Rethinking Autosegmental Phonology

William R. Leben
Stanford University

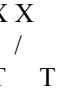
1. Introduction

It is striking how much generative phonology has changed over the past three decades or so. The major departures from the initial landmark of Sound Pattern of English (Chomsky and Halle 1968) have been as numerous as they have been radical. The assumption that the core of the phonological component defines a set of derivations has been all but abandoned. New principles of internal organization, such as phonological strata and "sympathetic" analyses, have been proposed. In addition to entertaining very new views on the organization of a phonological component, current generative phonology has also radically changed its ideas about the representation of basic phonological units, with metrical trees and grids for stress and rhythm and hierarchical representations involving moras, syllables, and feature geometry for other phenomena.

When we look at tone, the two most remarkable facts we notice are how much has been learned about tone systems in the past three decades yet how little the accepted notions of the *representation* of tone have changed. In the representation of tone the last major milestone was Goldsmith 1976. Certainly many small changes have come in since then. A wide variety of new proposals have been made about tone features, about the internal geometry of the TBU, and (probably most frequently of all) about where specific conditions on autosegmental behavior are universal. But most has been cast in autosegmental terms. More radical changes have been proposed, but none has gained general acceptance. This paper seeks to examine whether autosegmental theory of tone has lagged behind the rest of phonology and is in need of an overhaul.

2. Autosegmental phonology: early results

In a nutshell, autosegmental phonology posits representations conforming to the template in (1):

(1)


Tones (T) and tone-bearing units (X) are represented on separate tiers. Tiers are connected by association lines. We can read association lines in (1) as saying that the domain of the tone designated by the T on the left is coextensive with the tone-bearing units (TBUs) on the upper tier. Tones may also float, as does the T on the right in (1). In this case subsequent changes may apply (e.g. docking, stray erasure).

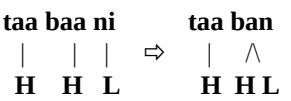
An attractive aspect of autosegmental phonology pointed out by (Poser [1984] 1991) is that it posits a simple, elegant theory of operations. These are restricted to the following:

- add a tone
- add a link
- delete a tone
- delete a link

These simple operations provide an explicit way of framing analyses. For example, note how elegantly the autosegmental formalism captures the Hausa contraction rule in (2):

 Selected Proceedings of the 35th Annual Conference on African Linguistics, ed. John Mugane et al., 1-9. Somerville, MA: Cascadilla Proceedings Project.

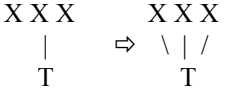
(2) Hausa contraction



The rule neither affects tone nor is conditioned by tone. It simply states that pronouns such as first person singular *ni* lose their vowel (in specific cases and not generally). Note in (2) that when the vowel is deleted, its tone remains. Due to a general principle of Hausa, this tone is unable to remain as an unassociated or floating tone, and so it must link up to some TBU. Due to the general constraints of autosegmental phonology, the only available candidate is the TBU of *ban*.

Tone Spreading (a term used by Hyman and Schuh 1974 that went on to receive wide application) also illustrates the autosegmental knack for capturing generalizations graphically: Here is a typical spreading rule, which extends a tone rightward from one TBU to include all of the remaining TBUs in its domain:

(3) Spreading



What is particularly nice is that this spreading rule is not only simple and highly graphic, but it is widely attested in languages. Below is one example. Of course the literature on tone is filled with these.

(4) Spreading in Mende (SW Mande, Sierra Leone)

a. **kɔ** ‘war’ **hu** ‘in’ **kɔ-hu** ‘in war’
H **H**

b. **mba** ‘rice’ **hu** ‘in’ **mba-hu** ‘in rice’
LH **L H**

Another early result that helped to seal the reputation of the autosegmental approach was the ability to capture a connection between downstep and Low tone by delinking the L yet retaining it in the tonal tier so that it would be present to cause a following H to downdrift.

(5) a. **hɛn** ‘their' **adan** ‘house’
H **L H**

b. **hɛn adan**
H L H = **hɛ́n ↓ádán** ‘their house’

In (5b), the L is not associated but its effect is heard as downstep on the following word. This downstep is symbolized as ↓ in the IPA representation on the right.

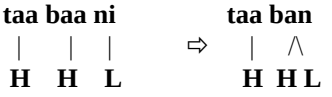
This is not to say that *all* downstep necessarily arises from floating tones. Hyman 1985 and others have brought up phenomena such as "double downstep" where floating tones do not suffice as sources of downstep.

3. Theory or descriptive framework?

Autosegmental phonology not only provides a simple, graphic way of capturing generalizations but also constitutes a theory of how phenomena represented autosegmentally ought to behave. For example, autosegmental tone ought to display frequent independence from tone-bearing units. Such independence can manifest itself as tonal stability, as Goldsmith 1978 pointed out, referring to an observation by Paul Kiparsky. The Akan example (5) above illustrates this perfectly, since a L tone of the prefix survives even when the segmental portion of the prefix assimilates to the following H.

Another familiar example comes from an optional contraction in Hausa, which can delete the vowel of certain pronoun objects, leaving their tone..

(6) Tonal stability (cf. retained L in (2) Hausa contraction):



That is not to say that tone will *always* behave independently of the TBU to which it is attached. The autosegmental formalism is rich enough to allow a TBU and tone combination to be inserted or deleted together or to condition some other change. For example,

Another clear prediction was first brought up by Odden 1986. (7a) shows that words in KiShambaa, a Bantu language of Tanzania, have two kinds of High-toned sequence. In the form on the left, this H pattern surfaces as level H; the form on the right, by contrast, surfaces with a downstepping pattern.

(7) KiShambaa (Odden, 1986)

a. *Underlying forms:*

/ **nyoka** / 'snake' *vs.* / **ngoto** / 'sheep'

H **H H**

Surface forms:

nyóká **ngó↓tó**

[¯ ¯] [¯ —]

b. *Tonal foot boundaries:*

/ **nyoka** / 'snake' / **ngoto** / 'sheep'

(H) **(H)(H)**

From the contrasting tone melodies H and HH in (7a), Odden concludes that the Obligatory Contour Principle (OCP), which reduces both H and HH to the same melody, is not applicable universally.

But suppose that the OCP is universally valid across a tonal foot and that in some languages like KiShambaa, words may consist of more than one tonal foot. The form on the left in (7) would be analyzed as having a single tonal foot while the form on the right would be parsed into two tonal feet, as shown in (7b). Odden 1995:456 in fact notes, in a different connection, that the environment of KiShambaa downstep insertion could be expressed as being across a boundary between tonal feet. If

this is so, what actually varies from language to language may simply be the number of tonal feet in a word; the OCP itself could be universal, but only within the confines of the tonal foot.

However this case is analyzed, Odden's example clearly illustrates a prediction made by autosegmental phonology.

4. A basis for a variety of theories

The predictions about tone in section 3 are impressive enough and accurate enough. Still, rather than being tightly tailored to fit the demands of tone, autosegmental phonology is a rather loose-fitting garment. To a great extent, it has simply provided a descriptive framework in which a variety of theories can be couched--for example, the theories of Pulleyblank 1986, Yip 1989, de Lacy 2002, Zoll 2003, along with many others.

That is why it was unsurprising to find that association lines and placing phonological features on distinct tiers helped to describe assimilations involving segmental features like voicing, nasality, and so on. With appropriate assumptions about representations of segments, the behavior of these sorts of assimilations was easily handled under autosegmental assumptions. But if segments and tone both yield to autosegmental treatment, this hardly indicates a deep similarity between them. Rather it indicates that the autosegmental framework isn't terribly demanding.

Among recent studies seeking to revise autosegmental phonology more fundamentally, one stands out, in my opinion, for successfully tailoring the theory to the demands of tone in a specific and significant group of languages. This is the Optimal Domains Theory of Cassimjee and Kisseberth in a large body of work. The papers consulted here and listed in the references will be referred to as C&K.

5. Optimality and Optimal Domains

Optimal Domains Theory, (ODT) is an optimality-based alternative to autosegmental representations that accounts for widely attested generalizations in Bantu that have eluded autosegmental phonology. The example in (8) from C&K 2001: 333-337 shows how a set of constraints describes the domain of tone spreading in a number of different Bantu languages. These constraints apply to what C&K call H domains, a sequence of moras, the leftmost of which is H in underlying representation and the rightmost of which is H on the surface.

(8) Constraints on H domains
   a. INITIAL H: H domain is initiated by a mora that is H in the input.
   b. NON-FINALITY: H domain does not include a word-final syllable.
   c. MINIMALITY: H domain is at least two syllables long.
   d. AVOID PROMINENCE: H domain does not include a prominent syllable.

As in other OT approaches, a constraint is violable only if some other constraint takes precedence.

These constraints account for the complex pattern of tone spreading in Swati in (9), as Cassimjee & Kisseberth 2001:334, 337 show. First a few points on notation:

- a H domain is enclosed in brackets [ ]
- within a H domain, the underlined vowel, as in *b<u>a</u>,* indicates a TBU that is underlyingly H. A vowel with an acute accent, yá, indicates a TBU that is H on the surface
- the penultimate vowel is followed by a colon and is prominent

(9) a. **[b<u>a</u>-yá]na:tsa** '[cl.2] are drinking'
[b<u>a</u>-ya-khilikí]dze:la '[cl.2] are stumbling'

b. **ngi-ya:-[ph<u>á</u>]** 'I am giving'

c. **u-ya-[b<u>ó</u>:]na** 'you see'

d. **si-ya-[s<u>e</u>b<u>é</u>:]nta** 'we are working'

The forms in the first set, (9a), are consistent with *all* of the constraints (8). At the left edge of the H domain is a mora *ba* that is underlyingly H (8a). The right edge is not final (8b). The H domain is two syllables long in the first example in (9a) and five syllables long in the second, satisfying (8c). And the H domain does not include a prominent syllable--here, a syllable with a long vowel (8d).

(9b) violates constraint (8b) because the H domain is at the right edge. But any form with a monomoraic H verb stem will behave in this way, so it would be wrong to treat this behavior as exceptional. This case simply shows that INITIAL H (constraint (8a) takes priority over NONFINALITY (constraint (8b). Similarly, the H domain in (9b) has only one syllable, in violation of MINIMALITY (constraint (8c). Once again, this is typical of monosyllabic verb stems, showing that INITIAL H always takes precedence over MINIMALITY.

(9c) violates MINIMALITY and it contains a long vowel, in violation of AVOID PROMINENCE (constraint (8d). These two constraints therefore must also yield to INITIAL H (constraint (8a).

PROMINENCE. The expression si-ya-[sɛ́bɛ́:]nta obeys the first yet violates the second. To conclude this example, note that all of the violations noted for (13a,b,c,d) follow from a single ranking of the constraints, and that ranking corresponds to the order in which they are given in (8).

6. Optimal Domains Theory: an alternative to autosegmental representations

Obviously the use of ranked constraints plays a crucial role in the above result. But another key element in the success is C&K's notion *H domain*. Compare the schematic autosegmental and ODT representations in (10). C&K informally represent surface H with an acute accent as we see in (10b) with *x́*. The question of the *formal* representation of H tones remains open, and an autosegment remains a possibility.

(10) a. Autosegmental representation

X X X X X X X X
\ | / \ /
H H

b. ODT representation

x (x̲ x́ x́) x (x̲ x́) x

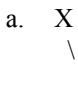
The two representations convey some of the same information. Both have a span of surface H tones running from the second to the fourth TBU and another running from the fifth to the sixth. But the ODT counterpart also reveals at least two other kinds of information--where the *underlying* H tones are and what kinds of *units* the tones are grouping into.

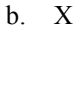
As noted above, ODT representations are tailored to observed tonal behavior, most notably in Bantu, as well as to the demands of OT. In (10b), (x̲ x x) and (x̲ x) represent H domains. The underlined character x̲ is a TBU that is H in the input. The right-hand domain boundary is set at the rightmost TBU that receives a H by rightward spreading. What claim does the representation (10b) make? It expresses a commonality between two languages, language L_1 which H spreads from TBU_i to TBU_j, making these plus all intervening TBUs H, and language L_2, where H also spreads from TBU_i to TBU_j, but only TBU_j surfaces as H. The commonality, of course, is that in both languages the H domain is the same, even though the surface tones are different.

Now compare the two representations in (11):

(11) a. (x́ x́ x́)
b. (x x́ x́)

Both are H domains containing three TBUs with an initial H TBU in the input, but they differ in which TBUs are H on the surface. How would autosegmental representation capture (11)? It could not capture the difference between an underlying and surface H, except with rules to relate the underlying and surface representations. Furthermore, it could not capture the fact that (11a) and (11b) are closely related. The surface autosegmental representations of (11) are in (12):

(12) a.


b.


The representation in (12a) does not share a left boundary with the one in (12b). To at least this extent, there is no sense in which autosegmental phonology says that (12a) and (12b) are related, while ODT says that (15a) and (15b) *are* related. Which statement is correct? The following comparison, comprising examples from C&K 2001, shows how closely the H domains are related in three Nguni languages. C&K actually note that H domains are constructed identically across all of Nguni, and their body of work demonstrates that the basic principles for constructing H domains have wide applicability in Bantu:

(13)

Swati	Durban Zulu (L. Downing, cited in C&K 2001)	Ngoni (Donnelly 1997)
a. **[ba-yá]na:tsa** '[cl.2] are drinking'	**[ú-kú]-cu-la** 'to sing'	**[kú-límí]si:sa** 'to cultivate intensively'
b. **[ba-ya-khilikí]dze:la** '[cl.2] are stumbling'	**[bá-yá-límá]li:ma** '[cl.2] are cultivating' (rep.)	**[bá-yá-límé]la:ma** '[cl. 2] cultivate for one another'
c. **si-ya-[sebé:]nta** 'we are working'	**si-ya-[lalé:]la** 'we listen'	**si-ya-[bóní:]sa**

(13) indicates that the H domains in all three languages are constructed in accordance with the constraints in (8)--this despite differences in whether on the surface the underlying H shifts or spreads or both. The differences are summarized in (14).

(14) a. In Swati, input H *shifts* to the final syllable of the H domain.
b. In Durban Zulu, input H *spreads* from prefixes to the final syllable of the H domain and *shifts* from input stems to the final syllable of the H domain.
c. In Ngoni, input H *spreads* to the final syllable of the H domain.

The fact that the initial and final locations of Nguni H tones are all determined by the principles in (8), despite the surface differences in (14) suggests that H domains are an important aspect of Nguni tonology. This apparently falsifies one of the (few) tenets of autosegmental phonology, but below I offer a reinterpretation that is consistent with autosegmental phonology.

7. What is a H domain?

Is a H domain a new kind of tonal constituent, unrelated to any other unit, or is it an element of a larger pattern? A first step toward an answer would be to consider other types of phenomena that H domains resemble. One such phenomenon is (12a), the multiply linked H tone. In fact, past autosegmental studies analyzing tone shifting derivationally as spreading followed by delinking normally regarded multiply linked tones as stage (15b) in the process leading from (15a) to (15c):

(15) a. X X X b. X X X c X X X
\ \ | / /
T T T

Another frequent case is the grouping of individual *lexical* H tones into a H domain that is *intonationally* unitary. In Hausa, an unbroken string of H tones in a phonological phrase combine into a H domain that is intonationally unitary. Any H tone can be "emphasized" by raising it to extra-H. But in that case, every following H in the phonological phrase is raised as well. In other words, even if a following word is not emphasized, its H tones will be raised, but only so long as no L tones intervene. This was described by Inkelas and Leben 1990.

(16a) shows a Hausa sentence with normal intonation and no special emphasis. Lexical Low tones are marked with a grave accent and appear only on the first two syllables. These are followed by a string of lexical High tones, which are transcribed with no accent mark. (16b) shows the same sentence, but with emphasis added to the H-toned word *neemoo*. The emphasis causes the H tones of *neemoo* to raise to extra-High. Even though the words following are *not* to be interpreted as emphasized, they are nonetheless raised to extra-High as well. This would not happen unless the H tones of the entire sequence formed (in some sense) a single domain.

(16) *Normal:*

a. Bà mù neemoo awarwaron Maanii ba

'We didn't look for Maanii's bracelet'

Emphatic:

b. Bà mù *neemoo* awarwaron Maanii ba

'We didn't *look for* Maanii's bracelet' (*neemoo* = 'look for')

In Baule, an unbroken string of H tones in a phonological phrase combine into a H domain that is intonationally unitary. We know that it is unitary because it is within this domain that we find "upsweep," ((Ahoua 1996, Dafydd Gibbon p.c.), a gradual rise in the pitch of H from the first H to the last H in the domain. Here is an example from Leben and Ahoua 1996:

(17) Ákíssí Bólí (proper name)
[– – – ¯ ¯]

In light of these examples, let us re-examine the Nguni cases described earlier for alternatives to C&K's conclusions. As with the Hausa and Baule examples, the Nguni cases involved a sequence of TBUs beginning with an input H on the left and ending with an output H (or, depending on the language, the last output H of a sequence of output H's). What the Hausa, Baule, and Ngoni cases all have in common, surprisingly, is is an intonationally unitary domain with left and right edges defined by H tones. The difference lies in how the tones in the H domains are interpreted phonetically. In Hausa they are High in normal intonation and extra-High in emphatic intonation. In Baule, they rise from H at the beginning to extra-High at the end. In Nguni, the interpretation varies from language to language: Swati ((14), col. 1) has a rise from L to H across this domain. Ngoni ((14), col. 3) has a level H across the same domain.

C&K find it necessary to regard the H tone of the left edge of the Nguni H domains as underlying, because in some Nguni languages the C&K surface tonal transcriptions give these as L. But perhaps there is a different interpretation. Suppose that in Nguni the theory of tone forced us to describe a H domain as being constructed of *surface* High tones. Hausa and Baule suggest a way to make this work for Nguni. Recall that a Hausa surface H-toned sequence has the intonational realization High (normal intonation) or extra-High (emphatic intonation. In Baule, the corresponding H domain is realized as an intonational rise.[1]

Let us consider an alternative for Nguni wherein spreading and shifting are not phonological rules at all, and where the role of the phonology is simply to assign domain boundaries according the constraints in (8). On this analysis, the surface forms of Nguni would contain H domains as given by brackets in (13), yet these domains would not be marked for surface tones at all! Instead, the tones would be assigned by intonational rules. As we have seen, Ngoni interprets a H domain intonationally as a level H. For other Nguni languages (Swati, for instance), the only difference would be that a High domain is interpreted intonationally as a rise, beginning with Low and ending with High.

This proposal is summarized in (18), where each H domain is given in brackets and a line above indicates the approximate pitch curve:

(18) a. Hausa normal intonation: a H domain is High-pitched:

xx[HHH]xx

b. Hausa emphatic intonation: a H domain is extra-H-pitched:

xx[HHH]xx

c. Baule intonation: a H domain rises in pitch from Low to High:

xx[HHH]xx

d. Ngoni intonation: a H domain is High-pitched:

xx[HXX]xx

e. Swati intonation: a H domain rises in pitch from Low to High:

xx[HXX]xx

Under this interpretation, ODT would still represent an important innovation over autosegmental phonology, and the innovation would still involve the way that tonal domains are defined. As C&K have shown, the position of the right bracket in (18d,e) is established *not* by where the rightmost H tone is but by an independent set of constraints, (8). By contrast, what autosegmental phonology provides, as seen in (12a), is a tonal domain (if it is fair to call it that) comprising the sequence of TBUs linked to a given *tone*. What is new under this reinterpretation is that a H domain need not simultaneously refer to underlying and surface tones in order to capture the facts in (13). Whether this reinterpretation holds for other Nguni languages or other Nguni phenomena is something that merits further investigation.

[1] This rise is described variously as going from Low to High (Ahoua 1996) or from mid-High or High to extra-High (Leben & Ahoua 1997).

References

Ahoua, Firmin. 1996. Prosodic Aspects of Baule, with special reference to the German of Baule speakers. Cologne: Koeppe.

Akinlabi, Akinbiyi, and Mark Liberman. 2001. Tonal complexes and tonal alignment. NELS 31.

Cassimjee, Farida. 1998. Isikhosa Tonology: An Optimal Domains Theory Analysis. Munich: Lincom Europa.

Cassimjee, Farida, and Charles W. Kisseberth. 1998. Optimal Domains Theory and Bantu Tonology: A Case study from Isixhosa and Shingazidja. In L. M. Hyman and C. W. Kisseberth, eds., Theoretical Aspects of Bantu Tone. Stanford: CSLI Publications, 33-132.

Cassimjee, Farida, and Charles W. Kisseberth. 1999. Tonal variation across Emakhuwa dialects. In S. Kaji, ed., Proceedings of the Symposium, Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Typology, and Related Topics. Tokyo: Tokyo University of Foreign Studies,261-287.

Cassimjee, Farida, and Charles W. Kisseberth. 2001. Zulu tonology and its relation to other Nguni languages. In S. Kaji, ed., Proceedings of the Symposium, Cross-Linguistic Studies of Tonal Phenomena, Tonogenesis, Japanese Accentology, and Other Topics. Tokyo: Tokyo University of Foreign Studies, 327-359.

Chomsky, Noam, and Morris Halle. 1968. The Sound Pattern of English. New York: Harper & Row.

de Lacy, Paul. 2002. The interaction of tone and stress in Optimality Theory. Phonology 19.1-32.

Donnelly, Simon. 1997. Aspects of Tone and Voice in Phuthi. Ms. of U. of Illinois Ph.D. thesis.

Goldsmith, John. 1976. Autosegmental Phonology. MIT dissertation.

Goldsmith, John. 1987. Prosodic trends in the Bantu languages. In H. van der Hulst and N. Smith, eds., Autosegmental Studies on Pitch Accent. Dordrecht: Foris, 81-93.

Hyman, Larry M. 1985. Word domains and downstep in Bamileke-Dschang. Phonology Yearbook 2.47-83.

Hyman, Larry M., and R. G. Schuh. 1974. Universals of tone rules: Evidence from West Africa. Linguistic Inquiry 5.81-115.

Inkelas, Sharon, and William R. Leben. 1990. “Where phonology and phonetics intersect: the case of Hausa intonation.” In Proceedings of the First Conference on Laboratory Phonology, ed. by M. Beckman and J. Kingston. Cambridge University Press, pp. 17-34.

Leben, William R., and Firmin Ahoua. 1997. Prosodic domains in Baule. Phonology 14.113-122.

Odden, David. 1986. On the obligatory contour principle. Language 62.353-383.

Odden, David. 1995. Tone: African languages. In J. A. Goldsmith, ed., The Handbook of Phonological Theory. Cambridge: Blackwell, 444-475.

Poser, William J. 2001. Phonetics and Phonology of Tone and Intonation in Japanese. Stanford: CSLI Publications.

Pulleyblank, Douglas. 1986. Tone in Lexical Phonology. Dordrecht: Reidel.

Yip, Moira. 1989. Contour tones. Phonology 6: 149-174.

Zoll, Cheryl. 2003. Optimal tone mapping. Linguistic Inquiry 34.225-268.

Selected Proceedings of the 35th Annual Conference on African Linguistics: African Languages and Linguistics in Broad Perspectives

edited by John Mugane,
John P. Hutchison, and Dee A. Worman

Cascadilla Proceedings Project Somerville, MA 2006

ISBN 1-57473-410-5 library binding

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Leben, William R. 2006. Rethinking Autosegmental Phonology. In *Selected Proceedings of the 35th Annual Conference on African Linguistics*, ed. John Mugane et al., 1-9. Somerville, MA: Cascadilla Proceedings Project.

or:

Leben, William R. 2006. Rethinking Autosegmental Phonology. In *Selected Proceedings of the 35th Annual Conference on African Linguistics*, ed. John Mugane et al., 1-9. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1291.