Fiengo, R. (1977) "On Trace Theory," *Linguistic Inquiry* 8, 35–61.
Jaeggli, O. (1980) "Remarks on *To*-Contraction," *Linguistic Inquiry* 11, 239–245.
Pullum, G. K. (1979) "The Nonexistence of the Trace-Binding Algorithm," *Linguistic Inquiry* 10, 356–362.

The Generative Capacity of Word-Formation Components

D. Terence Langendoen, Brooklyn College and CUNY Graduate Center

Questions of generative capacity have arisen so far only for the class of grammars of natural languages as a whole. However, if a particular component of a grammar is viewed as a grammar in its own right (so that a grammar as a whole is viewed as a system of component grammars), questions of generative capacity also arise for the classes of the various components of grammars. In particular, we may ask what the weak generative capacity of the class of word-formation components of grammars of natural languages is.

Clearly it must be at least finite-state, to accommodate the word-formation component of English, which generates such infinite sets of words as (1).

(1) $\{(\text{great})^n\text{grandparent} : n \geqq 0\}$

Need it be larger? I know of no attested natural language the word-formation component of whose grammar must be more powerful,[1] but it is easy to make up languages whose words

[1] Bar-Hillel and Shamir (1960) point out that the compounding process in English that gives rise to (i) extends to form (ii), and in general to form the infinite set of words (iii).

(i) antimissile missile 'missile designed to intercept and destroy another missile'
(ii) antiantimissile missile missile 'missile designed to intercept and destroy an antimissile missile in flight'
(iii) $\{(\text{anti})^n\text{missile (missile)}^n$ 'missile designed to intercept and destroy an $(\text{anti})^{n-1}\text{missile (missile)}^{n-1}$ in flight' $: n \geqq 1\}$

If so, and if none of the expressions of the form (iv) are well-formed words in English, then the set of English words is not a finite-state language.

(iv) $\{(\text{anti})^m\text{missile (missile)}^n : m \neq n\}$

However, I would argue that the elements of (iv) *are* English words. Most native speakers of English I have questioned agree with my judgment that (v)–(vii) are well-formed on the interpretations indicated.

(v) antimissile 'directed against enemy missiles'
(vi) antiantimissile 'opposed to that which is directed against enemy missiles'
(vii) antimissile missile missile 'missile designed to intercept and destroy an antimissile missile in flight'

More complex elements of (iv) would be interpreted along the same lines.

Hence, (iii) fails to establish that the vocabulary of English is not a finite-state language.

cannot be generated by finite-state or even context-free word-formation components. Consider first the hypothetical language Inglish in which agentive nouns are productively formed from verbs by the suffixation of *-er*, and in which verbs are productively formed from nouns by the prefixation of *en-*, with the meaning 'act like . . .'. For example, given the noun and verb stem *fish* 'fish', we have the words *fisher* 'one who fishes', *enfish* 'act like a fish', and *enfisher* 'act like one who fishes; one who acts like a fish'. More generally, the result of intersecting the vocabulary of Inglish with the regular set (2) results in the set (3), which cannot be generated by a finite-state word-formation component.

(2) $\{(\text{en})^m\text{fish}(\text{er})^n : m, n \geqq 0\}$

(3) $\{(\text{en})^m\text{fish}(\text{er})^n : m, n \geqq 0 \text{ and } |m - n| \leq 1\}$

Hence, the vocabulary of Inglish cannot be generated by a finite-state word-formation component.

Next consider the hypothetical language Anglish, which is just like Inglish, except that nouns and verbs with the meanings 'a lot of . . .' and '. . . a lot' are formed by reduplicating the noun and verb stems. For example, Anglish contains the words *fishfish* 'a lot of fish; fish a lot', *enfishenfish* 'act a lot like a fish', *fisherfisher* 'a lot of ones who fish', and *enfisher-enfisher* 'a lot of ones who act like a fish; act a lot like one who fishes'. However, no reduplicated form enters into other word-formation processes; the Anglish vocabulary contains neither **fishfisher* 'one who fishes a lot', **enfishfish* 'act like a lot of fish', **enfisherfisher* 'act like a lot of ones who fish', nor **en-fishfisher* 'ones who act like a lot of fish; act like one who fishes a lot'. The result of intersecting the vocabulary of Anglish with the regular set (4) results in the set (5), which cannot be generated by a context-free word-formation component.

(4) $\{(\text{en})^m\text{fish}(\text{er})^n(\text{en})^p\text{fish}(\text{er})^q : m, n, p, q \geqq 0\}$

(5) $\{(\text{en})^m\text{fish}(\text{er})^n(\text{en})^m\text{fish}(\text{er})^n : m, n \geqq 0 \text{ and } |m - n| \leq 1\}$

The word-formation processes used to form the vocabularies of Inglish and Anglish—suffixation, prefixation, and stem reduplication—are widely found in attested natural languages. Are there any attested languages in which they interact as they do in Inglish or in Anglish? If not, is their absence accidental, or is it a consequence of some yet-to-be-formulated principles of word formation that limit the weak generative capacity of the word-formation components of the grammars of natural languages to that of the class of finite-state or context-free grammars?

Reference

Bar-Hillel, Y. and E. Shamir (1960) "Finite-state Languages: Formal Representations and Adequacy Problems," *Bulletin of the Research Council of Israel* 8F:155-166. Reprinted in Y. Bar-Hillel (1964) *Language and Information*, Addison-Wesley, Reading, Massachusetts.

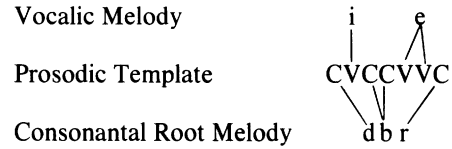
The Representation of Consonant Length in Hebrew

John J. McCarthy, University of Texas at Austin

Leben (1980) presents a stimulating contribution to the problem of segmental quantity cast in terms of a metrical theory of syllabification, which I will not summarize here. Justification for this theory comes from consonant quantity paradoxes—instances of apparently contradictory treatment of long consonants as single segments and as clusters—in Hausa and Biblical Hebrew.[1] In this squib, I will offer an alternative theory of quantity which, while incorporating many of Leben's insights, constitutes just one aspect of the broader prosodic theory of morphology of McCarthy (1979b; to appear). An analysis of the Biblical Hebrew data (hereafter, Tiberian Hebrew or TH) illustrates this proposal. A treatment of the Hausa facts under the same theory can be found in Halle and Vergnaud (1980; in preparation).

The basic properties of the prosodic theory as applied to the characteristic morphology of the Semitic languages are as follows. Verbal and nominal morphological categories stipulate prosodic templates, composed of the archisegments C and V, which give the canonical pattern of each form. Melodies on separate autosegmental tiers make up the consonantal roots and affixes and the vowel patterns that are typical of a language like TH. The regular universal autosegmental conventions for association of Clements and Ford (1979), augmented by a few language-particular rules, effect a mapping between consonantal melodies and C positions of the template and between vocalic melodies and V positions of the template. Under this theory, the representation of a TH form like *dibbēr* 'he said' will be roughly as given in (1):

(1)



The major problem in TH consonant quantity concerns the proper formulation of a rule regularly spirantizing oral stops

[1] Leben's (1980) Hebrew evidence is drawn from Sampson (1973) and Barkaï (1974).