Thank you!

John Goldsmith

September 26, 2019

This is a preliminary version, sent just to the contributors to this book, because I have included a couple of questions to them, and I hereby ask all of you to let me know if I have misunderstood or misrepresented something in what I have written.

Contents

1 On opening the book

Trepidation – that's what I open this book with. And quite a bit. It never occurred to me that there would be a festschrift for me, until there was one. And when Diane and Jackson gave me the book, hardcover and all, all I could see on it was the sticker saying "Open me with trepidation."

I don't even know why–after all, haven't I been saying to everyone who will listen that our discipline is best understood as a great big conversation that we are all part of, and hence we all contribute to? Isn't this wonderful book a testimony to that, and more?

Indeed it is. I tell you what: let's get started; we can come back to this question later. I am delighted to engage in a conversation with all of the contributors to this volume.

2 John Coleman's secret history of prosodic and autosegmental phonology

I'd like to start with a word about John Coleman. When Jason Riggle, Alan Yu, and I invited a whole flock of phonologists to contribute to a handbook of phonological theory (one that would update the one that I had edited in 1995), we explained to them that we were looking for a study of how each subdiscipline had formulated its own deep and interesting questions, and we wanted to learn how well those questions had been answered. This is not the way everyone sets off when writing a chapter for a book such as this; most people think first of their own vision of what direction the field should take in the future, and less about noting and evaluating what had been accomplished. In the end, we did not get what we were hoping for from all of the authors, but when we invited John Coleman to write such a chapter on computation and phonology, he wrote exactly the kind of chapter that we were looking for. I would have been very happy if he had written all the chapters. It might have taken him a few years to do, and it would have been a delight to read.

In John's chapter on the secret history of autosegmental phonology, he does what I can't: compare what I was doing in 1975 to what Trubetzkoy was doing some decades earlier, what J. R. Firth was doing. John begins his story in just the way that I came to phonology, which means listening to what Will Leben and Edwin Williams had to say about treatment of tone in several African languages, in the first couple of years of the 1970s. Will Leben, Edwin Williams, John Goldsmith: what a small stage. Three students spending a few wonderful, intense years at MIT, studying phonology with Morris Halle. I do remember asking Morris what would be a good topic for a phonology paper (I had no idea, myself, knowing next to nothing about phonology). Morris handed me a copy of Will's dissertation, suggesting that I read it and find a question to pursue. Will had left MIT just before I got there, so I had not met him. But his dissertation intrigued me. He demonstrated that a vowel could be realized with a sequence of tones (Low plus High, or High plus Low), and there was no way to represent that in the formalism that Morris and Noam had developed just a few years before, in *The Sound Pattern of English*. *SPE* was published in 1968; this was five years later, in 1973. As an aside: I don't think any of my classmates thought that we were supposed to be studying SPE. I know that I bought a copy of it at the MIT Coop bookstore, but I don't think we were expected to read it.

John Coleman begins his chapter by quite rightly pointing out that the official statement of generative phonology, SPE, acknowledged that its authors did not offer a framework with a clear way to deal with issues of prosody. That's not *quite* right; Chomsky and Halle thought that they had exactly the right way to talk about stress in English. But they did not presume that the answers had been given to how their analysis could be extended to problems of prosody more generally.

Morris knew that this was an important problem. I can only imagine he had encouraged other students inbetween 1968 and 1973 to work on tone languages, students such as Nancy Woo, Will Leben, and Edwin Williams. I have a vague recollection that Morris told me that whoever worked out this problem would win the Big Prize (the Oscar, or whatever was the equivalent in linguistics).

John Coleman wrote, "as Goldsmith knows and has written about himself, he built upon a wide range of earlier work on prosodic features, such as Zellig Harris's *long components* (Harris 1944) and J. R. Firth's *Prosodic Analysis* (Palmer 1970)." But there is more to the story than this (even though what John wrote is quite true). When I was working on my dissertation, I did not have the slightest understanding of how knowledge accumulates in a discipline. Sure, I had read Kuhn's book on scientific revolutions; I had read it in 1969, and soaked in all of the glory it shone down on revolutionary scientitsts who strike out on their own.[1] I'm sure my modesty would never have allowed me to think that I was a revolutionary in Kuhn's sense. I was a member of a small group of people working to move forward the principles of generative phonology.

The first intimation that I had that there were linguists who outside of that small group at MIT (other than those who were hoping to get inside) was when Morris told me that I had to write something about how autosegmental phonology related to the work of other linguists in the 1940s, such as Bernard Bloch and Zellig Harris. If I didn't address that question, Morris explained to me, others would say that what I had written was warmed over Bloch and reheated Zellig Harris.

This came totally from out in left field, as far as I was concerned. I had *no idea* there were other people on our island (if I may switch metaphors), and that they had been there since before I was born. So I went home and read Harris's paper on simultaneous components in phonology, and a few other papers from the period. Yup, Morris was right. I had to explain how my work related to Harris's and Bloch's, because they certainly were trying to address similar questions (even as Stephen Anderson was doing at the very same moment, I vaguely noticed). So I wrote an introductory chapter in my dissertation in which I tried to express similarities between that work in the 1940s and what I had done.

It took me another ten years to begin to understand how knowledge grows, and the relationship between linguists and the papers they publish. By then I had worked on the tonal analysis of several Bantu languages based not on published grammars but on working with speakers, and I was beginning to appreciate the complex path from observation to description. What I began to understand was the enormity of the intellectual and conceptual work that Green and Igwe (for example) had put into their grammar of Igbo, the West African language; it was their grammar that had given me confidence in the power of autosegmental analysis. Their grammar was their work.

I tried to express this realization of the obvious in a paper that I published in the *Journal of Linguistics* in 1991. It was intended as something close to an apology aimed at those who considered themselves followers of prosodic analysis in England, or if not quite an apology, a expression of considerable gratitude on my part towards the published work of the students and grandstudents of J. R. Firth.

All of which is to say that while what John said about me is true, none of this was obvious to me at the time when I wrote my dissertation (even if I did my best to not let my ignorance show).

Perhaps there will never be a better moment to pursue a bit more the value of the advice Morris Halle gave me when he directed me to read Bloch and Harris in 1976. Bloch's work and Harris's work was almost thirty years old at that point, so it was really ancient history. (Most people whose dissertations were thirty years old were dead, I imagine. This gives us pause.) I don't think it ever occurred to me that I should have known that work, or that Morris should have let me know about it earlier. While working on my thesis, I would meet with Noam, though not as often as I met with Morris, and Noam never mentioned Zellig's work on simultaneous components. I do recall that when I defended my dissertation — an open event at MIT, with no cover charge and all invited — I did talk briefly about Zellig Harris and what he had tried to do, with my eyes on Noam to check whether he agreed with me. Many years later, all of this came back to me when I heard Noam expressing regret and unhappiness that Zellig had never told him to read Leonard Bloomfield's last paper, Menomini morphophonemics. I have a sense that I may have hurt someone's feelings when I observed that Noam didn't point out Zellig's work to me either; I don't know.

John Coleman's paper gives enough to whet the intellectual appetite of a phonologist, with the promise of seeing careful intonational transcriptions of speakers from several centuries ago, such as those given in Steele 1775. He also draws the reader's attention to the connections between where linguists were employed during the Firthian years in England and the particular problems that they were drawn to. There is much to think about when one looks at the professional and intellectual connections. John emphasizes two salient features of prosodic analysis: that any phonological aspect of an utterance might be the target of a prosodic analysis, and that prosodic contrasts in one position of the word (or the grammar) need not be made commensurable with prosodic contrasts elsewhere in the language: word-medial consonantal contrasts need not be aligned with word-final contrasts, perhaps.

I think that there is another, equally great difference that separates prosodic analysis from the families of phonology that emerged out of contrast-based theories of phonology (and by that phrase, I mean to include Saussure, Trubetzkoy, Sapir, and Bloomfield): prosodists are interested in generalizations even if they are

[1] I even remember reading Gerald Holton on the very day that I failed in my defense of my first paper on autosegmental tonology.

only marginally related to true *contrasts* in a language. Bantu tonologists know that they share that feeling with Firth, and I suppose any autosegmental phonologist will share the feeling too. Autosegmental analysis requires the analyst to figure out how the phonological pieces fit together, and that fitting together often amounts to adding and deleting association lines, though there can be other sorts of fitting as well. Sometimes it is all "mechanical" and not based on expressing or maintaining contrasts, but it is still of great interest to the phonologist. Pre-generative American phonology really did prioritize phonological contrast, and by prioritizing it, made other aspects of the phonetic symbol less interesting. When we read Sapir today, we can see something akin to glee when he points out that speakers of one language really don't care about some aspects of the sounds of English, just as speakers of English don't care about some aspects of the other. We could call that a kind of phonological relativism, and I think in pre-war America that would have been just fine. It was just this relativism that Trubetzkoy and Jakobson hoped that they could overcome, though.

As I say, John notes that Firth was more than willing to allow a contrast in one part of a word to be really different from a contrast in a different part, and he draws a parallel to Saussure's willingness to focus on a difference (but not a contrast!) between syllable-initial and syllable-final /p/'s (and all the other consonants which appear in both positions). Bernard Laks has long defended the view that Saussure's view, that elements in syllable-initial position were really different from elements in syllable-final position, can naturally be viewed in the context of dynamic computational nets, where an abstract notion of sonority, which can be measured and shared. That will probably come back below.

John has managed to pull together published and not-so-published material to allow us to much better understand where Trubetzkoy's thinking was in the few short years before his death. Trubetzkoy certainly could not abide fools, but he did not feel that he had to stop there. In a letter that Trubetzkoy sent to Jakobson in 1934, he expresses the impression that there were "no real linguists" in English. But Firth he could learn from, Coleman notes, and his story is quite interesting.

I suggested just above that the pre-generative American phonologists were relativists in a certain sense, and John suggests that in much the same way, Firth was one too. Firth and his students applied their method across a range of quite different languages, but there was no need to see the results of those analyses fill out some analytic table of possible differences; he was not, as John notes, a Jakobson, nor a Greenberg, nor a Chomsky. Or so it seemed, until John discovered a page of notes that Firth had sketched for a paper published in 1948, though the notes did not make it to the final version. In these notes, Firth looks like he is moving on to a new stage, in which he looks at the totality of the analyses that he and his students, and others, had discovered.

3 Will Leben

Will begins his chapter by reminding us how phonology looked to linguists in the wake of the publication of *Sound Pattern of English*. So many phenomena could be treated in new ways, but tonal phenomena remained recalcitrant. Will brings up the treatment of tone in Tonga, a Bantu language spoken in Zambia. It's a language which had been studied by a number of linguists, including Hazel Carter and A.E. Meeussen, and closer to home (where "home" meant "MIT," of course) by Jim McCawley and Michael Cohen. It is a language which bears great similarity to Luganda, an even better studied language, but the similarities are hidden; they emerge only once one has been able to undo some changes that were (as they entered the language) sound changes, in the traditional sense. Working on the language using the materials published by Hazel Carter, and also using additional materials provided by Jerome Hachipola, who was a grad student at Indiana University and a speaker of Tonga, was for me a great experience, and I think that experience made a phonologist out of me in ways that no other experience has.

Carter and Meeussen made sense out of the Tonga data by noting that from a tonal point of view, morphemes in Tonga are typically divided into two types, but neither type is naturally characterized as "low tone" or "high tone." One of the groups gives the impression of being *active*, contributing to the tone pattern of the language, while the other gives the impression of being *passive*, accepting a tone assigned to it by general principles. The active group generally corresponds to morphemes with a High tone in reconstructed Proto-Bantu, while the second generally corresponds to Proto-Bantu Low tone. Both Cohen and I analyzed the system as reflecting the presence of, in each word, one or more packets of High plus Low tones (written symbolically as $H\ \overset{*}{L}$, where the asterisk indicates that the Low tone is to be associated with the accented vowel position on the opposite tier).

There are a couple of things about that analysis that seem odd, and bear discussion right from the start. Why should the term *accent* enter into the discussion, if we are looking at what seems on the face of it to be a tone language? If there is a prosodic contrast between two sets of morphemes (the contrast that is the descendant of proto-Bantu High and Low), just what is that contrast? Is it *accented vs. unaccented*, or *HL versus* $\emptyset$, or something else, like *H vs. L*?

I thought the right answer was *accented vs. unaccented*, and used the asterisk as a symbol for accent on the grounds that this formalism was just like what I had suggested be used for the treatment of English intonation (Goldsmith 1980[1974]), and I'll return to that just below. Will pointed out that Doug Pulleyblank proposed a different analysis, based in part on implications of Paul Kiparsky's influential model of lexical phonology; Doug's account also included a recognition of greater variance across languages of how the spreading effects of the Well-formedness Condition performs. Some effects of the WFC seem well-motivated in some languages, and yet those same effects arguably are not observed at all in other languages. Thus there must be some way for the grammar (the grammar is learned, let's not forget) of a tone language to employ the WFC in language-particular fashions.

Lexical phonology, Doug noted, could be interpreted as ruling in favor of High as the sole feature value (of a feature of tone, where lexical phonology could be interpreted as allowing only one specification – only plus, or only minus – of each feature, probably in a contexts dependent way; the opposite value would be called the "default" value) that is active in the lexical phonology. If that is the case, Tonga has to be rethought, and on this rethinking, Tonga would retain underlying the same tone assignment that had been found in Proto-Bantu. The vowels that we called accented are, rather, High toned vowels, and not Low toned vowels, and these High tones may spread over an unbounded number of vowels to the left; in addition, that High tone will disassociate from the vowel it is associated with underlyingly.

This investigation, or exploration, spoke to whether universal linguistic theory should decide what phonological contrasts looked like at an underlying level. In more recent years, in work that I've done with Bernard Laks (the greatest part of which is in *Battle in the Mind Fields*), I've come to see the respects in which Trubetzkoy and Jakobson did not agree with each other about precisely this question. Jakobson found very appealing the view that all features have the same kind of logic, while Trubetzkoy thought that studying a language allowed one to see the feature logic each language devised for itself.[2]

Laks and I emphasize in our book the degree to which the life of a discipline is best understood as cluster of conversations bringing together both people who agree with each other and people who very much disagree. Regardless of whether someone agrees or disagrees, the conversation cannot continue if we each fail to understand what the other is saying, and we fail most spectacularly if we fail to understand important advances that others have made. My own experiences in phonology helped me to understand this, even if it sometimes took me a decade or two to understand. I met Will at the Conference on African Linguistics in 1975, and through him got to meet Larry Hyman and Jan Voorhoeve, whose early interest in autosegmental phonology was also critical for the direction my work took. As we come to better understand the growth of knowledge, we appreciate better how these personal contacts and conversations form an integral part of the evolution of disciplinary knowledge.

In his chapter, Will reconsiders arguments for and against invoking the notion of accent in analyzing languages whose prosody seems at first and even second glance to be traditional tone languages (even if we recognize that that phrase requires some unpacking: if something is traditionally viewed as X, does that make any claims above and beyond the observation that others in the past have had a certain belief?). This question can be seen as closely tied to how we understanding tone languages in which the underlying tonal contrast is the contrast between possessing a High tone and possessing no tone at all. Some (but not this writer) would suggest that systems with accent ought to reflect a strong preference for each word to have no more than a single accent, what Trubetzkoy called *culminativity.* I believe that the right way to identify an accentual pattern is to find a prosodic generalization that identifies a location in a word, and uses that location to associate the tones of certain (tonal) morphemes, when that tone has to be marked or indicated explicitly in the language. While that sounds rather abstract, the intent is quite simple: English has a range of intonational melodies, and in some we find that a High tone is associated with the accented syllable of its phrase, while in others, it is the *Low* tone that is associated with the accented syllable. That difference

[2] *Battle in the Mind Fields*, Chapter 9.

is reflected in the tonal melody by noting which tonal segment is the associater, and by parity of reasoning, we use the same mechanism to mark accented moras, that is, we mark accented moras with an asterisk. As Will notes, I tried to make that argument clearer in the treatment of Kintandu. (Still, my confidence in any of these solutions is weakened by the fact that a consensus has not been reached on the treatment of accent despite a lot of time and effort that has been put into the question. It may be that we will not understand the prosody of accent until we better understand rhythm in languages.)

So just to mark the temporary end of the discussion of tone and accent, I believe that we find languages that have tonal morphemes consisting of more than one tone, and that in some of these languages, one of the tones is marked as the accented tone. In systems with accented tones, there are necessarily accented syllables (or moras), and association principles associate corresponding accented vowels and tones. This perspective strongly suggests that the analysis of Tonga is one in which all (non-default) tones arise through the association of a $H \overset{*}{L}$ melody.

Will draws our attention to work that shows that not all structure in tone languages is easily reduced to cross-tier association; sometimes tonology appears to either give rise to, or else to be sensitive to, a "domain" of a sort, a domain which is not morphological in origin but which nonetheless allow certain generalizations to be naturally expressed, as has been argued in the context of "optimal domains theory." He also emphasizes that there can be a healthy trade-off between the economical accounts offered by optimality theory and those provided by using autosegmental represenations.

My own view of how to think about autosegmental phonology is that it reflects (as if in fossilized form) part of the way in which the structure of language is discoverd by each individual. As humans, we come to the world overwhelmed by the range of sensations that beset us and actions that we can take in every single muscle of our body. We learn to organize the sensible world, and we learn to fit muscular actions and gestures together into higher level executions. Imagine a baby learning to clap, for example; it will be months before she can succeed in clapping her hands, time that she spends learning to coordinate the muscles in her shoulders, her elbows, and her hands. Once she has that down, she can learn to beat them in a sequence, and then evenually she can beat them to the rhythm of a song or a dance. But that's only just the beginning, when we learn to clap our hands to the beat. The next step is to learn to clap your hands on the off-beat (the back beat, rockers say). When you clap your hands in a syncopated fashion, you are not ignoring the beat; you are deviating from it in order to gain something from this more complex coordination between what is thought (or felt, if you will) and what is done (or heard). To put it more succinctly, we learn the sequentiality of phonology by virtue of creating a simplified model of phonetic reality, and in that simplified reality each segment is preceded and followed by exactly one other segment. But some languages allow an autonomous status for certain phonetic components, then ones that we call autosegmentalized. In like fashion, we learn the dynamic and chronological patterns of language by recognizing multiple patterns of strong and weak beats, but we allow ourselves to speak or sing in syncopated ways, recognizing the shift that such syncopation consists of between two patterns, one internal and one external.

4 Larry Hyman on Lusoga

I *imagine* that there are more Bantu languages that Larry Hyman has not analyzed than there are that he has, but I'm not at all sure about that, and in the end it is only a matter of time. There are languages that he has come back to time and again, but still the list lengthens of Bantu tone languages which he has brought to the attention of phonologists the world around. Larry's work (like so many of the contributors to this volume, and many others in African linguistics) always shows complex relationship between analysis and theory. He knows the questions to ask, and then he both uses the theory to provide an analysis and he uses the analysis to shape the theory. His work stands as a healthy antidote to the empiricist belief that there is a straight-forward division that can be made between description and analysis.

In his chapter, Larry addresses the question of dealing with a Bantu language, Lusoga, that resembles Luganda (a language that Larry has analyzed on several occasions) but in which etymological High tones are realized on a Low. He begins with the tonal pattern of the infinitive:

(1)

Proto-Bantu	augment	prefix	stem	FV
Low	ò	kú	H^n	á
High	ò	kú	L ((L) H^n)	á

(2)

LH (1):

Proto-Bantu	augment	prefix	stem	FV	Proto-Bantu	augment	prefix	stem	FV
Low	ò	kú		á	High	ò	kú		à
Low	ò	kú	H	á	High	ò	kú	L	á
Low	ò	kú	H H	á	High	ò	kú	L L	á
Low	ò	kú	H H H	á	High	ò	kú	L L H	á
Low	ò	kú	H H H H	á	High	ò	kú	L L H H	á
Low	ò	kú	H H H H H	á	High	ò	kú	L L H H H	á

Many Bantu languages in this area have maintained a vowel length contrast, and the tone pattern with a long vowel in the first root syllable is the following:

(3)

LH's (2):

Proto-Bantu	augment	prefix	stem	FV	Proto-Bantu	augment	prefix	stem	FV
Low	ò	kú	HH	á	High	ò	kú	LL	á
Low	ò	kú	HH H	á	High	ò	kú	LL H	á
Low	ò	kú	HH H H	á	High	ò	kú	LL H H	á

There exist roots that have a short vowel in the first syllable and a long vowel in the second syllable:

(4)

LH's (3):

Proto-Bantu	augment	prefix	stem	FV	Proto-Bantu	augment	prefix	stem	FV
Low	ò	kú	HH	á	High	ò	kú	L LL	á
Low	ò	kú	HH H	á	High	ò	kú	L LL H	á
Low	ò	kú	HH H H	á	High	ò	kú	L LL H H	á

Larry notes that in the last two cases on the right, we would expect *ò kú L LH H á*, and that as in Luganda and other languages of this area, expected rising long vowels are realized as long Low, and there is no way to generate a long rising (Low High) vowel.

When the root is vowel-initial, then the prefix *ku* merges with that initial vowel, and the prefix's High tone appears on the first mora of the long vowel. The second mora of that long vowel is then Low. If there is one more mora after that long vowel before we get to the FV, then it is Low, but any further vowels are High. What a beautiful system!

(5)

LH(4):

Proto-Bantu	augment	prefix	stem	FV	Proto-Bantu	augment	prefix	stem	FV
Low	ò	kw	HH	á	High	ò	kw	HL	á
Low	ò	kw	HH H	á	High	ò	kw	HL L	á
Low	ò	kw	HH H H	á	High	ò	kw	HL L H	á

Larry suggests a vision of this data in which *all* of these words have a phrase-final High tone, which is realized as such in the left column of the tables of data just above (i.e., the proto-Bantu Low column) over all of the syllables past the first, and on the right over a reduced set of syllables (these reduced syllables are competing with the Two Low Tone constraint), and *all* of the words have a phrase-initial Low tone. To repeat: all of the forms described here show the effect of a prefix Low and a suffix High, both of which are part of the phrasal analysis and not strictly part of the phonological information contributed by the word as such.

Second, in the right column (i.e., the proto-Bantu High case), that proto-Bantu High is associated with the infinitival prefix *ku*. In *every* context, the radical High is followed by a Low toned syllable, and in most of the cases it is followed by two Low toned syllables. What determines how many moras or syllables will associate with this Low tone (whose origin we have not yet seen)? And about that – where and when did this Low tone come from?

Larry points to his analysis of Luganda to justify the insertion of a Low tone in the cases on the right, based on the proto-Bantu High radicals.[3] One of the reasons why this is of consequence is that we might be tempted to think that this language contains a violable constraint (which Larry calls *2LTR*, or Two Low Tone Requirement) whose effect is seen in all of the high toned verbs as long as there are enough moras (since it is really a requirement regarding moras, not vowels). The constraint, if it exists, is violable, in that there are well-formed words which are too short to permit two moras to associate with the Low, and yet the word is grammatical.[4]

Larry suggests a scenario which consists of a sequence of sound changes, and something like this scenario will also emerge in rule ordering in the synchronic account of this language at present. The first change is one which lowers the High tone of the augment (we have not yet seen the reason to think there was a High tone on that morpheme, but we will return to that), and the second is one that assigns a Low autosegment to the mora after the (sole) High tone. That High tone may be the verb root's lexical tone, or it may be the tone of the augment, as we will see just below. The rule will not apply if that High is (already) followed by a Low tone, Larry suggests, because

The third and most important part of Larry's tone-shift scenario is an account of the reassociation of a High tone to the vowel preceding vowel, which he calls *High Tone Retraction rule*. The scenario is one in which a sequence of sound changes have occurred (and therefore, it appears on this account, present-day rule ordering reflects the chronological sequence of the sound changes), and the first one is the rule that deletes the High tone of the augment. We actually have not seen a High tone on the augment yet in *any* forms, but we will shortly.

1. Delete High tone from word-initial augment.

2. (Meeussen's Rule of $H\ H \rightarrow H\ L$ will be here shortly.)

3. *Low Tone Insertion*: Insert a Low tone immediately after a High tone, and associate it with the mora immediate after the High tone; but only do that to the rightmost High tone of a word (for that last condition, see discussion of Object Markers below).

4. *H Tone Retraction* (p. 50) Retract (leftward) association of a High tone, associating the now-toneless vowel (produced by the preceding rule) to the Low on the right.

5. Associate phrasal tones on the left and right to all accessible toneless vowels.

There are, however, two ways to formalize Larry's account of the tone shift, and Larry appears to prefer hypothesis 2: he refers to it as the $H > L$ *change*. I have used the color *blue* to indicate the root in the following words, and also to mark the underlying High tone of the root. The black Low tone that follows

[3] For any non-Bantuist who is reading this and feeling impatient with people talking about proto-Bantu High when they should be ignoring historical linguistics and doing real phonology, the reason Bantuists do this is the same reason that dialectologists talk about the vowel in HAT—it is an easy way for everyone to understand what we are talking about at a particular moment.

One of the reasons that this question is of consequence for phonological theory is that some theories (like lexical phonology, as we briefly noted above) make predictions about which parts of a phonology may employ only one value of a feature, and which parts two values of a feature. If the High tone and the Low tone are autosegments consisting of a single feature which can take on two different values, then we have a clear case to examine.

[4] This is seen notably in High toned verbs such as *-bon-*, which surface as *ò-kú-bòn-á*.

each of the root's High tone is inserted by rule, according to Larry's proposal, which he suggests is just like what is found in Luganda. Please note my notational convention, whereby the accents placed above the vowels indicate the surface tone, which is generally not the same as the authosegmental representations being discussed. I find that this is the most convenient way to indicate material, and it highlights the changes that a form undergoes as well. In the case of many Eastern Bantu languages it suffices to mark High tone, and leave Low tone unmarked.

(6)

input:

ò-**kú**-**bòn**-á	ò-**kú**-**ghùlìr**-á	o-**kú**-**kàlàkát**-á	ò-**kú**-**lím**-á
H L	**H** L	**H** L	

hypothesis 1:

ò-**kú**-**bòn**-á	ò-**kú**-**ghùlìr**-à	ò-**kú**-**kàlàkát**-á	ò-**kú**-**lím**-á
H L	**H** L	**H** L	

hypothesis 2:

ò-**kú**-**bon**-á	ò-**kú**-**ghulir**-á	ò-**kú**-**kalakát**-á	ò-**kú**-**lím**-á
H L L	H L L	H L L	

The right-most example (ò-**kú**-límá) has a High tone associated with the augment at the deepest level, on Larry's account, but it is deleted before the process described in the chart immediately above.

I am interested in the difference between the two manners of description that we see in these two hypotheses, 1 and 2. Granted, the difference seems so slight that perhaps only a copy editor would care—but I care, and for the reason that autosegmental theory commits itself to the idea that getting the representations *right* will have important (rather than peripheral) consequences. From an autosegmental point of view, hypothesis 1 is possible, and hypothesis 2 is not (hypothesis 2 is not possible, in the presence of a great deal simpler rule, the one that is at work according to hypothesis 1). However, we should look at Larry's analysis of the augment's tone before proceeding further.

Larry points out that we can consider nouns preceded by the preposition *na*, both with and without a following argument. Here are the toneless and High toned infinitives following the preposition, nearly as Larry presents them (as above, I have written the correct *surface* tone with acute accent (High) and grave accent (Low) over the vowels). Larry notes that the expected falling tone on the *na-o* in *ná-ó-kú-bòn-á* is eliminated by a larger generalization, that word-internal sequences of H-L-H are simplified to H-H-H, though the phrasal %H does not count as a word-internal segment, and does not bear on this generalization (HLH is OK if the second H is a phrasal High). I have ignored the syllable-internal change of *a-o* to *a-a* in the data.

(7)

Toneless or Low tone verb

Before tonal rules		After tonal rules
ná-ò-**kù**-**ím**-á	→	ná-ò-**kù**-**líma**-á
H		**%L** H L **%H**

High tone verb				
Before tonal rules		Intermediate		After tonal rules
ná-ó-**kú**-**bòn**-á	→	ná-ó-**kú**-**bòn**-á	→	ná-ó-**kú**-**bòn**-á
H **H**		HL H**L**		**%L** H H**L** **%H**

(8)

Here are the same forms, but without the augment:

Toneless or Low tone verb		
Before tonal rules		After tonal rules
nà-**kù**-**lím**-á	→	nà-kù-**lím**-á
		%L **%H**

(9)

High tone verb				
Before tonal rules		Intermediate		After tonal rules
nà-kú-**bòn**-á	→	nà-kú-**bòn**-á	→	nà-kú-**bòn**-á
H		H **L**		**%L** H **L** **%H**

(10)

Larry suggests (p. 52) that the augment does not bear a High tone more generally, and this is shown by looking at verb-subject sequences where the subject noun which *does* begin with an augment, does *not* assign a High tone to the previous syllable, while a subject-verb sequence in which the verb begins with an underlying High tone (such as the class 2, also known as human 3rd person plural, prefix *ba*) will in fact place a High tone on the final syllable of the preceding noun.[5]

All object markers (OMs) behave the same way from a tonal point of view, and Larry argues that they behave like morphemes descending from an earlier High tone. When followed by a toneless verb radical, the word is realized with a High on the syllable preceding the OM, and when the radical is High toned, it behaves the same; there appears to be a tonal neutralization of the root's underlying tone when it is preceded by an object marker. The effect by which a tonal neutralization occurs after a High toned Object Marker is referred to as *Meeussen's Rule*. Larry assumes that the formal effect is *not* quite tonal neutralization; he takes it that Meeussen's Rule changes a High to a Low tone. Here is what we see, exemplified by the 1st person plural prefix *-tu-*, on p. 53 of Larry's chapter.

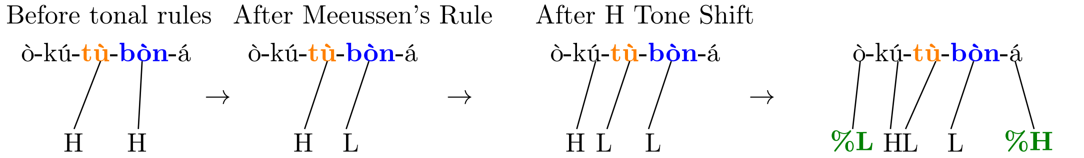
Toneless or Low tone verb		
Before tonal rules		After tonal rules
ò-kú-**tù**-**bàl**-á	→	ò-kú-**tù**-**bál**-á
H		**%L** H L **%H**

(11)

[5] *o-kú-bon-a a-bá-kal-í* "to see women" versus *a-bá-kal-í ba-sek-á* "the women laugh".

High tone verb

Larry's account, in his (10)



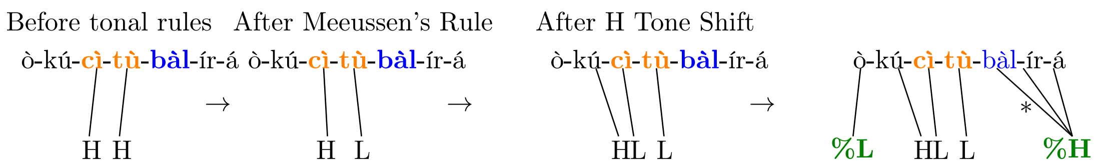
(12)

Why, in the case of the toneless verb, is the Low tone on *-tu-* only associated with *-tu-* and not also with *bal*? We will have to come back to that. It is worth noting that the autosegmental geometry is being fully exploited, in the sense that all three kinds of tonal association are playing separate roles. Vowel positions may be unassociated, or associated with an H tone, or associated with an L tone. Larry takes Meeussen's Rule to change an L to an H, though the theoretical possibility remains that the rule might instead delete the tone, leaving the vowel position toneless.

When two Object Markers (OMs) appear, we again find tonal neutralization of the two tonal classes of roots, but the analysis till now does not predict that. The tone of the root is Low, whether it is underlyingly Low or High, while the analysis predicts H. I have marked the wrong prediction association line with "*". The surface Low on *-bal-* indicates that it is associated with a Low, and for this reason cannot associate with the phrasal High.

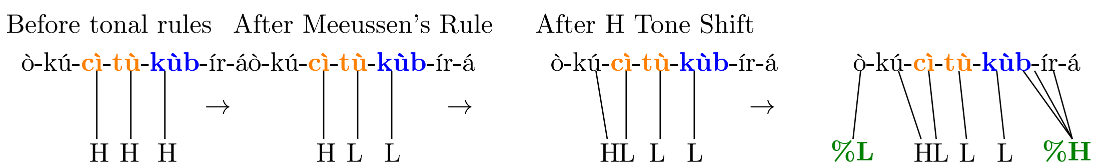
Low tone verb, 2 OMs

Larry's account, in his (12a)



(13)

High tone verb, 2 OMs



(14)

Larry suggests that the correct analysis derive from a cyclic analysis in which a domain consisting of the rightmost Object Marker and the stem, the rule of Low Tone Insertion applies, giving the root *bal* a High tone early in the derivation.

This is a move we would only want to make cautiously. We have two processes that in a derivational fashion ensure that H tones and L tones should be paired. One case is *Low Tone Insertion*, which only applies to a right-most H, and pairs it with an inserted L tone (and it will not apply if there is already a Low there), while the other case is *H Tone Retraction*, which shifts the association of a High, and derivationally supplies it with a Low tone just to its right. So derivationally, the tonal tier is readjusting itself to be sequences of H's and L's, and the phonological role played by those L's is to block the association of a following floating H from associating with them.

Larry draws then the natural conclusion that some, and perhaps all, of the H-L sequences are not produced derivationally, but instead, the H L sequence is the synchronic descendant of an earlier High tone, and furthermore, it is the Low tone of this sequence that is tonally associated with the vowel of the morpheme that it is part of. We might imagine marking that L tone with an arbitrary mark (such as an asterisk *) to indicate this (if we do not, then we might be forced to engage in some odd formalism in which a tone has half an association line attached to it).

The material that Larry discusses works well with an $H\ \overset{*}{L}$, as I suggested for Tonga, though Larry and I have never seemed to agree about whether it is a reasonable thing to call the asterisk here a mark of "accent" (see his p. 67).

Larry suggests (p. 59) that Meeussen's Rule must be made slightly more complex under the $H\ \overset{*}{L}$ hypothesis, though I'm not quite sure I see that. He notes that if we take $H\ \overset{*}{L}$ to be the underlying tone in a sequence of morphemes (such as object markers) we will have unassociated High tones that seem to be stuck with no vowel to associate with, and he seems to think either that this could not be the surface representation of a sequence of Low-toned morphemes (because the High has to heard?) or else a correct analysis requires that such High tones be deleted. I don't think that either of these implicit assumptions are correct, however.

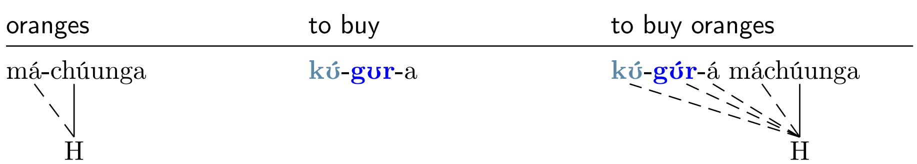
Larry draws his discussion to close with a mention of a trend that he sees in phonology, "a definite trend of disinterest in, if not active opposition to abstract representations in phonology," (65) and suggesting that examples like this one need to taken seriously by linguists who welcome and linguists who challenge this trend. He is certainly right. At the same time, we should not lose sight of the fact that autosegmental representations are by construction less abstract than most strictly segmental accounts; an autosegmental analysis incorporates the claim that this is what the surface looks like (i.e., consider any of the output structures in this paper).

5 David Odden on Logoori

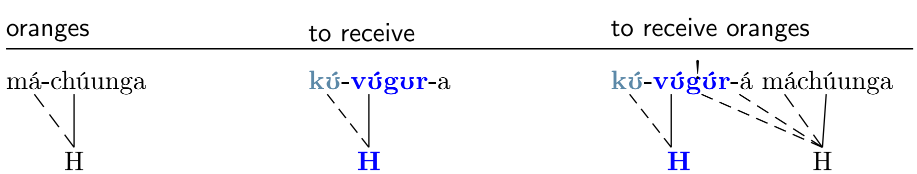
David Odden presents an analysis of Logoori every bit as beautiful as Larry's of Lusoga. They both involve a rule of leftward High-tone spreading from objects onto verbs, and both language do not allow such a spreading High to associate with a vowel that was underlyingly High, even if that High is removed by the effects of Meeussen's Rule (M's Rule is the rule that deletes High after High, or neutralizes the H/L contrast immediately after H). Does this make the system more abstract, or less abstract? The answer, to be sure, depends on how we define abstractness. But a good case could be made that this characteristic makes the language more learnable: even if the High tone of verb stem is not realized, its presence is indicated on the surface by the blocking of a tone coming from the right.

The Logoori tone system is familiar in that the various tenses (using the term broadly) display different tonal patterns that divide into roughly four groups: the "basic" pattern, which displays the underlying tone of the verbal root; the Melodic Pattern 2, which adds a High tone on the second mora of the verb stem when the verbal root is toneless, and on the final vowel when the verbal root is High-toned; a Melodic Pattern 3, which adds a High tone on the second mora of the verb stem; and a few other patterns, including the subjunctive. All this is quite familiar to Bantuists. But this is just the beginning.

There is a general rule of leftward tone spreading, operating at a phrasal level:[6]



(15)



(16)

[6]This is an example to come back to in the context of Mark Liberman's ideas about post-lexical rules being matters of articulation.

There is a vowel length contrast, and only long vowels can display a Falling tone, but only on a penultimate syllable; Rising tones are found nowhere. There are cases where a long vowel in prepenultimate position is assigned a High tone on the first mora by the rules, but when that happens, the prepenultimate vowel is High rather than Falling in tone.[7] There is an issue that arises with respect to the tone of a bimoraic (i.e., long) penultimate syllable, and it is an issue that has arisen in the analysis of a number of languages relatively close to Logoori. That question is: when a High tone is assigned to the first mora of a long penultimate, when is this realized as a long High, and when is it realized as a Falling tone? This question will be asked in connection with each tense, and a final understanding of this remains illusive. For example, in infinitives, a High toned radical whose first syllable is long is realized with a Falling tone when it appears in penultimate position, as we see here:

(17)

ku-**híiz**-a — H

ku-**záázam**-a — H

ku-**sááŋgaar**-a — H

Meeussen's Rule deletes a High tone on the verb root when a High-toned Object Marker (OM) precedes immediately (while two High-toned OMs may be found in sequence with neither tone lowered). Notably, though, when a High tone spreads leftward into a verb from a following word, the High tone seems to be blocked by the vowel which undergoes Meeussen's Rule, suggesting that Meeussen's Rule changes a High to a Low; the structure that results is one in which the verb radical is associated with a Low tone, and thus not available for unbounded spreading.

(18)

to count it for them

kʊ́-**gí**-**vá**-**variz**-a — H H

to catch it for them

kʊ́-**gí**-**vá**-**ɲagull**-a — H H L ← lowered by Meeussen's Rule

David suggests that the absence of downstep between the two Hs of the Object Markers shows that the two OM Highs have merged, giving this structure:

(19)

to count it for them

kʊ́-**gí**-**vá**-**variz**-a — H

Quite strikingly, the phrasal leftward spread of High tone does not spread to the position where the High tone of a High radical root was located, as in (b):

(20)

to read it slowly

kʊ́-**ké**-!**sóóm**-á gáráha — H H

to shave them slowly

kʊ́-**vá**-**veg**-á gáráha — H L H

Subject markers (SMs) are associated with a High tone underlyingly which does spread to the left, but which in the examples David gives surfaces as Low:

(21)

they will catch for each other

na-**rá**-ká-**nágɪllan**-e — L H

[7] David, is there a typo on page 73, just below (4), ake[sooma]? It is written with an underline as if it were High, but I think it is toneless.

Check on these forms:[8] Check on these forms:[9]

A number of the tenses display a pattern that Dave calls *Melodic Pattern 2*,[10] which in general places a High tone (a *melodic* tone, in Odden's terminology) on the second stem mora when the radical is Low toned (or has no tone). When the radical is High-toned, that H is lowered to L, and a High appears on the Final Vowel. We see the SM surfacing with its High tone here. Note the realization of this High in *maní-vá!-kʊ́ʊ́t-a*, where the melodic H assocites with the second mora of the bimoraic radical, and then spreads leftward.

(22)

Toneless radical		High-toned radical	
maní **vá**!-**móróm**-a	they spoke	maní **vá** **vitán**-á	they passed e.o.
maní **vá**!-**vʊ́rʊ́ganyiran**-a	they stirred for e.o.	maní **vá** **gɪrʊ́ʊ́ngány**-á	they turned
maní **vá**!-**kʊ́ʊ́t**-a	they scraped	maní **vá** **deek**-á	they cooked

they stirred for e.o.	they turned
maní **vá**!-**vʊ́rʊ́ganyiran**-a	maní **vá** **gɪrʊ́ʊ́ngány**-á
H **H**	**H** **L** **H**

(23)

Dave notes that two tenses that fall into this pattern, the indefinite future and the persistive, the melodic High does not appear on the surface, and yet there is good evidence that these tenses belong to the same tonal pattern. The evidence appears when such verbs are followed by an adverb with a High tone that spreads leftward into the verb. In such cases, the leftward spreading is blocked by the Low tone of an underlyingly High radical (though not that of a toneless radical), and the High on an Object Marker is also Low (Low, not toneless, because the leftward spreading from the next word goes as far as the toneless radical, but not further into the OM), strongly suggesting that the melodic H is present to trigger the effects on the preceding radical tone and the preceding Object Marker, but no longer present at the phrasal level, when a High tone spreads leftward onto the verb from the following adverb (I use the square brackets to indicate a tone that is deleted, in this case as a function of this tense):

they will shave (H-tone vb) e.o. slowly	they will stir (toneless vb) slowly
va-ri- **vegán**-á gáráha	**va-ri**- **vʊ́rʊ́gány**-á gáráha
L **L** **[H]** H	**L** **[H]** H

(24)

But in the presence of an Object Marker (OM), the observed pattern is a bit different. In the case of a toneless radical, the OM's H tone becomes L—we say that this, rather than saying that the High deletes, because the the melodic H tone spreads leftward to the radical's first vowel, but not further to the left onto the OM:

they counted us
maní **vá**-**kʊ**-**várís**-a
H **L** **H**

(25)

[8](11) kʊʊ[káraanji] – should the prefix be High? ditto: koo-ké[rori], and next example in 11c, and second example in 11d.

[9]in 12, the Tense Marker -áa- is indicated thusly; should it be -áá-?

[10]Which I have sometimes called the *Complex Pattern*.

But when the radical is High-toned, something unexpected occurs: when an OM is followed by a High-toned radical and a melodic High tone immediately follows, then from that triple emerges only a single High tone, which is associated with the first mora of the radical:

they fried it

maní vá-ki-káraang-a

?

H H H H

(26)

Somehow these three adjacent High tones result in a single H-tone on the surface. Why is this tone described as the realization of the melodic High, rather than the radical's own High? In this regard, David notes that this High tone has one peculiarity that it shares with a melodic High, rather than a radical's own underlying High: when this High appears on a long penultimate syllable, it is realized as a long H, rather than a falling tone:

they cooked it

maní vá-ke-déék-a

?

H L H H

(27)

When a radical's High tone is realized on the first mora of a long penultimate, it appears in a falling syllable, as in ku-híiza, and the only case of a Long penult that we have seen is from a melodic tone associated with the second mora of the stem. So on balance, Dave's case for this being the result of a melodic High seems natural.

There is certainly much more to say about this very rich language and its tone system.

6 Bob Ladd on features and autosegments

Bob Ladd published a book in 2014 on simultaneous structure in phonology that was a great pleasure to read, and I remember writing a review of it for *Phonology* in which I washed away all the reviewing sins of my youth, those jejeune opportunities to seem too clever by half, and I did it by inviting the reader to listen (along with me) to Bob Ladd's mature thoughts about the nature of phonological representations, and how simultaneity and other fundamental notions can be most effectively integrated into phonological theory. I hope somebody some day writes about a book I've written in terms like this, speaking of Bob's book:

> Robert Ladd has written the sort of book that senior scholars ought to write but rarely do: a monograph in which the central concerns of a field are laid out with the perspective that only a period of time measured in decades can offer. That alone is reason enough to say that this is a book that any serious phonologist should read. I will try to outline the issues that Ladd reviews, with the hope that this will make clear why I think these reflections deserve the attention of phonologists today.

Towards the end of his contribution to this volume, he makes reference to Steve Anderson's influential book on the history of phonology, published in 1985, where Steve frames developments with respect to two important concerns that phonologists continue to return to. These concerns involve the nature of phonological rules and the nature of representations. Bob draws our attention to the danger that lurks of reducing autosegments to features. He is certainly right. It is also true that we do not fully understand features, and that is in part due to the fact that Trubetzkoy, Jakobson, and Morris Halle (to say nothing of Noam Chomsky) did not agree on what phonological features are. Bernard Laks and I have tried to clarify a bit how Trubetzkoy and Jakobson disagreed about the nature of phonological contrasts and features (Goldsmith and Laks 2019, chapter 9). But there is much that needs to be done.

Intensional definition and extensional definition. In *Battle in the Mind Fields*, chapter 9, Bernard Laks and I discussed some of the ideas that went into Trubetzkoy and Jakobson's conception of phonological oppositions. The most important thing to bear in mind is that oppositions are relations between pairs of objects, and features are properties of oppositions (not of phonemes). The structuralist conception of a system of phonemes is a multidimensional graph, whose nodes are phonemes and whose edges are oppositions. This view is grounded in the logical analysis of ideas that Edmund Husserl famously pronounced in his *Logical Investigations* (1904-05), itself heavily influenced by the work of Husserl's teacher, Franz Brentano, an expert on Aristotle. Laks and I did not emphasize there what perhaps should be emphasized, and that was that traditional Aristotelian logic made clear that there were two ways to describe collections of objects: one could either specify them one by one, or one could provide a set of properties that they all shared (but which were not all shared by objects outsdie the collection). The first was called extensional and the second intensional. This view goes back to medieval discussions. Russell and Whitehead noted in the introduction to the second edition of the *Principia Mathematica*: "It is an old dispute whether formal logic should concern itself mainly with intensions or extensions. In general, logicians whose training was mainly philosophical have decided for intensions, while those whose training was mainly mathematical have decided for extensions." (Cambridge 1925, volI: 25.) Nicholas Rescher, in "The Distinction between Predicate Intension and Extension" (1959) p. 623, cites that passage, and continues, "I propose here to give some indications why the distinction between intension and extension, although of no critical importance for mathematics, not only is useful, but quite important for philosophical purposes." It is very unlikely that one could defend the notion that this distinction is not critically important for mathematics, but what is of interest to us here is that what linguists have come to call distinctive features is nothing more and nothing less than linguists' efforts to replace extensional language in phonology with intensional language. When phonologists talk about natural classes, they are looking at segments from an extensional point of view; when they talk about features, they are looking at collections of segments from an intensional point of view.

Realism and nominalism in phonology There is a traditional distinction in philosophy (in metaphysics, actually) which may be helpful, and which in any event deserves to be mentioned. This is the distinction between nominalism and realism. It is relatively easy to give a simple description of the difference, but it is not at all obvious how this distinction can be clearly applied to the discussions of phonology. Nominalism and realism agree that particulars exist, but nominalism rejects the notion that there are *universals*, and universals are in some fashion over and above particulars. A philosopher trying to be didactic will say that we all (realists and nominalists) agree that there are green things, but only realists believe that greenness, or the color green, exists; nominalists will reject that. The nominalist will warn you that language may deceive you into thinking that greenness exists, since it is a word, and they will warn you not to let yourself be deceived.

Being and nothingness in phonology: segments are objects For phonologists, things are not so clear. Your typical phonologist is seeking to understand the system of sounds in a language, but in a very abstract sense: the phonologist talks about the sound /p/ in English, or the /ö/ in German, and the phonologist takes these to be very concrete things which can be measured. But to a non-linguist, the idea of the phoneme /p/ is wildly abstract, and what is concrete are the various utterances in English in which the sound /p/ has occurred. Bear in mind that what is concrete and particular are events in space and time: phonemes are not events in space and time. If the nominalist rejects abstract objects, those abstractions that are not catalogs of events in space and time, then the nominalist will have no room in their universe for phonemes, or really anything else that linguists generally care about.

And yet. Being a linguist involves signing on to a manifesto of sorts that phonemes (or phonological segments) exist. The description of a language includes statements about the inventory of sounds (at one or more levels of description) in that language. And since at least the middle of the 19th century (I am thinking of the work of William Dwight Whitney) it has been clear to phonologists that sounds can be organized into various subgroups, such as vowels, semi-vowels, nasals, voiced obstruents, and so forth. For the working phonologist, the various sounds of a language are the concrete individuals of its phonology. The question that might separate the nominalist from the realist (within the discipline of phonology) would be whether there are things in phonology other than the phonemes themselves. It was the structuralists' contention (Trubetzkoy and Jakobson, for example) that the binary relations between phonemes (what they called oppositions) were real, and could be studied in a systematic way. They were thus realists among the realists (by which I mean, phonologists are already realists, and they were the realists among the phonologists).

Derivations and assembly-lines From Leonard Bloomfield's *Language* down through phonology until the rise of optimality theory, all eyes were on phonological analysis as a sequence of operations which *change* a basic, or underlying, representation. Certainly both Bloomfield and Zellig Harris were suspicious of the idea of phonological representation as a dynamic account in which one representation is turned into another, but both acknowledged that this was the best way to look at at least *some* things in languge. Edward Sapir, on the other hand, was perfectly comfortable with this way of looking at things, and (most surprisingly) it was Charles Hockett (1954) who first acknowledged his suspicion of such metaphors and then provided a way of restructuring the metaphor in a way that he was perfectly comfortable with. For Chomsky and Halle, the centrality of the derivation (take a basic structure, and then first do this, and then do that, and then ...) was part and parcel of maintaining that the most important components of a grammar are rules that generate the grammatical forms.

Once we have signed on to working with derivations and we have accepted the fundamental difference between a rule, which operates on representations, and (static) representations, which exist and but do not *do* anything, then we have accepted a certain kind of existence for the atoms of our representation. Something that is part of a representation and is not explicitly *changed* by a rule exists before the rule applies and *after* the rule applies.[11]

Binarism and information theory Something happened in the early 1950s which shifted the way in which phonological features were understood by Roman Jakobson and Morris Halle. It might have been due to the influence of Noam Chomsky; I think it was rather working with Colin Cherry to produce the paper Cherry, Halle, and Jakobson 1953. This was the moment at which features lost their fundamental, or essential, connection to oppositions, and became rather functions (as we would more likely say today) that had as their domain the segments of the language, and had as their range (or co-domain) the values +, -, or 0.

This led to a way of talking that encouraged phonologists to say unreasonable things like "segments are bundles of features." The reader can see already how this way of talking is closely related to the issue of nominalism and realism in phonology. To back up a step again, the nominalist believes the individual segments exist, but their properties or relations do not exist; they constitute ways of talking about segments. The realist believes that properties may well exist (they have the right to exist!). The Jakobson-Halle phonologist goes one step further and seems to be saying that the properties exist but not the concretes![12]

I frankly think that there has been some confusion in discussions of features and feature values, and that the confusion begins with the turn of phrase, "a segment is a bundle of features." If pressed on the subject, a defender of that view would undoutedly agree that the intent is better expressed as "a segment is a bundle of feature values"; a /p/ and a /b/ are not the same segments, even if they are expressed with the same features (it is the *values* of the features that distinguish them, they would agree). But that is not quite right either, since a value is either a +, a -, or a 0. What must be said is that a segment is a bundle of pairs of features and feature values.

This then has led to phonologists trying to remove (or add) features or feature values from a segment. Neither of these notions make sense, if features are functions from segments to {+,-,0}. Distributed morphology relies on this formalism perhaps even more than any theory of phonology; it is referred to as *impoverishment*, and is formalized in a fashion similar to the deletion of a phoneme in generative phonology, except that the left-hand side of the deletion rule is the name of a feature.

Two ways in which phonemes can be related Let's back up a step, and remind ourselves that there are two different ways in which linguists study relations between things like phonemes (just like relations among words). The traditional terms for these ways is *in absentia* and *in praesentia*. *In absentia* relations are relations between items in particular categories, and for phonologists, the clearest example is talking about oppositions among phonemes. For example, if a language with vowel harmony has the vowels i and I, with each appearing in the right harmonic domain, then these two vowels are in a relationship *in absentia*: where one appears, the other one does not. On the other hand, if a /t/ becomes /č/ before a high front vowel, then there is a relationship *in praesentia* between the high front vowel and the /t/, and that relationship can be expressed with a phonological rule (although that is not the only way to express it).

[11]Bear in mind that the same is not true for simple numerical quantities. The number 15, for example, like any integer, has a unique set of prime factors (3 and 5, in this case). Subtract 1 from 15, though, and those prime factors are gone; the prime factors of 14 are 2 and 7. Put another way, if we subtract 1 from 15, we have not done something to 15, and the resulting number, 14, has its own properties, and none of it carries over from what it used to be (a ridiculous thing to say!).

[12]The reader may remember the Cheshire Cat in Alice in Wonderland, whose smile outlasts its existence.

Structuralism Structuralism is a framework in which *most* of the interest lies in the relations *in absentia*, while generative phonology is a framework in which most of the interest lies in the relations found *in praesentia* — and indeed, generative phonologists rely principally on relations *in praesentia* (which we call *phonological rules*!) to be sure that they have gotten the analysis into features right.

As generative phonology shifted that focus toward relations *in praesentia*, it became evident that rules of assimilation were, along with rules of deletion, the most common sort of phonological rule being discovered, and rules of assimilation seemed to locate the position, in some sense, of where the assimilation took place. As Bob Ladd observes, when early textbooks of generative phonology described the phonological representation in the form of an array of feature specifications organized into columns, one column for each surface seg more than one way to think of this; "different textbook presentations treat the significance of such representations in rather different terms. Some contain passages implying that features are 'things' and segments mrerely composite or even epiphenomenal—at most, convenient abbreviations for bundles of features." (100) He cites Robert Harms, writing in 1968:

> The fundamental unit of generative phonology is the distinctive feature. . . The phoneme—specifically, the systematic phoneme—clearly has a secondary status, but, for matters of presentation, it is frequently convenient to refer to phoneems as the underlying segments used to designate or "spell" morphemes.

Other textbooks were clear about the idea that what is fundamental in phonology is the set of segments, which form a sequence, and features are a secondary matter. Ladd: "The difference between these two interpretations of segments betrays a crucial tension in the feature concept as it has developed since the 1930s." (101)

Parts, wholes, and mereology In short, just what is the relation of a phoneme to one of its feature specifications? Is it like the relation of a deck of cards to the queen of hearts card, or is it like the relation of an American flag to the color red? Readers of *Battle in the Mind Fields* will recognize this as one of the central questions of Franz Brentano, picked up and developed by several of his students, notably Edmund Husserl: for Brentano, this is the question of the logic of *parts* and *wholes.* There are many different ways in which a part relates to a whole (the whole of which it is a part). A card which is a part of a deck is easily removable, and has no physical attachment to the other cards, but if it goes missing, the deck is no longer a full deck. An object typically has a color, but the color cannot be removed from it (it can only be changed to another color). A sentence is constituted of words, but not the way cards comprise a deck: if we change the order of words, we get a new sentence, not the old one, and the words form intermediate sized groupings which we call syntactic constitutents. The list goes on and on; the search for a general logic of what it means for one thing to be a part of another is by no means a simple operation (and, as I mentioned above, that search was a part of Brentano's and Husserl's philosophical work).

As regards the theory of parts and wholes, the most fundamental division is between those wholes constituted of parts that can be logically removed (think cards in a deck) and those that cannot (think of the colors of a rainbow). For the phonologist, words are built of segments by the relation of concatenation, and a segment can be removed from such a string. Segments themselves, however, are the smallest thing that can be removed; a feature is not a part of a segment, but is rather a property of a segment, and the featural account of segments treats those properties as functions that map from segments to values such as + and -.

In this general context, autosegmental phonology can best be understood as a model of phonology in which segments are concatenated on more than one (autosegmental) tier, and features maintain the character of *features*, although in general each such feature-function takes as its domain only the segment that occur on a specific tier. Bob Ladd writes,

> Segments, in other words, are not merely convenient abbreviations, but are, rather, the atoms of phonology; features are the attributes that define the place of any segment in the overall sound system. (104).

As Bob Ladd wrote, "Feature theory is fundamentally an attempt at formalizing the relation between the simultaneous and the sequential. (88)" This is first and foremost true regarding relations *in absentia*: Trubetzkoy and Jakobson first viewed the relationship between phonemes of a language as forming a crystalline image, and the edges that appeared in this image, linking one phoneme to another, are best understood as features describing the opposition between the two phonemes that define that edge.

Halle (and also Chomsky, with him) developed generative phonology as a means of employing features to describe relations in presentia as well, and in saying that, I am just referring to the ways in which underlying segments are realized in particular phonological contexts. It is this notion "... in a particular context" which is our way (first used by Zellig Harris) to formalize the notion of relationship *in praesentia*.

Is it information theory? Cherry, Halle, and Jakobson was probably the first paper to explore any of the consequences of the new information theory for modern linguistic analyses. Reading it more than 60 years later, some of the choices made are difficult to fathom — and this includes some of the most important choices, relating to the notion of "binary feature." There was, at the time, a relatively new method of creating an encoding system (called Huffman encoding) which allowed one to take a set of options (like the letters of an alphabet, or the phonemes of a phonological system) where we know the probabilities of each option, and create a binary tree in which each node corresponded to a yes or no question, and in which lower probability options would appear lower in the tree (so that they would be encoding with longer strings of 0 and 1) and higher probability options would appear higher (so that they would be encoded with shorter strings of 0 and 1). But it was eventually realized by people working on compression that these yes/no questions missed the point, and were neither necessary nor even useful. Instead, what was used (and is used) is a system that assign

I also think that there is room in phonological theory for a different model, one in which phonological representations and phonological rules become one. OT eliminates rules, and replaces them with constraint ranking. A dynamical systems approach to phonology suggests a different direction in which to go, one in which the very substance used to inscribe representations has itself a dynamic to it—a suggestion that would have resonated well with the gestalt psychologists of the 1920s and 1930, trained as so many of them were in physics. I've tried to propose an analysis along these lines under the rubric of dynamic computational networks, which several people developed to some degree in the early 1990s, including Gary Larson (as Caroline Wiltshire points out in her chapter of the present book).

7 Diane Brentari

Diane's reflections on the role of autosegmental phonology in the study of the phonology of sign languages draws us into some of the same questions that Bob Ladd's remarks do. Looking at work in phonology in the 1970s, 1980s and 1990s, she draws our attention to the difference between the association lines in autosegmental representation — which are the clearest examples one could hope to find of relations emphin praesentia—and the lines which organize tiers in a model of feature geometry (of the sort envisioned by Mohanaon, Sagey, Clements, and others). The central concern of feature geometry has been to provide a universal account of what groups of feature specifications spread in various rules of assimilation across languages. In this sense, the goal of feature geometry is to find a single Venn diagram of all of the subsets of the set of all features (not feature specifications) in which in such a way that only these chosen subsets act as a team in an assimilation process, and if two subsets have a non-null intersection, then one is a subset of the other (which is simply to say that the features could be organized in a tree-like manner).

At the same time that autosegmental analysis seems to be focused on breaking down the traditional phoneme, it also changes what we must mean when we say that a given phenomenon is phonologically local. We are accustomed to observing that most phonological effects involve two adjacent segments, but two segments on a particular autosegmental tier may be adjacent even if they are associated with other autosegments on a different tier that are not adjacent (for example, in the case of two autosegments in a language with nasal or with vowel harmony).

Diane emphasizes that the criteria for autosegmental analysis are not physiological but strictly phonological, just as Bob Ladd emphasizes that if we treat all cases of assimilation as *ipso facto* cases of autosegmental spreading, then we have virtually lost as much as we have gained from the point of view of an explanatory theory of phonology. And yet — it is quite tempting to take advantage of the simplicity of autosegmental notation to describe assimilation even in the absence of other telltale signs of features organized around a distinct tier. My own take on this is by no means worked out in any detail, but it is based on the observation that what the phonologist traditionally takes for granted—which is that a great deal of information in a phonological representation can be packed into a representation in which segments follow one another, like letters on a page — is not easy to justify in a straightforward articulatory or acoustic fashion. Mapping from a spectrogram to a sequence of phonemes is not at all obvious. Why do children succeed in such a mapping? There's nothing wrong with saying that they succeed because of something that is innate, but that's no more helpful than saying that mountain goats climb up sheer mountain cliffs because of something that's innate to them. Perhaps children succeed in segmenting each acoustic/articulatory channel into successive periods that are homogeneous, and then focus their attention on taking this orchestral score — with one part of the score for each acoustic channel — and reducing it to a single staff in which chords are notated. This image of *deautosegmentalization* (which I sketched in the last chapter of my dissertation) suggests that adult language structures that are autosegmental (such as tone in an African language) are the result of the realization (so to speak) of language learners that the segmentation of laryngeal (pitch) information cannot satisfactorily wrapped into the segments that organize all of the oral articulators.

Diane sketches the remarkably history of linguistic analysis of American Sign Language, in which the first remarkable step was that of William Stokoe, who insisted that it had to be possible to ask what minimum articulatory differences are used to distinguish lexical items in ASL. That was the question that American descriptive linguistics took to be central to the analysis of a language, and if ASL was to be treated as a language, that question had to be asked, and indeed, answered. Stokoe's proposal was that signs must be distinct from one another along one or more of three dimensions: handshape, place, movement. It is a simple proposal, but which took deep insight to get to, and the fact that it *failed* to be the final answer (it left unsaid how orientation, for example, could be contrastive) makes it all the more remarkable as a first major accomplishment in the analysis of ASL.

If it were true that all ASL lexical items were specified in precisely this tripartite way (a specification for handshape, for location, and for movement), that would still leave us scratching our heads about how this architecture relates to what we are accustomed to finding in spoken languages. Is handshape a feature with a large number of values (perhaps an indefinite number of values), or is it decomposable, and if it is decomposable, how does a subset of features get together to work on a single articulator (feature geometry would offer an answer to this question, but much later in time)?

I remember hearing from Scott Lidell about the Hold-Movement model of ASL that he and Bob Johnson were developing, beginning in the late 1970s, using autosegmental representations to allow them to express *non*-simultaneities that a traditional segment-oriented notation did not permit. This work was part of the blossoming of linguistic analysis of ASL during the 1980s, which found in autosegmental analysis a way to focus attention on literally *smaller* aspects of articulation, such as the selected fingers employed by a given lexical sign to express its contrastive essence. What was also crucial was the insight that certain aspects of a handshape were best understood as simultaneous, and others as sequential, and that a formal model was needed which would allow for such a difference.

Throughout the work on ASL and other sign languages researchers have kept in mind the question as to how similar and how different the linguistic tools are that are needed to shed light on the structures inherent in signed and spoken languages. Diane mentions the striking observation that while length contrasts are common in spoken language phonology, no signed language has shown such a distinction to date.

Why is there phonology? It is tempting to think that the best answer to this koan-like question is: *there is phonology so that there can be fluent speakers.* No phonology, no fluent speakers, and no fluent listeners either. Diane explains how the size of movements can best be understood as running parallel to physical energy in spoken language (where the primary energy vehicles are vowels). Sonority-poor verbs (SIT) can be nominalized with reduplication, while sonority-rich verbs (INFORM) cannot. To speak a language fluently is to have entered into the tacit agreement among speakers of the language when and how phonological information will be offered by the speaker to the hearer—not too fast, not too slow, not too little, and not too much.

8 Bert Vaux and Bridget Samuels

Bert Vaux and Bridget Samuels address the question of how to think about syllabification in light of all that we have learned about phonology in the last several decades, and how to think about the decision to include information about syllabification in the basic form of a morpheme or a word. These are questions that lie at the heart of most of the papers in this book. In Larry Hyman's discussion, the question arose of whether the basic form in a tonal language could be built from a pair of High-Low tones, or does that redundnancy (why always High-Low? Why not just High in the basic form, and then add the Low by a mechanical rule, or just the Low plus a mechanical rule?), just as in Dave Odden's paper, where the effects of morphosyntactic features are broken up into two parts (one a melodic High tone, and the other a set of rules that apply mechanically once we take those features into consideration). And that's just to mention two.

Some phonologists acknowledge a wish (I almost wrote, "a pious hope") that reflections on the phonologists' basic forms are ultimately reflections on processes in the brains of speakers of a language. I think this is a dangerous hope. By *dangerous*, I mean *risky*, and if a serious cost-benefit analysis of the risk comes out in favor of assuming that basic forms in phonology correspond to structures that can be identified in the brain, then fine: take the risk. What makes this assumption risky is that one is obliged to learn and explore what we know about how strutural information can be embedded in structures of the sort that we have so far found in the brains of humans and species reasonably close to ours. One thing that can be safely concluded from our studies of the brain is that there is no paper in the brain, and neither are there whiteboards of any sort to be found there. Yet sequentiality is absolutely central to linguistic behavior! Can all three sorts of sequentiality be seen as reflections of a single device: there is the more or less fixed sequence of segments in a morpheme (and of morphemes in a word), there is the more or less fixed sequence of words as specified by the syntactic rules of the language, and finally—we must not forget!—there are the choices of the speaker, the decisions that each speaker makes when uttering a sentence, typically choices made within the limits set by the rules of the grammar of the language. All of which tells us with certainty that the brain is able to deal with sequential structure regarding phonemes and words (at the very least), but we haven't figured out how it does it without having any paper to draw on.

What makes this risky is that if we wish to link our ideas about basic phonological forms to the computations performed by the brain, we have an obligation to use the computational styles that the nervous system can use, as far as our best current theories go. If we do not want to do that (because, perhaps, we cannot think of a way to model sonority within the computational style that we understand for neural networks, in a biological responsible sense), then it seems to me that we are setting ourselves up for disappointment: we may *think* we are talking about the brain, but it might just turn out that we are just developing another computational model that can be implemented in Python but not in the style of computation employed in Wernicke's area.

These common-sensical observations passed through the head of this reader when I read that from some perspectives (that of Bromberger and Halle 1989, for example) the linguist should be concerned with the "need to reduce long-term memory burden." (147). What would long-term memory burden be, from the perspective of a phonologist? Should it be identified with a cost attached to assigning a + value to a feature, and no cost to assigning a 0 value to a feature? On such a view, the linguist might try to figure out a way to have more 0's even if that means devising a rule that shifts those 0's to a + or a -. But that's not what garden-variety encoding theory suggests. Under most circumstances, the simplest way to encode data (where *simplest* here means requiring fewest bits) is to devise a probabilistic model of all possible messages, and partition the interval from 0 to 1 so that every possible message is associated with a unique subinterval. Then the encoding that we will assign to a message will be (roughly) a string of 0s and 1s which points unambiguously to the subinterval associated with that message. It's all algebra; and it can be done base 2, base 10, or any other base. The probabilistic model (not a frequency-based model!) will be able to assign shorter encodings to messages that it assigns higher probabilities to.

A consequence of this is that to reduce long-term burdens on memory, the right strategy is not to look at a phonological representation and see what could be left out of it, and then writing something that looks like a phonological rule to put it back in. I say *it's not the right strategy* not because I have a different strategy to offer to phonologists, but rather because we already know what the best methods are for data compression, and we know them because people have been working hard on this problem for the last 70 years.

To put it slightly differently, the compressed form of symbolic information does not look like the original with some symbols left out (it doesn't look like someone's home-made shorthand, so to speak).

> Unfortunately, none of what I have just said sheds any light on how an autosegmental (or more generally, a tier-based) model can provide an *explanation* for what a High tone spreading right to left (as in Logoori) will be blocked by the appearance of an association line that gets in the

way (as the radical's High tone, which is lowered to Low (by a melodic High) can then block the spreading into the verb of a High tone from the immediately following word. Do I think that some day we *will* be able to connect those to ideas (blockage of spreading by an association line in place)? I would say: I hope so.

I'm not sure what to conclude from Bert and Bridget's relatively brief discussion of tone in Makhuwa (163-64). Coda consonants are positions in which tone is realized, and they write that "these tonal specifications must therefore be stored with the relevant consonants in the corresponding lexical entries." They provide three nouns, two of which have class prefixes (*n-latu, ma-kristao*) and one that does not indicate whether it has an overt noun class prefix (*nantana*); *na-* is probably not a noun class prefix, but it might be a conjunction. It appears that the /k/ of *kristao* bears a High tone and is syllabified with the preceding syllable, but it's not clear that there are any other stem-initial *kr-* clusters; it's a borrowing, of course, and we need to know more about the word before we can draw any conclusions.

Bert and Bridget's central point is that it is possible to find aspects of phonological structure that are both predictable, in and of themselves, and yet essential for specifying at least some contrastive phonological information. Perhaps something like the Makhuwa structure (with an eye to Larry Hyman's and Dave Odden's analyses) might provide a case of this sort. In many Bantu languages, all vowels are bimoraic before a nasal cluster (-nC) with no possiblity of an exception: a post-lexical phonological generalization if such things exist. Yet the assignment of tone to the verb stem (in particular, a High to the second mora of the stem), very much a lexical, morphophonological rule, needs to know that the first syllable of a stem like *-liimb-* is bimoraic, in order to assign the V2 H to the second mora of the long first syllable of the stem.

8.1 Who was that?

I can't say that I remember teaching the first class of the year when Bert Vaux was a student at the University of Chicago, but there is one thing that has stopped me cold in my tracks– it's that footnote 1, in which I was 15 minutes late to class. How is that possible? I have no patience for people who show up late to teach their class. Yet could Bert's vivid memory be *wrong*? We are going to have to track down the other students in the class and see what they remember. In the meantime, I'm awfully sorry. I won't do it again.

9 Caroline Wiltshire

Caroline's discussion of sonority waves in such languages as Spanish and Malayalam underscores the possibilities of thinking outside the current boxes with regard to what syllabification *is*. I said just above, thinking about Diane's remarks about sign language syllables, that phonology is there to allow for there to be fluency in speaking, and in signing. And fluency is in large measure a matter of producing syllables of one's language in the proper way. And so it is perplexing that it is so common in languages to find conflicting constraints on possible words and on possible phrases (or utterances), especially in view of the fact (I think it is a fact) that constraints on words can be either more or less strict that constraints on possible utterances.

One of the perplexing challenges for dealing with syllabifications, as phonologists look at things, derives from the perception that syllabification is not part of the linguistic system aiming at marking differences between words (or, if you prefer not to talk about words in phonology, then we can talk about differences between morphemes). (Why is that perplexing? Because phonologists focus on those aspects of phonetics whose function is to distinguish two utterances.) As Bert and Bridget were just pointing out in the preceding chapter, we don't find languages using syllabification to mark a contrast between [$ CV $ CVC $] and [$ CVC $ VC $]. We find things a little bit like this when we compare *nightrate* and *nitrate*, but that's not because English uses syllabification to contrast, but because the difference of syllabification here results from the presence of the morphology of compounds in the *nightrate* but not in *nitrate*.

So what difference (or differences) do differences in syllabification make? They identify differences between *languages* (which makes them naturally the subject of work on typology) rather than differences between words (or morphemes). Where does that kind of information *go* in a grammar? A framework like *principles and parameters* has an easy answer to that question: that kind of information is precisely the selection of values for each parameter. Phonological theories generally don't have easy answers like that. Optimality theory can point to constraint ordering as that which marks the difference between two phonologies.

Caroline's discussion of L1 and L2 syllabification is another way to arrive at the same question. L2 syllabification serves as a testbench for principles of syllabification in the learner's L1.

I was struck by something that Caroline wrote (though she says that I wrote it first)—that syllabification is not just an effect of which the sounds are the cause. Linguists certainly employ the notions of cause and effect, but not in the way that social scientists regularly do. Social scientists (really without exception) have signed on to a positivist view of the world in which reality is the totality of events that occur in space and time, with the understanding that space and time are shared by all of us. While some linguists agree with that—I would say, for example, that people doing Labovian sociolinguistics put themselves in that category—most linguists do not, especially if they take the object of their study to be individuals' knowledge of language. But the notions of cause and effect are just too hard to drop. It is difficult to imagine what would count as a satisfactory *explanation* of something if we could no longer speak of causes and effects. And so, we linguists find a way to divide the world of linguistic observations and linguistic generalizations into causes and effects. If we are interested in syllabification, we may notice the absence of sC-initial words in Spanish, and account for their absence by hypothesizing a change (in this case, an epenthesis of *e* word-initially before an "impure s" sequence, i.e., sC) caused by such a structure, and in like manner, as Caroline notes, a semi-vowel /j/ may strengthen to a fricative in Porteño Spanish when it occurs in syllable onset position. We speak of this fortition as "attributed to sonority," and "attributed to an OT constraint" which takes the form of an order: "be strong in an onset." (An order is not quite a statement of cause and effect: if I order you to shoot a gun, and you do it, it is only in some unusual situations where we would agree that I caused you to shoot the gun, and the same is perhaps true if I made you shoot the gun, or had you shoot the gun, or even got you to shoot the gun.) Caroline suggests that "consonants must be parsed in onsets to be evaluated by this constraint [HONSET]." Constraints in OT have little force that they can muster *against* a representation that violates them; a constraint's objection to a particular representation is a bit like passive aggressiveness: a constraint may claim that its violation runs against the pattern of well-formedness (and not just in one language, but necessarily in all languages), but it's not that constraint's fault if its violation had no impact on identifying what the best output representation turned out to be.

10 Mark Liberman

What is the difference between autosegmental phonology and articulatory phonology, as developed by Catherine Browman and Louis Goldstein? The biggest difference is the commitment of articulatory phonology to develop a way for phonological knowledge to connect directly to the human articulatory apparatus. How can we figure out the optimal way to maintain a robust theory of phonology which has a responsible attitude towards articulatory detail? Mark reminds us that this is an issue that is always with us.

I can imagine someone (even myself) saying that the difference between autosegmental analysis and articulatory analysis is the difference between what the speaker *knows* and what the speaker *does*. But I would not put too much stock by such a statement, because we don't have as clear a handle on that difference as some might think we do. There has long been a tendency to try to separate phonological effects into two sorts, with the hope held out by some that one part is what is known, and the other what is done: for those coming of age in the 1980s or so, it is natural to wonder whether the difference between what Paul Kiparsky called lexical phonological rules and post-lexical phonological rules was not just the difference we were looking for, though we can find a very similar tendency as far back as the work of Baudouin de Courtenay a century earlier.

Mark asks the phonologist to take the bull by the horns, which means here that post-lexical phonology *is* articulatory phonology. Or maybe not post-lexical phonology, but the treatment of non-contrastive elements of phonology.

I think that any phonologist or phonetician who has any interest at all in English must come back, again and again, to the question of the flap in American English. Mark must feel the same way. He points out that *at all* is pronounced with a flap, while *a tall* never is.

> The relevant generalization is roughly that all non-onset consonants are weakened, and in the case of intervocalic /t/, the closure is weakened to a ballistic tap and the laryngeal gesture is

weakened to the point of disappearance. For phonetic interpretation to be an option in such cases, phonological structure must be available to be interpreted. (204)

When I read that, I had to stop and think about it for quite a while. For two reasons: the first is that this statement has to be made a bit more complex to handle the facts, and second, that the complexities we need to add do not make Mark's statement any less true. And for a third reason, I suppose: how sure are we that syllable structure is an example of "phonological structure," if by phonological structure we mean something which can be expressed over the symbolic units (phonemes and constituents made of phonemes, and the like)? Lurking behind all this is the question of what *kind* of structure is appropriate for syllabification: is it like constituent structure in syntax, and if it is... what does constituent structure in syntax really look like? Let me explain.

Complexity of the context for flapping: Word-internally, flapping is conditioned by the stress on either side. If the vowel following the /t/ is stressed, flapping is never possible (though obviously it *is* possible across word boundaries, as in *at áll*. If the vowel following the /t/ is not stressed, then flapping is obligatory when the preceding vowel is stressed, and optional when it is unstressed. But these generalizations surely concern not flapping directly, but rather reflect the complexities of syllabification, both inside words and across word boundaries. And it does not stop there: when a word is *t*-final and is followed by a vowel-initial word, then the *t* typically flaps, unless it has been glottalized (an option which is, I think, reflective of a speaker's option regarding phonological phrasing). This is the case which is operative for *at all*. And worst of all, or best of all, is how word-initial *t*s are treated in American English when the preceding word ends with a vowel. Such *t*s *never* flap (*the tomatoes*)—unless the *t* in question is part of the preposition *to* or the prefix *to-*, as in *tonight, tomorrow, today, together....*

In my opinion, all of these (rather crazy) conditions teach us directly about how American English syllabifies sounds, and only indirectly tells us anything about flapping. Flapping "simply" applies when a *t* is ambisyllabic (as Dan Kahn taught us so many decades ago). Now Mark in this paper is exploring how we can reasonably create a theory of spoken language which contains within it both a set of symbolic representations and a model for the implementation of corporeal gestures. I've put it that way to leave wide open the question of how the symbolic/gestural parts of the model relate to one another. In our current views, so heavily influenced by traditional computer programming, most of us are inclined to phrase such a question in terms of an interface between two components, and to believe that the best answer (to the question, What does such an interface look like?) is a statement about the syntax of a representation that can be generated by one component (symbolic phonology) and accepted/implemented by the other (gestural phonetics). I actually don't think this is a good way to pose the question, or to pose the answer. But even if we agree that this *interface* understanding is a good first approximation, where does syllabification fit in? Are there two different representations, one that employs notions like constituency and another that unrolls dynamically in time?

Mark reviews some recent discussions of the interaction between diphthongal raising (usually called "Canadian raising") in words such as *writing*, a word whose diphthong is different from *riding*, and asks (on p. 212) why don't we say that the interface representation between phonology and phonetics has ambisyllabic consonants, and the same vowel (i.e., *aj*). The phonetic component will deal with the allophones of *aj* and the flapping of the ambisyllabic consonants. From Mark's point of view, that is the null hypothesis, too.

If we had the opportunity to ask Zellig Harris or Charles Hockett what they thought of such an interpretation of phonology, they would undoubtedly say that this eliminates phonology as they know it, leaving only morphophonology *and* the computation of various kinds of phonological structure (in this case, syllabification, but no doubt also processes like tone spreading). They would both say, "and when I say 'phonology,' I mean, of course, phonemics. Gone are the statements about phones and their distribution, and their falling into various groups with complementary distribution." It's gone, because we think we can find a better way to account for allophony. If we were to ask Morris Halle in 1960 what he thought about this, he would have said, "No, I don't think this is right; all the rules of allophony are cut from the same cloth as the rules of morphophonology, and even Sapir was coming round to accepting that when he died. There is only one kind of phonological rule, and it has a deep underlying representation, and is turned into a surface representations by means of a sequence of ordered rules."

We don't know exactly how a phonetic implementation component would be structured, but it is not an unreasonable guess to say that it would not contain many ways to allow for its version of rules to be ordered; indeed, people have generally imagined that such a phonetic implementation component would look to the outside world like a component in which all of the internal generalizations are in counter-feeding/counter-bleeding orders (which is what you get in a rule system if rules apply all at once, working off a common underlying form).

As an interesting aside, I would mention that this is how Zellig Harris thought of rules of allophony—there is just one side to his view that (even though he announced it quite explicitly) others would view as shocking, though I don't think people quite picked up on it. In particular, he thought the rules of allophony were conditioned by segments at the phonemic level, not at the phonetic level—quite different from the way the arch-Bloomfieldian Hockett viewed things. Mark agrees with Harris in this regard, in the sense that Mark's suggestion is that the symbolic representation that is computed

Mark's second example, second case study, is that of *s*-aspiration in New World Spanish, a process found in a large number of Western Hemisphere dialects of Spanish. In its simplest (and historically earliest) form, it is a pronunciation of the phoneme /s/ as an [h] in preconsonantal position ([ahta] "hasta"); in some dialects this is generalized—to phrase-final position and even word-final position, and it is often described as a suppression of the oral articulation associated with the *s*, leaving only a laryngeal gesture as its phonetic realization. I do not know of a reason why this could not be treated in a non-segmental model which had access to structural information (such as syllabification and word boundaries).

It would not surprise me if a person's willingness to shift as many processes as possible to a non-segmental phonology were deeply influenced by their disciplinary allegiance. If a phenomenon can be equally well treated in two different ways, and the different ways involve different journals and annual meetings, will not an individual be inclined to argue for as many phenomena as possible to be treated within their discipline? One would have to think (and tread) carefully in such cases. (I think Mark is alluding to much this point in the section he entitles "A little disciplinary history.")

Doug Pulleyblank was the first phonologist to focus on the differences between tonology within the lexical phonology and within the post-lexical phonology (in the sense of 1980s lexical phonology). A lot of puzzles remain here, and there is a lot to be learned by trying to be clear on what tonal processes are lexical and which are post-lexical (or whether that difference hinders more than it helps, in some cases) – but by and large, the tonology "in the lexicon" and post-lexical phonology look an awful lot alike. We know for sure that rules of bounded and unbounded tone spreading occur post-lexically, spreading comfortably from one word to the next (often in an anticipatory fashion, which is to say, from right to left as we write it). On the other hand, working on Bantu tonology is great fun precisely because every Bantu tone language has a set good-sized set of different tonal patterns used in different morphosyntactic contexts, which can in no way be viewed as a part of post-lexical articulatory phonology. Larry's paper and Dave's paper towards the beginning of the volume illustrate this, and I've put a few forward references there to this discussion here. I've wondered for a long time whether in Bantu tone languages, tones are associated to a single vowel in the lexical phonology, and then spread only in the post-lexical phonology (it may be that Doug Pulleyblank prosposed that somewhere). In Tonga, there is a tone rule that has a big impact on the verbal tone, a rule which deletes the first High of a verb. Since that High is often associated with a sequence of word-initial syllables, this rule has the effect of making a sequence of verb-initial syllables Low (rather than High).

Mark looks at s-aspiration in Spanish, just as Caroline does. (I'm guessing that neither of them knows that one of the first papers I wrote on autosegmental phonology addresses that topic — it appeared in the LSRL volume edited by Donna Jo Napoli, way back in 1981, which is going on forty years ago.) Mark notes that this is a good candidate for a rule that need not be written symbolically, and could be an example of what a speaker knows how to do. In one respect, this phenomenon resembles flapping in American English (at a pretty high level of abstraction, though). I believe that in some dialects of Spanish (though not the Porteño Spanish that I had studied), aspiration is possible in word-final position, even when a vowel follows (though of course not inside a word when a vowel follows). This at least suggests that we consider the possibility of syllabification having two sides to it, so to speak, for the same reasons that one must say something like that in English. In English, the case is clear that a word-final t, when followed by a vowel-initial word, resyllabifies —this is the "at all"' case; and that resyllabification does not in any sense undo the affiliation of the *s* with the preceding syllable (it it did, it would guarantee that the *t* would not become a flap) In the Spanish case, if a dialect allowed aspiration in a phrase such as "sus abuelas argentinas" (her Argentinian grandmothers) it would be the result of a syllabication that took only a word as its domain. It is well-known, I think, that Spanish phonology rarely takes word domain into consideration (the best studied phenomena in Spanish phonology include the stop/spirant alternations and nasal assimilation).

11 Jackson Lee

Jackson writes about topics that he and I have discussed many, many times: how to do better linguistics by incorporating machine learning – and not just that, but also the methodologies that have grown up in the machine learning community, which emphasize the importance of sharing code and data as part of the what it means to be a responsible member of an academic community.

Minimum Description Length (MDL) methods can teach us a lot, but as is well known, they do not give us any insight into how hypotheses are formulated by a machine that is attempting to model a large amount of data. There is a long history (which, in my own history, began with more than one paper by Jim McCawley on language acquisition, in a day long gone by) of visualizing language acquisition as a sequence of small steps, each one of which is less miraculous, but which taken together give the appearance of being little short of miraculous. If this were so, it could be paraphrased as saying that learning is local, which is to say, structure will undoubtedly be added as the sequences of the grammar become more complex, but what we would not expect to find is that a false hypothesis is adopted early on, and only later is the false hypothesis removed and replaced by a quite different hypothesis. I'm thinking of an example like learning word order in German: a learner hearing only simple sentences could well imagine that the basic word order is SVO, but later see data that includes sequences of sentence-final verbs that give rise to a wholesale reanalysis of word-order in which SOXV becomes the new best hypothesis. This reanalysis (if it is, or if it were, possible) would run counter to the "each step in learning must be small" view because it requires the learner to come up with two quite different but major changes at the same time: changing the underlying word order, and developing a hypothesis that moves the finite verb to second position. In the work that I have done, it has never been necessary to require the learner to give up a hypothesis that seemed to work well, but which ultimately has to be abandoned in favor of a better approach. It will be very interesting when the day comes that we can observe that.

MDL, as Jackson emphasizes, places equal weight on hypothesis-simplicity (as measured by minimal bit length of the grammar in a universal Turing machine) and the degree of fit of the data to the hypothesis (as measured by the inverse log probability assigned by the grammar to the observed data). As I look at the successes that various generations of Linguistica have display, and their failures as well, it seems to me that the first of those two – grammar length – has more to teach us than the second, the information content of the data, given the grammar. The conceptual substance of that second term is a measurement of how large a part of the total grammar space is filled in by observed data. The probability mass of everything generated by the grammar remains 1.0, and the larger a proportion of that space is demanded by the observed data, the better the system works. As we let the grammar generate more data (but we do not let the size of the data grow), the smaller the probability assigned to the (observed) data, and thus the probabiity of the observed data goes down (and its inverse log probability gets larger).

There is a different way to measure the goodness of fit of data to grammar, one which I have been exploring in the most recent version of Linguistica. It is consequence (at least in the way that I arrived at it) of the learning strategy that can be expressed roughly as: whatever language you want to learn, start with the low-hanging fruit first. That is, languages will differ in fundamental ways as regards where the simplest generalizations to learn are. Accept that, and do not insist on learning phonology before morphology, or syntax before morphology. Allow the learner to try to learn several generalizations at the same time, and measure how effective those generalizations are, allowing the most effective learners to be the first to have their generalizations integrated into the overall grammar.

That general spirit of learning suggests a different way of evaluating goodness of fit of theory to data: evaluate a hypothesis not on the basis of how well it assigns probability to all of the data, but rather: ask each hypothesis to list predictions it makes, and rank hypotheses on that basis at each point. For example, a morphology-learner might learn a set of signatures in such a way that the signature set overfits the data; this is in fact what linguistica always does. In a language in which there are multiple layers of morphology (true of almost all languages), linguistica produces a finite state automaton with far too many states, and linguistica can formulate certain hypotheses as the collapsing of two states. Such collapsing always predicts the generation of words that were not previously analyzed but were observed, and we can use the count of newly analyzed words as a measure of the successful predictions of a hypothesis (again, the hypothesis might be as simple as the collapsing of two nodes).

12 James Kirby and Morgan Sonderegger

I came to the University of Chicago in 1984, moving there from Indiana University, where I had been hired in 1976 as part of a binge hire—two ph-linguists hired from East Coast institutions—that was Bob Port and me. The 1980s were an exciting period of time in Bob's lab, with a lot of energetic grad students working, students who would continue to be active in the field in later years. Bob was showing that phonologists had missed what he called "incomplete neutralization" – that contrasts that we had thought were lost in phonologically defined positions were not really, not completely neutralized. James and Morgan return to this debate and use it as a case study of how to think about experiments that fail to show results, alongside experiments that do show results (in short, the relationship between type 1 and type 2 errors). Statistical power (I did not know the term) is the probability of committing a type 2 error. We would like to not commit type 2 errors; we would like to avoid concluding that there was nothing to observe when indeed there really was, and we can say that an experiment has high power if it is designed to help us avoid that pitfall.

I learned a lot from reading this paper by James and Morgan (and I am sure that if I knew more before reading it I would have learned even more from the reading). It leads me to think more about type 1 and type 2 errors in the context of unsupervised learning in linguistics—take the case of unsupervised learning of morphology, such as that done in Linguistica (version *.*). Type 1 and Type 2 errors happen all over the place.

Linguistica's primary hypotheses about morphological structure are its signatures, which are sets of stems and affixes which appear together (that is way too sloppy a definition, but leave it at that for the time being). For all sorts of reasons (Zipf's law frequencies being the main one), stems typically do not appear often enough to appear with all the affixes they in principal could appear with, and Linguistica typically hesitates to go beyond the data it has seen to predict unseen forms (though it's getting smarter — but that's not my point here). Linguistica continues to make far too many signatures, which could certainly be interpreted to mean that each signature is a hypothesis about the affixes that will appear with the relevant stems.

How should we think about Linguistica's hypotheses about the morphological structure to be found in a large corpus of words of a language, and what would a reasonable null hypothesis be? I'm not quite sure. Obviously a null hypothesis like "there is no morphological structure in these words" is not going to be helpful or useful; evidence against it is just too easy to find.

Can we say meaningfully speak of hypothesis testing as we consider a sequence of more and more uncertain signatures? The more stems associated with a signature, the more secure we are in its existence. Well, not exactly. Consider a case like Swahili verbs, which could begin with any subject marker out of a set of 18. There will be many, many cases where only a subset of those 18 prefixes appears on the stems that follows, and Linguistica will consider each subset to be a separate linguistic category (what it calls a *signature*). For Linguistica to declare that such categories exist (when it doesn't really: it's just an accident that the Class 8 prefix did not appear, let's say) feels partly like a type 1 error and partly like a type 2 error. A type 1 error finds evidence for a hypothesis that is wrong, and this hypothesis is wrong (the hypothesis is that the stems in this signature will not appear in the future with the missing Class 8 prefix). So the more stems Linguistica finds in this signature (the one where the Class 8 prefix did not appear in the corpus), the more evidence it has in support of a false hypothesis.

Part of the problem with what I'm saying is that it's easy to be confused about what should count as the null hypothesis in the case I'm describing. It's fine for a system to conclude that there should not be a category of stems that take all subject markers except for class 8, but what hypothesis do we fall back to? Certainly it makes no sense to say that we fall back to no hypothesis at all (Linguistica can be perfectly certain that it has properly identified the word-initial subject markers themselves). It would be totally reasonable for Linguistica to fall back onto the hypothesis (which is correct) that these stems just *happen* to appear without the class 8 prefix.

The problem is how to relate a large set of data to a simpler underlying model—though put in those terms, that seems like just about any problem we'd every want to deal with.

You can see where this is going: systems like Linguistica need to be informed about how to calculate and then use the power of the experiments that it is running on the data that it analyzes. Thanks, Morgan and James (from Linguistica)

13 Khalil Iskarous and Louis Goldstein

Talk about trepidation. Khalil Iskarous and Louis Goldstein look at the model of dynamic computational networks (DCNs) that I worked on with Gary Larson in the late 1980s and early 1990s. That was a period during which there was a lot of excitment about works with neural networks, and I was reading a lot of it. It was also a period in which there was considerable interest in phonology in looking for ways to integrate constraints and rules in phonology. My work on DCNs was intended to address the question as to whether the rules and representations could be integrated, so that there would be no difference between the two. The *difference* between rule and representation seemed to be very solidly built into generative phonology (indeed, generative grammar more generally). Even lexical phonology, the most influential (and interesting) theory of phonology developed in the 1980s, kept rules and representations just as different as SPE phonology did, though it succeeded in merging the notion of morpheme structure conditions (constraints) and lexical phonological rules, which was a big, big step forward.

To think about neural nets, it's important to bear in mind that there are three things to look at, in a model: first, how are the units connected (which is to say, what are the edges in the network)? Second, what are the weights associated with each edge? Edge weights are understood as changing only slowly (if at all), and constitute long-term knowledge. Third, what are the activation levels of each unit, and what do they *mean*? And then there are possibly two more questions: just how are the activation levels for each unit computed? This question is sometimes answered obviously, but in the case of DCNs the computation is not exactly like any of the standard or familiar models. Then finally, how are the weights determined, i.e., learned?

As Khalil and Louis point out, DCNs also represent an effort to keep computation *local* while producing global patterns (for example, global patterns of stress assignment to words in a language). (The same is true to some degree of SPE-style rules with iterative application, but in a very different fashion.)

I have to quote most of a long paragraph from this paper:

> the basic mathematical tool at the heart of the dynamical systems analysis of nature is the notion of a differential equation, originally proposed by Newton. He, along with the great mathematical physicists of the eighteenth and nineteenth centuries, developed the idea that the fundamental laws of physics, which describe how physical fields such as positions of particles, temperature, fluid density, and so on, depend on space and time, are fundamentally local in nature. These laws, stated as differential equations, describe how the physical field-dependent variable relates to the physical field at immediately neighboring nearby instants of time and points of space. What the matheamtical physicists were interested in was the global behavior of a physical field-dependent variable, such as the entire trajectory of a paricle, the flow of blood in an entire vein, or the temperature field of an entire bar of material. But Newton's insight, applied initially to the trajectories of particles due to applied forces, was that the laws of physics do *not* directly govern the global behavior of the field, rather they only specify the local relations between neighboring values of the physical field-dependent variable. Mathematical tools such as calculus can then be used to obtain the global behavior from the local description given by the physical laws.. . . sonority in syllable structure and prominence in metrical stucture are dependent field variables, and that the possible patterns of syllabification and metrical patterns we observe typologically are the result of local relations between syllables. And we hope to show that deep appreciation of how local computaiton leads to a global pattern in mathematical physics can be greatly rewarding for a linguist.

There is a lot in that passage that goes beyond what Larson and I were aware that we were trying to do, including first, that the network-activation-update rule that we developed could and should be thought of as an approximation to a differential equation, and second, that we should find a way to think about the problem as one involving an underlying field. Let's think about what a field is, in the physical world.

There are two physical fields that we learn about in elementary physics: a gravitational field, and an electromagnetic field. It was Michael Faraday who brought the notion of field to physicists, in the first half of the 19th century. Introducing a field is a way of rethinking forces that would otherwise be thought of as arising between various pairs of objects, just as Newton's law of gravity said that there was a force between two objects that was directly proportional to the mass of each. In a complex system with several objects, one would compute forces between every pair of objects, and then add them (why add? Because [and this is a statement about physics, not mathematics] *forces add* — and this is why the DCNs are linear, so that their pseudo-forces add).

When we think about force *fields*, however, we don't think so much about forces between pairs of objects as much as the way in which all of the objects present jointly conspire (so to speak) to create a field over space. This field may be a vector field, or a scalar field (meaning that its value at any point in space may be a vector or it may be in a real number).[13]

And then, in turn, each object interacts with the field (rather than interacting with each of the other objects). This amounts to a massive conceptual simplification of the physical problem.

It did not occur to me, ever, that the DCNs were a step towards viewing syllabification and stress assignment as properties of a field. Why? The field interpretation effectively says that the interactions between adjacent segments (involving sonority or stress) are *not* that: they are interactions between "places" in the field that the segments happen to occupy.

I wish I had seen this, because it has as an immediate consequence a way to think about the most troubling failure of the DCN approach, which is the *appearance of adjacent maxima.* That is, I used DCNs as a way of computing where peaks of activation occurred, and while they worked well to do that in a wide range of cases, there were cases where adjacent segments had to be identified as peaks. But this is not possible; a peak is a segment whose activation is greater than that of the units on either side, so two adjacent segments cannot be peaks. But if the right way to think of sonority and stress is as a field, then the computation of peaks involvs finding peaks of the field, and it is perfectly reasonable to suggest that linearly adjacent segments could be positioned at successive peaks of the field.

Khalil and Louis begin by asking us to imagine a string being picked up at one specific point along its length, in preparation for plucking it. They give a discrete version of the Poisson equation, where y_x is the height of the string at point x, and f is the force applied at that point x:

$$-2\left(\frac{\frac{1}{2}(y_{x-\Delta x}+y_{x+\Delta x})}{\Delta x^2}-y_x\right)=f \tag{1}$$

We can make this simpler, perhaps, if we emphasize the discrete character of this description; we could consider only a set of points $\{x_i\}$ along the x-axis, separated by $\Delta x = 1$:

$$\frac{y_{x-1}+y_{x+1}}{2}-y_x=-f \tag{2}$$

This equation expresses the close relationship between the force (on the right) and (inside the parentheses) the difference between y_x (the actual position of the string at position x) and the average of the positions just to the left (at moment $x-1$) and to the right (at moment $x+1$).

In just a moment, another form of this will be useful, in which the y_x position is described in terms of the other variables:

$$y_x=\frac{y_{x-1}+y_{x+1}}{2}+f \tag{3}$$

Bear in mind that when we say that the string was "picked up" at point y_x, we mean that a force was applied to the string at that point. We do *not* mean that the string was picked up to a specific height; we rather say that a given force is applied at that point, and the height (along the y-axis) that the string goes up there) is the result of the elastic character of the string (and Hooke's Law[14]).

And the notation here does not underscore the fact that the force f on a given position is a function of (depends on) the position, so we probably ought to be writing:

$$y_x=\frac{y_{x-1}+y_{x+1}}{2}+f(x) \tag{4}$$

[13]When that field is a scalar, then it may turn out that the *gradient* of the field, which is a vector, is of great interest: that is, the scalar field conceptually creates the vector field.

[14]Where is Hooke's Law in this derivation, you may ask. It is hidden; it is used in the derivation of equation (1) and the equations derived from it.

Quick digression Khalil and Louis point out that the continuous version of this equation is $\frac{d^2}{dx^2} = -f(x)$ (they write f, not $f(x)$). This will become important below, in the context of linear differential equations and Green's functions. But how did this simple discrete version turn out to be related to the second order differential operator? The key is this: when we discretize a function, we think of a grid (it could be 1 dimensional, as in the current case) of successive points, each of which has a numerical value. The first derivative of the values is the difference of the successive values (think of it being well-defined inbetween the i^{th} tick and the $i+1^{th}$ tick), because the derivative is a limit of $\frac{f(x+1)-f(x)}{\Delta x}$, and here the denominator is always 1 (this is, after all, discretized). The second derivative in this discretized world is the difference of two adjacent first derivatives, divided by 1. So here the difference is $(f(x+1) - f(x)) - (f(x) - f(x-1)) = f(x+1) + f(x-1) + 2f(x)$, which is twice what we see on the left side of equation (2). In short, equation (2) says the second derivation of the function we are looking for must be -1 times the set of boundary conditions, i.e., the underlying phonological conditions. **End of digression**

They then make a surprising observation (that stunned me). Suppose we apply the same force to two adjacent position on this (discretized) string (positions 60 and 61, they say, along a string of length 100). The result is *not* a string with two points (60 and 61) at the same height — no, because applying the same vertical force to two adjacent points does not lead to an equilibrium state where the string is at the same height in these two places.

Why is this? Let's think about what the height will be of the string at position x when if we apply a unit force at that position (and that position only). Will the height be different at positions 1, 50, and 99? Our familiarity with rubber bands allows us to say firmly that the position will be highest if the force is applied at position 50, and lowest if it is applied at positions 1 or 99. At position 1, if the string is perfectly elastic, there will be a much greater force pulling the string down when we apply the unit force at that point. More generally, the downward pull from the left-side of the elastic will be inversely proportional to the distance from the 0 point, and the downward pull from the right-side of the elastic will be inverely proportional to the distance from the 100 point (the right end of the string). So if we apply a unit force at points 60 and 61, the downward force on the string at point 61 from the right-end point is greater than the downward force on point 60 from the left-end point. Wow! This remarkable result is due to the Khalil and Louis's decision to focus not on how high the string was plucked, but rather on the height that would be attained if a fixed *force* were applied at various points.

They point out that what is called a relation method can be used to find the values that y takes on, given a known set of forces applied at particular locations inbetween 1 and 99. A relaxation style of computing consists of a sequence of assignments of values to y_x in which at each computational moment, each unit updates its value based on the local computation given in (4). And that is exactly the algorithm that we used in determining the activation levels of the DCN.

13.1 Green's function

Khalil and Louis then turn to another approach to the general problem we are considering, this approach involving Green's function (which I had never heard of).[15] These methods are typically used in continuous models, rather than discrete, but Khalil has made the case that it is useful to discretize models to gain a good intuition about what is going on.

It is often the case in physical systems, we can describe a problem by using a linear differential operator $\mathcal{L}$ — not something familiar to most linguists. Such an operator applies to a function to give another function, and a good example of a linear differential operator is differentiation. You can take the function $y = x^2$, apply the differentiation operator (which is linear, too), and get a new function $y = 2x$. The second derivative is also a linear differential operator (mapping x^3 to $6x$, for example). Typically we want to find a function $f(x)$ which has the property that a particular linear differential operator (like differentiation) maps it to a particular known function.

Consider two simple examples. In the first, we take $\mathcal{L}$ to be the first derivative: so $\mathcal{L}(f(x)) = \frac{d}{dx}f(x)$. Suppose we want to find a function $f(x)$ which has the property that $\mathcal{L}(f(x)) = f(x)$ — find a function which has the property that it is its own first derivative. The answer (of course?) is $f(x) = e^x$.

[15]What I learned I have learned from *Mathematical Methods for Physics and Engineering*, K.F. Riley, M.P. Hobson, and S.J. Bence, pp. 492ff.

A more relevant example is the case where we are interested in a linear differential operator which has the property that it maps a function to -1 times that function: a different $\mathcal{L}(f(x))$, now, and $\mathcal{L}(f(x)) = -f(x)$. If $\mathcal{L}(f(x))$ is a second differential equation, then it includes second differentiation (but nothing higher than that). In particular, if $\mathcal{L}(f(x)) = \frac{d^2}{dx^2} f(x)$, then $f(x) = sin(x)$ is a solution, and so is $f(x) = cos(x)$ (as is their sum).

Often the big (linear differential) equation we care about might look like:

$$\mathcal{L}(f(x)) = \lambda f(x) \tag{5}$$

and this looks just like the equation for an eigenvector of a linear transformation, so we talk about eigenvectors of linear operators too. (An eigenvector of a linear transformation is a vector whose direction is left unchanged by the linear transformation. Remember that a linear transformation can be fully described by a matrix. Some readers will remember that singular value decomposition shows how linear transformations can be thought of as decomposition into a set of matrices which crucially involve eigenvectors.)

So where we stand now is here: we would like to specify a linear differential operator (and there is no one to tell us what this is! We have to think it up all by ourselves) $\mathcal{L}$, and we would like to think of the underlying linguistic information as *boundary conditions* — these are facts that just have to be true of any solution we care about. In a lot of cases, the boundary condition is that a certain function has to take on the value 0 at the boundaries, the edges, of whatever system we are looking at. These two things, these two pieces of knowledge, allow us to express so-called Green functions.

The Green functions are based on an eigenvector decomposition of the linear differential operator $\mathcal{L}$, and a certain additional condition on $\mathcal{L}$ allows us to be sure that these eigenvectors span the whole space of allowable functions (i.e., the functions we are looking for). This, I think, is the key point to the graphs given in Figure 12.6, on page 266: these functions are orthogonal and form a basis set by which any solution can be expressed as a linear combination.

Khalil and Louis's paper becomes somehow magical on page 270. Here is how I read it, though I might not be getting it right. Because the DCN model restricts its positional boundary conditions to initial and final positions, the solutions to the problem we are looking at will fall within a relatively small subspace of the entire space spanned by the eigenvalues (given in Figure 12.6), and in particular it is the first and last eigenvectors (G1 and G7) that span the space we care about. Different solutions will emerge because α and β are independent parameters that can vary, and Figure 12.9 on page 271 gives a visual representation of the stress patterns that emerge over the different values possible for α and β.

13.2 Digression on eigenvectors and spectral methods

Time for a digression on eigenvectors and the like. Why eigenvectors? Because a relaxation method is an operator looking for an eigenvector with eigenvalue 1: it is looking for a set of values which will remain constant when we appy the same function to it, over and over.

This way of rethinking the problem allows us to see that we have looking at is very similar to what we do when we look for *eigenmaps* on graphs, as Belkin and Niyogi made clear about 20 years ago. A graph has nodes and edges, and a function (or *map*) on a graph is an assignment of real values to each node. This is much like what we have just seen for DCNs, though their graphs are simple (they are just linear). Each graph has a matrix associated with it called its *Laplacian*, and there is also an operator called the *Laplacian* L that can be applied to a function f on a graph. We consider first the case where all of the edges have the same weight, 1.0. We write $x \sim y$ if x and y are nodes connected by an edge.

Let's start with the Laplacian of a *graph*, and then we can ask about using the Laplacian to act as an operator on functions on the graph. The Laplacian is a matrix L, and the values on the diagonal of L are positive: the value of $L(i,i)$ is $deg(i)$, the number of edges that come to x_i. The value of an off-digonal element $L(i,j)$ is -1 if the nodes x_i and x_j share an edge, and 0 otherwise.

This matrix appears in Khalil and Louis's paper in footnote 3!

In most cases of interest, we want to associate a weight with any particular edge, $w(e_{x,y})$. In that case, the Laplacian has the value $L(i,j) = -w(e_{x_i,x_j})$, and the diagonal $L(i,i)$ is the sum of the values of all the edges leading to x_i, which is to say, $\sum_{y \sim x} w(e_{x_i,y_j})$.

Given a function f mapping from the nodes of a graph to the reals, $L(f)$ is a new function on the graph. It is a function that applies to each node x as follows:

$$L(f)(x) = \sum_{x \sim y} (f(x) - f(y)) \tag{6}$$

If each edge $e_{x,y}$ which joins x and y is associated with a weight $w(e_{x,y})$, then the Laplacian reflects this.

$$L(f)(x) = \sum_{x \sim y} (f(x) - f(y))w(e_{x,y}) \tag{7}$$

Now, in the case of DCNs, there is a weight associated with each edge between adjacent segments. If the edge goes to the left, then it links node x_i to node x_{i-1}, and that value is α, fixed for a given language, and if it links node x_i to node x_{i+1}, then that value is β. So the preceding equation can be written for the DCN as:[16]

$$\begin{aligned} L(f)(x_i) &= [f(x) - f(x-1)]\alpha + [f(x) - f(x+1)]\beta \\ &= \end{aligned} \tag{8}$$

(not finished...)

And an eigenmap of the Laplacian operator with an eigenvalue of 1.0 is a steady state of the DCN.

When we study graphs in general, the analysis of the Laplacian operator into eigenmaps tells us a lot about the graph structure (which, in the case of DCNs, is very simple). **End of digression, for now**

What's next?

What's next, I think, may be pushing this model towards being a field model, in the way I alluded to at the begining of my remarks on this paper. This would mean computing a wave – a function – in one dimension (which is time), and the segments would be associated (in some fashion) with subintervals of time on the wave. But they would not necessarily be a segment associated with each peak or trough. As in familiar quantum mechanics, the energy level of the representation would be higher if the frequency were higher (and the wavelength was lower), and energy minimization would push towards a sequence of segments being syllabified in as few waves (i.e., syllables) as possible, just as a sequence of syllables would be footified in as small a number of feet as possible.

14 Bernard Laks, Basilio Calderone, and Chiara Celata

I've learned most of what I know about French phonology from Bernard Laks, and it's a great pleasure to read this paper, one of whose co-authors if Basilio Calderone, with whom I participated in Bernard's seminar in Paris in 2006.

The phenomenon of consonant liaison in French is one of the oldest puzzles in modern phonology. It is closely related to the phenomenon of *enchainement* (which is perhaps among the most universal of phenomena in the world): a word-final consonant syllabifies as the onset of a syllable if the next word begins with a vowel. Liaison involves those cases where a consonant can be identified that does indeed undergo *enchainment* in this sense (such as in the phrase *petit⌢enfant*), but fails to be realized (i.e., fails to be pronounced) if it does not undergo liaison (and it will not undergo liaison if the next word begins with a consonant or an *h-aspiré* (the *t* is not pronounced in *petit garcon*). It is convenient to be able to refer to liaison occurring between Word 1 ($W1$) and Word 2 ($W2$).

The authors point out that there are (at least) three things to bear in mind: (1) how to represent the *process* of *enchainement* (and they suggest that a unitary account of *enchainement* is appropriate for liaison and non-liaison cases); (2) how a liaison consonant is represented, in contrast with the way in which a simple word-final consonant is represented; and (3) what the grammatical context is in which liaison occurs.

The answers to (2) divide into those that (i) have different phonological representations for $W1$ depending on whether $W2$ begins with a vowel (thus permitting *enchainement*); (ii) have a way to indicate in the basic (i.e., "lexical") representation that a word-final consonant is of this special "latent" sort; (iii) include a rule or process of epenthesis. Approach (iii) may be appealing in the case of analyses of dialects whose liaison

[16] We have to define values for -1 and 101 in the natural way, or something equivalent.

phonology does not reflect standard orthography, and has been considered as a possible analysis by which a word-*initial* consonant is epenthesized in cases like *nous⁀avons* or *petits⁀enfants*.

The authors write, "In other words, the proclitics are part of the same construction as the following word and *petits⁀amis* 'small friends' ... are one word each." I'm not sure that the case has been made that they form one word each. I think in particular of the phenomenon in American English whereby the flap is formed. When $W1$ ends in a vowel and $W2$ begins with the morpheme *to*, then the *t* undergoes enchainement (but becomes *ambisyllabic*), as in *go to school, go tomorrow, go tonight.* However, if $W2$ begins with a *t* that is not part of this morpheme, the flap does not occur (*see tomatoes, see Topeka, through topology*). I do not see a reason to say that the form in which the flapping occurs is a single word (though it is true that the phonological structure at the boundary looks like a word-internal sequence—is that all that is needed to argue that a single, larger word occurs? That does not seem right.)

The authors point to previous work of theirs, done in the context of the large, pan-dialectal survey of PFC corpus (Phonologie du francais contemporain), in which they have found two distinct domains in which liaison occurs. One part is constant and stable across time, geography, and social factors, while the other is variable (often very variable) across the same set of factors.

One of the very interesting aspects of the PFC is its rich variation in speaker age. Linguistic differences that appear in speakers of different ages can be interpreted in (at least) two ways: (1) these may be differences that reflect differences among the speech patterns *heard* by late adolescents (the age at which language usage becomes relatively fixed) in different periods of the recent past (thus using present day variation related to age as a tool to observe language in the past: if the present age of someone is A, then their speech may reflect the language they heard A-20 years ago), or (2) it may reflect differences between speakers of different ages in a way that has been constant over time. In this second case, the differences could be implicitly or explicitly tied to "well, that's how old people speak," or it *could* be linked to aspects of language that take longer than 20 years to learn!—if such things exist. If they do, liaison might be one of them.

15 Aris Xanthos

Aris's paper in this volume discusses several recent approaches to the unsupervised learning of Arabic root-and-pattern morphology. In some ways, it reflects the wonderful spirit of the working group we had in 2006, some of which led to Aris's 2008 dissertation on this subject, but the present paper moves forward from that, using again Minimum Description Length (MDL)-inspired learning methods.

One of the things I believed I have learned from ongoing work developing the Linguistica morphological analyzer is that one should not require that a given word be analyzed in only one way. In the earliest work I did, I though that insistence on a unique analysis would allow the learner to begin with the clearest and (in some sense) most convincing analyses, and it could later return to dealing with words which seemed to be uncertain between two different analyses. That was, I think, an error; if a verb like *tip, tips, tipped, tipping* has two forms that are doubly analyzed (*tipped, tipping* are each analyzed with both the stems *tip* and *tipp*), that is fine, because the system can (once it has produced that first pass of hypotheses) look to see if there are patterns of repeated multiple-analyses of the same word, and in the case of English, it will discover that there are a large number of words that appear both with the signature *NULL-ped-ping-s* and the signature ed-ing. It would be very interesting to see if this style of reasoning would help Arabica's analysis in a similar fashion.

Is there a linguistic connection between two things that Arabic (like other similar, unrelated languages like Sierra Miwok) does: (i) allow the vowels to "vary" in morphologically determined ways, and (ii) allow a particular consonant of a consonantal morpheme to appear sometimes in onset position and sometimes in coda position? If both conditions were not met, I think we would not be drawn to view the system as one of root-and-pattern. That the first (i above) is a necessary condition is fairly obvious, I suppose, but I think that the second is a necessary and sufficient condition for identifying the pattern as root-and-pattern. What does this tell us about syllabification? Does it, in particular, give us hints about the way in which segments interact with a syllabification algorithm, perhaps of the sort that Iskarous and Goldstein describe in their paper above?

16 Thank you, again

Thanks to Diane and to Jackson, who poured so much effort into making this book a reality (and I'm so pleased that the University of Chicago Press was eager to see the project succeed). *Ça a de la gueule.* When I look at the topics I have worked on in linguistics, I see three large groups (and one or two small ones): phonology, unsupervised learning, and in the last several years, social epistemology and the history of the mind sciences. In phonology, this has meant developing autosegmental phonology and its application to African (especially Bantu) tone languages, and some efforts to apply insights from neighboring fields to problems of phonology. First and foremost are the dynamic computational models that I worked on with Gary Larson, which constituted an effort to develop a theory of phonology in which representations and rules are integrated into a single formal object, and harmonic phonology, an effort to apply to generative phonology insights developed by early connectionist modelers, including Paul Smolensky. Work on language acquisition using machine learning models struck me in the 1990s as potentially the source of a deep contribution to linguistics, if these models could help us make the problems of language acquisition more concrete. As I write this, *Battle in the Mind Fields* has only been out a few months, but I hope it will have an effect on how linguists look at their own work, and how they see the relationship between what they are doing and what they have learned from the linguists who preceded them. I'm grateful to be told that my work had an impact on the people who contributed to this volume, and I am equally pleased to say that I have learned from them as well.

All science is a great conversation going on every day, every week, every year. We listen, we think, we even pipe up when we think we have something to say. We agree, we disagree, we change our minds. I learned so much from listening to, and then talking with, my teachers in graduate school (thank you, Morris, Noam, Paul, Haj) and then later listening to many others who had left their ideas behind in writing (thank you Zellig, Edward, Chas). This book allows me to see that I've been part of that conversation that is still rolling on, still in business 24 hours a day.