea of doing this; that is to say, that it can think of this, but has no faculty of thinking of anything else. The chicken you say pecks by instinct. But if you are going to think every poor chicken endowed with an innate tendency toward a positive truth, why should you think that to man alone this gift is denied? I am sure that you must be brought to acknowledge that man's mind has a natural adaptation to imagining correct theories of some kinds".

Peirce's chicken, who manages to have the idea of picking up food, and not countless other ideas, has something in common with the Chinese youth, who manages to have the idea that when her parent says *tiantian*, that it reflects a general process of repeating a set of phonemes, and is not a paralinguistic outburst, or a process limited to nouns with diphthongs. What they have in common is that the number of hypotheses they consider is restricted. That is, what their "innate tendency toward a positive truth" does is not lead them *towards* the positive truth persay, but rather *away* from numerous other irrelevant untruths.

Many critics of domain-general learning theories argue that domain-general architectures, such as neural networks, secretly "build in" a lot of innate structure, so that their simulations which appear to be learning complicated linguistic phenomena are doing so with a headstart. I want to argue that these critics are focusing on the partly-full nature of the glass, but that what is *not* there is even more important. What allows connectionist networks to succeed, when they do, is not what they have been built to bring to the task, but rather what they are specifically built *not* to bring to the task. It is for this reason that every connectionist network that has ever been programmed into a computer to learn, say, how the past tense is formed in English, through statistical tendencies, has already been built with no predicates or functions that can count the number of syllables in the input, with no representation of prosodic stress, and with no subroutines that can determine whether the word is palindromic. If connectionist networks kept statistics about *every* linguistic property inherent in the data, they would never be able to make any generalizations.

The function of what is called Universal Grammar, then, is not really to *provide* a grammar, but rather to provide a set of constraints on what can and can't be a possible grammar. What I want to explore today is how Universal Grammar aids the learner of reduplication the most: not by giving him or her the answer magically from scratch – in fact, as we will see, the learner may *never* learn a uniquely correct answer. But what Universal Grammar can do is *prevent* the learner from ever considering a host of irrelevant answers. And I do mean prevent. The closest analog to Universal Grammar in another species is probably in the songbirds. W.H. Thorpe's 1958 study of trying to teach chaffinches the songs of the tree pipit and seeing their failure to learn it (although it was sensorily perceptible and motorally executable) led him to conclude that "the chaffinch has the inborn blueprint conferring on it a tendency to learn to pay attention to certain kinds of sounds and certain types of phrase only" (p.84). Thorpe's quote, and indeed his views, emphasize that the defining feature of songbird learning is what they cannot and do not try to learn.

On the picture of Universal Grammar developed here, it is not a set of *do*'s in theory construction, that guide the learner of reduplication through a specific set of steps in learning reduplication, but rather, by virtue of its circumscribed search space, rather, in effect, a set of *don't*s for what hypothesis never to bother with. I will try to convince you that the actual mechanism that the learner exercises in pursuing "a positive truth" in reduplication is actually quite dumb: it's what computer scientists call a "British Museum search", because it is likened to trying to find a pottery shard in the British Museum by simply looking through the whole museum. However, what I want to argue here is that even the “British Museum” part of the “British Museum search” is a powerful enough restriction on any search (say, because it excludes the Tate Gallery): excluding irrelevant domains of search, and automatically dismissing irrelevant sources of generalization, is all that UG seems to do for the learner, at least with respect to the phenomena we will discuss here.

The problem of underdetermination is aided, then, by restricting the family of possible generalizations to only include certain factors. We know, by looking at reduplication across the world’s languages, that it contains a stunning array of variation. And even within languages sometimes: while Chinese uses *total* reduplication to convey ‘every X’ for nouns, it also has a process of *interleaving* reduplication, used to form adverbs from some adjectives. Thus, while *gaoxing* means ‘happy’, *gaogaoxingxingde* means ‘happily’. When we turn our attention to other languages, we find apparently tremendous variation from these patterns: Turkish intensifies the adjective *dolu*, meaning ‘full’, and the adjective *beyaz*, meaning ‘white’, as *dopdolu* and *bembeyaz*, respectively. The generalization for Turkish can be extracted from these two examples kind of easily: reduplicate the first consonant and vowel, and add a fixed bit of segmental material[1]. Finally, we can consider Lakhota, which repeats everything after the first vowel in a subset of the lexicon (Adam Albright, personal communication): *naxmá* becomes *naxmaxma* and *yámni* becomes *yamnimni*. What all of these patterns have in common, however, is that the amount of copied material is blind to the segmental properties of the input, and blind to a ton of other potentially linguistically relevant factors, such as whether the word has an odd or even number of consonants. The reduplication learner, then, will best be helped by only considering hypotheses for how the process works that involve delimiting where in the word the copied sequence begins and ends. The Chinese learner needs to know that one pattern involves copying from the first segment to the last segment; the Turkish learner needs to deduce that one pattern involves copying the first vowel and first consonant, and the Lakhota learner must figure out that one pattern involves copying from after the first vowel until the last segment. What we will turn to next is an attempt to help the learners a bit more (though not too much) by deducing these endpoints from a set of primitives.

2. Ambiguity. Underdetermination, as we saw, refers to the problem of deciding what of hundreds of relevant factors to consider in hypothesis formation. Ambiguity, on the other hand, refers to the problem of deciding of which of dozens of apparently equivalent rules or structures might actually underlie the phenomena at hand. One of my favorite examples from syntax has always been *I saw the man with the telescope*. The ambiguity of this sentence arises because English has two different structures that underlie verb phrases: one with only a direct object, and one with a direct object and a prepositional phrase. Ambiguity thus refers to the fact that the same string can correspond to two different linguistic structures: one is VP NP PP and one is VP NP, where NP is further decomposed into NP PP. Importantly, ambiguity has nothing to do with “two different meanings”, persay. It is only because these two different syntactic structures provide the

---

[1] More precisely, look at the upcoming consonant. If it is a liquid or a nasal, preprend the labial stop /p/, otherwise, preprend the labial sonorant /m/.

foundation for semantic composition that meaning turns out to be involved. But consider the famous reversible goblet illusion, where in a single image, viewers perceive either two faces, or one goblet. The ambiguity is one of assigning structure to a visual input, and "meaning" is only involved insofar as faces or goblets are then treated as items in the world with a significance. The focus in linguistics on syntactic/semantic ambiguities has led many people to associate ambiguity with "two possible meanings", but I will employ its more general use here, one in which an observed phenomenon can correspond to two or more distinct mental structures, regardless of their particular representational vocabulary. I will adopt this use for reasons to become immediately clear.

Suppose that we limit our vocabulary of reduplication to include phenomenological concepts such as the following (to be revised):

(1) total reduplication: Chinese *tiantian*
(2) first-syllable reduplication: Ilokano *kal-kaldin*, *tak-takder* (Hayes and Abad 1989)
(3) last-syllable reduplication: Kaingang: *va-va, jemi-mi, vasan-san* (Poser 1982)
(4) iterative reduplication: Chinese *gaogaoxingxing*
(5) stressed-syllable reduplication: Chamorro *hu-ga-gándo* (Topping 1973)

The problem of ambiguity rears its head forcefully in this case. Consider the Kaingang learner, who, as it turns out, happens to hear *va-va* as his first encounter with verbal plurality. Since he has UG on his side to take care of underdetermination, he is luckily able to ignore extralinguistic factors, and in fact, he even deems irrelevant the fact that this word ends in a low vowel and that it bears a falling tone. If (1)-(5) are among the analyses of the phenomena that UG has left open, he will not be able to choose among any of them based on this example alone.

That is, Kaingang *va-va*, just like English *I saw the man with the telescope*, is ambiguous. It looks like total reduplication. It also looks like first syllable reduplication, operating on a word that happens to be monosyllabic. It also looks like last syllable reduplication, since in monosyllabic words, the only syllable is indeed the last syllable. The same remarks hold for iterative reduplication, a process that reduplicates each syllable one-by-one, applying trivially here. Finally, stressed-syllable reduplication is a possible hypothesis here as well. The string *va-va* is thus (in the range of options delimited here) five-ways ambiguous: it corresponds to five different structural representations.

But on to those representations, which we haven't really formalized yet. As I mentioned above, "iterative" and "total" are phenomenological predicates. What reduplication does is repeat a sequence of phonemes within a word. Hence one basic starting point for any learner's hypothesis about reduplication in his or her language is to generalize over the end points of the copied sequence.

Typological investigations of reduplication reveal that it does not target any random substring within the word. In fact, the domain of "anchor points" that delimit the beginning and ending of the copied sequence come from a finite list:

(6)

- Starting at, or immediately following the first consonant
- Starting at, or immediately following the first vowel

- Starting at, or immediately following the first segment
- Starting at, immediately preceding, or immediately following the stressed vowel
- Starting at, or immediately preceding the last vowel
- Ending at, or immediately following the first vowel
- Ending at, or immediately following the stressed vowel
- Ending at, or immediately preceding the last vowel
- Ending at, or immediately preceding the last segment
- For every vowel, starting at the segment immediately preceding the vowel and ending at the segment immediately following it

One question arises is whether this list represents a random collection of anchor points within the word, or whether there are any systematic regularities. And of course, there are. The predicates *first, last,* and *stressed* play a defining role. These points are known to be perceptually prominent, as discovered in the work of Pierrehumbert and Nair (1995), and to play an important role in phonotactics distribution, as explored in the PhD thesis of Jill Beckman (1998). Even more immediately relevant for our research here is that these predicates play an important role in determining the distribution of infixes, as shown in a forthcoming book by Alan Yu of the University of Chicago. If UG is to further constrain the search space of reduplication hypothesis, it should limit the domain of guesses to around the neighborhoods of the first and last vowel and consonant, and the stressed vowel.

When establishing linguistic generalizations takes place through varying the values of certain predicates, we can speak of a parameterization of the set of generalizations. The set of generalizations about the anchor points targeted in reduplication can thus reduce the unstructured list above into a set of values for the following open properties of the beginning and ending of the reduplicative sequence:

(7) Parameters of Anchor Points (first approximation)

- First, last, stressed, or every
- Vowel, consonant, or don't-care (i.e., segment)

The next point of interest in the unstructured list from above is that it often contains reference to the segment immediately preceding or immediately following a given anchor point. For example, a generalization about Kaingang forms such as *jemi-mi* and *vasan-san* is that the copied sequence begins with the segment immediately preceding the last vowel, and continues to the last segment in the word.

(8) Kaingang reduplication: Copy the sequence that **Begins** with the Segment Immediately-preceding the last vowel and **Ends** with the Last Segment

Consider next what is called *Out-of-Control* reduplication in Lushootseed by Bates, Hess, and Hilbert in their comprehensive dictionary. Descriptively, it copies the first VC sequence: ʔibəs, meaning 'to walk', can undergo reduplication and become ʔibibəs, meaning 'to pace back and forth'.

(9) Lushootseed Out-of-Control reduplication: Copy the sequence that **Begins** with the first vowel and **Ends** with the segment immediately-following the first vowel

Like many other Salish languages, Lushootseed has more than one type of reduplication. Thus, the same word ʔibəs can undergo what is called *distributive* reduplication, and is used to describe a verb action distributed in space or time. Distributive reduplication is descriptively a copy of the initial CVC sequence:

(10) Lushootseed Distributive reduplication: Copy the sequence that **Begins** with the segment immediately preceding the first vowel and **Ends** with the segment immediately-following the first vowel

A brief excursion into a few more reduplicative patterns has revealed that while the anchor points are crucial in determining the area within a phonological word where reduplication will begin and end, it is nonetheless also the case that the segments immediately flanking each anchor point on both sides are relevant landmarks for reduplication as well. The following parameterization thus characterizes all of the target sites for reduplication:

(11) Parameters of Anchor Points (second approximation)

- The segment **defined-by**, **immediately-preceding**, or **immediately-following**: {0,1,2)
- The **Segment, Vowel**, or **consonant** {0,1,2}
- That is the **first**, **last**, **stressed**, or **every** in the word {0,1,2,3}

Reduplication can now be described as a pair of values within a discrete search space of twenty-four values for each endpoint. It is crucial to emphasize that parameterization, on this view, assumes a fixed set of representational objects (i.e. stressed vowel), and that all of the variability in hypothesis formation is in the *modes of combination* of these basic predicates.

(12) Total reduplication: <0,0,0> to <0,0,1>
Chamorro stressed CV: <1,1,2> to <0,1,2>
Lushootseed first CVC: <1,1,0> to <2,1,0>
Lushootseed first VC: <0,1,0> to <2,1,0>
Kaingang last CV: <1,1,1> to <0,1,1>
Iterative reduplication: <1,1,3> to <2,1,3>

Before proceeding, I should mention that I have not yet specified the mechanism by which speakers "copy a sequence". I assume a *Multiprecedence-and-Linearization* model, in which reduplication results from the addition of new precedence relations to an existing word. In other words, while the anchor points specify the endpoints of reduplication, I do not assume that they do so merely like the junctures |: :|, which, in musical notation, specify the start and endpoint for a sequence. Instead, I assume that the most primitive and irreducible relation in phonology, immediate precedence, is employed in characterizing reduplication. Thus, total reduplication is not implemented in the phonology by a statement saying “Copy from here to there”, but rather by the addition of a basic relation: in addition to the first segment immediately preceding the second, the second immediately preceding the third, and so on, until the next-to-last immediately precedes the last, I assume that the phonology of total reduplication involves the addition of one simple statement: that the last segment immediately precedes the first segment.

This view of reduplication as multiple-precedence, which leads to multiple copies on the surface, may be likened to movement relations in syntax, where multiple-dominance within a tree leads to multiple copies of a phrase: it is called “movement” by some syntacticians, but in reality, all that the phrase marker contains is a new relation between two representational objects, with particular consequences once it is realized at the surface (such as the pronunciation of the phrase in only its “higher relation” in the case of overt movement). The multiple-precedence view of reduplication yields a deterministic output, as I have shown in computer implementations with Aaron Iba, demonstrated at the LSA this year. There are a number of important consequences of this model, but in the interest of time, I will sketch only one here.

Recall that Lushootseed has two kinds of reduplication: *Out-of-Control* and *Distributive*. It turns out that, just as when we use the adverbs “uncontrollably and repetitively” in sequence for a single verb, the Lushootseed speaker sometimes chooses to apply both Out-of-Control and Distributive reduplication to a single verb at once. In other words, the speaker who produces *bali* to mean ‘forget’, and *balali* to mean ‘suddenly and unexpectedly forgetting’ in the Out-of-Control, and *balbali* to mean ‘forgetting more than one thing’, in the diminutive, may want to express ‘suddenly forgetting more than one thing’, with both diminutive and out-of-control reduplication applying to the verb root *bali*.

The simplistic *start-and-endpoint* view of reduplication will yield a structure as in (12c) for the combination. It is clear that no musician would be able to interpret this structure in a deterministic fashion, and that a range of interpretations might be assigned to it, yielding, for example, *balbalali, balalbi*, *bbalali*, and so on. Lushootseed speakers, however, are unanimous on this and all comparable forms: they are *balalbali*. This result is entirely natural in the view of added precedence relations, and I refer you to the paper for a discussion of a deterministic algorithm that linearizes correctly, as that is the subject of another lecture altogether.

(12) a. Lushootseed first CVC: <1,1,0> to <2,1,0> **|:** bal **:|** i
b. Lushootseed first VC: <0,1,0> to <2,1,0> b **|:** al **:|** I
c. Combination: **|:** b **|:** al **:|** **:|** I

Today’s talk is not about the linearization of multiple reduplication, but rather the problem of learning any kind of reduplication given the number of possible compatible hypotheses. The most important point for our present purposes, then, is that reduplication is defined by the anchor points within the parametric system we have developed here, and that it is implemented by an elegant view of precedence structures within phonology.

We cannot, however, abandon out Kaingang learner, whose plight arises when first encountering *va-va*. He or she will face a parametric ambiguity, as the output form, and any attempt at subsequent generalization, is compatible with:

(13) Kaingang *va-va* compatible with:
<0,0,0> to <0,0,1>
<1,1,2> to <0,1,2>
<1,1,0> to <0,1,0>
<1,1,1> to <0,1,1>
<1,1,3> to <0,1,3>

The difference between underdetermination and ambiguity, then, is that ambiguity is always within a parametric system that has domain-specific objects and variable modes of combination. UG, by providing the representational objects within which the hypotheses can be couched, solves the problem of underdetermination. But once the modes of combination are allowed to vary freely, UG can do nothing to solve ambiguity. That is, there is nothing in Universal Grammar that will tell the Kaingang learner which of these parameter values is the correct one. Recall that the view developed here is that UG limits the space of possible hypothesis, but does nothing more. I submit that UG does not and cannot help the Kaingang learner, nor any reduplication learner, any further with reduplication beyond already providing the parameters in (11). In the next section, I will argue that there is empirical support for this view.

3. Variation as the Consequence of Parametric Ambiguity. How is the Kaingang learner to choose among all of the options in (11)? The suggestion here is that he does not. That is, the model developed here is one of a *parallel* search. It leads to the view that speakers have “multiple grammars” simultaneously – though of course, by grammar, we are referring only to a very small subfragment of the phonology: how reduplication works in the ambient language. On this view, the variation that has been observed in real production data, most thoroughly by sociolinguists such as Labov, is the consequence of multiple solutions to the same problem, or, put another way, variable formulations of the rules.

A convergence of recent research has suggested that the parallel-search view accords well with what is actually observed in production data. Anttilla (1997) treats free variation in the selection of Finnish genitive plural allomorphs as the result of crucially unranked constraints; more generally, this represents a two-grammar view. Boersma & Hayes (2001) examined the distribution of light and dark [l] in English. Their subjects reliably chose light [l] in word-initial position (e.g. *light*) and dark [l] in preconsonantal or word-final position (e.g. *bell, help*). However, they reported free variation for intervocalic, posttonic [l], in proper names such as *Mailer* and *Greeley*. Their idea is that among the constraints determining the allophony of [l], there are two constraints: “[l] is dark” (context-free) and “prevocalic [l] is light” are freely ranked with respect to each other. In most cases, the ranking chosen between these constraints won’t matter, as higher constraints choose the allophone anyway. But in the case of *Mailer*, no higher-ranked constraints determine the allophony, and the freedom of ranking yields either *Mailer* with a light [l] or a dark [l]. Before the details of these analyses concern us too much, it is perhaps important to emphasize that pairwise rankings are almost always intertranslatable with binary parameters. For example, the well-known *head-parameter* in syntax could of course be modeled in a constraint-framework as *Head-Initial outranking *Head-Final, or vice-versa. The point is that the examples of variation, modeled by crucially unranked pairwise constraints above, are formulable as multiple parameter-settings as well.

Returning to the model at hand, the guiding idea is that variation is the result of maintaining multiple parameter settings simultaneously. Consider the variation in reduplication in Ilokano, reported in Hayes & Abad (1989), a descriptive article. As shown above, Ilokano reduplicates up to the consonant immediately following the first vowel. Thus, in *takder*, speakers copy the sequence *tak,* and in *kaldin* they copy the sequence *kal*. I will represent this as the set of three precedence relationships, for reasons that will become immediately clear

(14a) Copy the sequence from the first consonant to the first vowel
to the segment immediately following the first vowel
<0,2,0> to <0,1,0> to <2,1,0>
(14b) Copy the sequence from the segment immediately preceding the
first vowel to the first vowel to the
segment immediately following the first vowel
<1,1,0> to <0,1,0> to <2,1,0>

No reduplication of the form *tak-takder* can ever arbitrate between these two sets of parameter values. However, Hayes and Abad report an interesting phenomenon: with the language contact from Spanish within the Phillipines, Ilokano, a language that otherwise lacks complex consonant clusters in onsets, borrowed words that contained them, such as *trabaho*. The result was that these words were reduplicated in two different ways, with free variation: *tab-trabaho* and *trab-trabaho*. These two output forms are the results of the two hypotheses in (14). It makes sense to explain the behavior of Ilokano speakers, who variably produce both *tab-trabaho* and *trab-trabaho*, as possessing two different rules for reduplication, and applying one or the other parameter setting freely.

A similar example can be found in English, which has not only the type of reduplication described in a new article by Ghomeshi, Jackendoff, and Russell (e.g. *I don't need someone with a PhD, I need a DOCTOR-doctor*), but also *shm-* reduplication. Mark Southern, the author of a forthcoming book on the historical development of Shm-reduplication, suggests it has been present in Yiddish since 1600, and in mainstream American English since 1929, where historical events led a curmudgeon a short story by H. Ralph Goller to say "crisis, shmisis". A recent web survey of 240 American speakers who identify themselves as conversant in Shm- reduplication reveals that there is significant variation for words whose initial syllable and stressed syllable do not coincide. Thus, 44% of the respondents preferred *obscene-obshmene*, while 46% preferred *obscene-shmobscene*. Importantly, all of them agreed on forms such as table-shmable, which are parametrically ambiguous between stressed-syllable reduplication and initial-syllable reduplication. I consider these findings important because they reveal a "pure experiment" in language acquisition: there is no prescriptivism for *shm-* reduplication, and learners are free to posit whatever hypotheses they want. Clearly, parametric ambiguity in the data and the model developed here go a long way towards explaining these facts.

Briefly turning to use of the parameterized-rules framework for syntax, Charles Yang's recent book has demonstrated that the frequent variation between null-subjects and non-null subjects in child language is best modeled in terms of multiple parameter-settings being maintained simultaneously. This is also a *parallel search* view of syntax, and it stands in sharp contrast to the predominant view in the field, which is that children acquire syntax by maintaining only one value for each parameter at a time, or, more generally, one hypothesis the well-formedness of null-subjects in their language. The more traditional views of parameter-setting, such as that of Gibson & Wexler (1993), would have to posit that children, as they progress from one sentence with a null-subject to the next, with an overt subject, and back to a null subject, are wildly flipping the values of their parameters every time they utter a sentence. Yang's view, similar to the one developed here, as well as by Anttilla, Hayes, and Boersma, is that variation is the consequence of multiple grammatical options, and a production component which chooses among these variably.

For some cases of parametric ambiguity, learners will eventually converge on a single mode of combination. Thus, most English children will converge on the fact that it is a non-null subject language. Yang's hypothesis is that statistical correctness of each grammar will eventually favor one over the other. This is what can be called a *selectionist* view of parameter setting: every grammatical option is considered from the outset, and as each one is proved untenable, it is discarded.

Thus, the Kaingang learner, who finds *va-va* five-ways ambiguous need not worry. With each successive input, less and less of those five modes of combination become true generalizations in the language. Once the Kaingang child hears *jenmi-mi*, she can throw out the hypotheses of iterative reduplication, first syllable reduplication, and total reduplication. Selectionist learning in reduplication, due to the small size of the number of parameters, can proceed in a British Museum search. Learners maintain all possible parameter-settings, with an incremental and error-driven approach to language acquisition: each point in parameter space that turns out to be a dead end is thrown away. Some learners, such as those of Kaingang, will converge very quickly on a unique solution. Other learners, such as those of Ilokano, may never converge upon a unique solution. More precisely, their selectionist process may reach a critical stage and stabilize with two grammatical options, as also seems to be the case for English learners of [l]-allophony who, by and large, encounter and produce *Greeley* long after *light* and *bell*. As Peirce eventually concludes, "Let us suppose that there are thirty-two different possible ways of explaining a set of phenomenon. Then, thirty-one hypotheses must be rejected. The most economical procedure, when it is practicable, will be to find some observed fact which would result from sixteen of the hypotheses and not any of the other sixteen…of course, it will be understood that in the testing procedure itself there need be no such assumption of mysterious guessing powers". (p.250)

The conclusion relevant for our purposes here is that a judicious initial set-up of the correct parameters of variation in reduplication yields an almost trivial acquisition process. At least in reduplication, there need not be any special learning path, or any principled means of ignoring certain inputs until other parameters have stabilized. Numerous computer simulations of the acquisition of reduplication in Kaingang, Ilokano, English, Lushootseed, and even the iterative reduplication of Colombian Jerigonza have all confirmed that the approach developed here yields the attested target grammar for reduplication. UG provides no bias towards "marked" or "unmarked" types of reduplication, and contains no provisions for identifying and ignoring ambiguous input. But it excludes so many irrelevant hypothesis to allow convergence often within six to ten steps. It seems that the secret to successful acquisition is not in a clever learning-algorithm, but rather a careful delimitation of the parameters of variation.

4. Consequences for the Architecture of UG. An exhaustive search like the one pursued is manageable for reduplication, in which each anchor point comes from one of three parameters. However, the question arises, how is this embedded within a larger theory of language learning? I would like to surmise that it isn't, and that a surprisingly large amount of phonological learning is accomplished in isolation from other micro-modules. The most comprehensive computational model of the learning of metrical stress is that of Dresher and Kaye (1990), revised slightly in Dresher (1999). In Dresher's model, metrical stress involves parameter-setting, and deliberately assumes that the learner has already identified units such as syllables, and has a representational distinction of light, heavy and closed syllables. Dresher's algorithm is highly specific to metrical stress, and involves a dependent-order of parameter setting. In that domain, however, the parameters interact in a very different way, as principles of metrical grid construction involve several derivational steps. Moreover, there is far less empirical basis for concluding that the acquisition of stress involves variation and production based on parametric ambiguity, as shown in the collection of papers in Fikkert's book *The Development of Prosodic Systems*. Dresher asserts that a parameter-setting model for syntax may not look anything like the parameter-setting model for metrical stress, and concludes that "there is no parameter-independent learning algorithm" – in other words, the nature of each family of hypothesis determines differently how they should be arbitrated among.

The view developed here is, if not similar to the letter, similar in spirit. Consider the OT learning algorithm developed by Tesar and Smolensky, which, in order to rank n constraints, requires on the order of n^2 tokens of "informative" input, where informative often specifically excludes parametrically (or ranking-order) ambiguous forms. Incidentally, there is no model of a production component in Tesar & Smolensky's model, so at any intermediate stage, there is no concrete prediction of what learners may produce. In contrast, production, implemented as choice among allowable grammars, is available at every stage in the model developed here. The crucial difference is that Tesar & Smolensky's algorithm is trying to rank every phonological constraint in the grammar *at the same time*. With this assumption, the concomitant search size follows. In a British Museum search of the type proposed here, there would be 2^n grammars to select among, with the simplified assumption of n binary parameters. But what I am proposing instead is a micromodularization of learning. The fact is, most parameters within grammars simply do not interact. For example, in fact, many models of syntax-parameter learning (i.e. that adopted by Baker in *Atoms of Language*) will either statistically decrease or change the values of *all* parameters in the current grammar anytime an incorrect parse occurs. That means that parameters such as adjective-ordering within the DP will be altered even when the learner makes a mistake on *wh-* movement. What this suggests is that most parameters do not refer to global properties such as *polysynthetic-or-not* but instead relate to microvariation within modules. If hypothesis-testing is limited to micromodules, with limited interaction, then the search becomes much smaller, as shown in the following idealized equation:

(15) $m*2^{k} < 2^{m*k}$

What I contributed today is showing that many, if not all of the patterns of reduplication in attested languages can be parameterized, that parametric ambiguity is a desirable and realistic consequence of what UG does and does not provide, and that a learning model which accounts for variation is feasible and simulable by computet implementation. I hope to leave you, however, with the broader suggestion that we as linguists rethink our notion of parameters in all domains, and that in doing so, we may simplify the learning problem considerably.