Finite-state phonology predicts a typological gap in cyclic stress assignment

Bronwyn Bjorkman

Queen's University

Ewan Dunbar

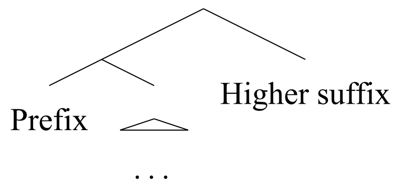
Laboratoire de Sciences Cognitives et Psycholinguistique, Ecole Normale Supérieure–PSL Research University / EHESS / CNRS

In this squib, we start from the empirical generalization that phonological grammars are all within the computational power of finite-state devices, a position nearly universally accepted in computational phonology. We show that this rules out certain patterns of morphologically sensitive stress assignment that are predicted by the *phonological cycle* hypothesis.

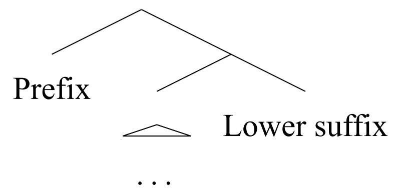
Given that morphological structure is hierarchical in complex words, the phonological cycle proposes that the phonological grammar reapplies successively to each morphological sub-constituent, starting with the smallest (Chomsky et al. 1956, Chomsky and Halle 1968). Variants on this basic idea have been widely influential in phonological theory, including versions with a Strict Cycle Condition (Kean, 1974; Mascaró, 1976), versions that group constituents into strata (Kiparsky, 2000; Bermúdez-Otero, 2011), and versions that make reference to phases (Marvin, 2002; Piggott and Newell, 2005). We claim that, while phonology is sensitive to some aspects of morphological structure, the true constituency is not used, as would be predicted by the phonological cycle: the structure seen by phonology is *non-isomorphic* to the full morphological structure in the sense of Scheer 2010.

Cyclic phonological theories predict that interactions at a distance between prefixes and suffixes can be sensitive to their relative height, as in (1).

(1) a. **Trigger**



b. **Non-trigger**



Cyclic reapplication allows a suffix to have a surface effect on a prefix only when it is introduced on later cycle B, not on cycle A, and thus, only when it is higher. Stress assignment in Chamorro (Chung 1983) has been described in precisely these terms, apparently requiring direct phonological reference to the morphological scope of suffixes with respect to stressed prefixes. However, Chamorro does not require any phonological *comparison* between prefixes and suffixes. Because of this, it can be described without referring to actual constituents, using only limited, local structural information. This puts it within the bounds of finite state computation.

This predicts a typological gap: patterns that, like Chamorro, are sensitive to relative structural height of a prefix and a suffix, are never sensitive to the phonological content of both; while patterns sensitive to the phonological content of both a prefix and a suffix (illustrated by stress assignment in Nez Perce) are never sensitive to their relative height.

1 The stress sytems of Chamorro and Nez Perce

We compare two stress patterns where prefixes and suffixes interact. In both cases this interaction arises as a result of lexical accent (lexical marking of particular items for stress). The first is Chamorro, analyzed cyclically by Chung (1983). In Chamorro, whether an accented prefix gets primary stress depends on its height relative to any suffixes.

Stress in Chamorro is generally penultimate, but can surface elsewhere in the presence of lexical accent. Certain prefixes are lexically accented, and these attract stress leftwards.[1]

(2) a. mantíka
/mantika/
'fat'

b. mímantika
/mí-mantika/
'abounding in fat'

This stress attraction is disrupted, however, by any suffix that is structurally higher than the marked prefix: in such cases, stress is again penultimate.

(3) a. mimantikáña
[[mí mantika] ña]
'more abounding in fat'

b. ákwentusi
[á [kwentus i]]
'to speak to one another'

This pattern suggests that the phonology is sensitive to morphological information. Chung claims that this sensitivity is due to the phonological cycle: stress is recalculated with the addition of each affix: only when a prefix is higher than any suffix can its accent escape re-assertion of the regular penultimate stress pattern.[2]

Notice, however, that Chamorro stress does not require any direct comparison between lexically accented prefixes and the suffixes that outscope them: the phonology needs to know that some suffix outscopes an accented prefix, not which suffix does so.

By contrast, stress in Nez Perce is sensitive to the phonological properties of both prefixes and suffixes, but is insensitive to their relative scope. As in Chamorro, stress in Nez

Perce is generally penultimate, but can be attracted by lexically accented morphemes, including prefixes. Default penultimate stress is illustrated in (4):[3]

(4) a. pískis b. pìskísne
/piskis/ /piskis-ne/
door door-obj

c. hàníisa d. hànisáaqa
/hanii-see/ /hanii-seeqa/
make-incomplete make-recent

Both roots and affixes can bear lexical accent. When multiple accents occur in a word, the default pattern is for the rightmost non-final accent to surface with primary stress.

(5) pàaynóosàqa
/páay-núu-seeqa/
arrive-toward-RECENT (Crook 1999: 458)

There is a subpattern for verbal prefixes with lexical accent, where stress appears on the leftmost lexically accented prefix (in the absence of any accented suffix).[4]

(6) a. cúukwece
/cúukwe-cee/
know-INCOMPLETE
'I know'

b. néesepèsilèwcùkwece
/nées-sepée-siléew-cúukwe-cee/
PL.OBJ-CAUS-by.seeing-know-INCOMPLETE
'I make you know by seeing' (Crook 1999: 465)

Lexically accented suffixes disrupt this leftwards stress shift, but do *not* need to scope higher to do so. In the examples below, the accented directional suffix *(n)úu* 'toward' transitivizes an intransitive root. It thus takes semantic (and, we assume, morphosyntactic) scope below both the causative prefix *sepée* and the plural object agreement prefix *nées*.[5] Though these prefixes consistently attract stress away from accented roots, as seen in (6), the accented suffix retains stress in the examples in (7).[6]

(7) a. hìnàspàynóosa
/hii-nées-páay-núu-see/
3-plob-arrive-toward-inc
'He arrives to them'

b. hìnàsàpapàynóoca
/hii-nées-sepée-páay-núu-cee/
3-plob-caus-arrive-toward-inc
'He makes them arrive to him'

What is important about these cases is that the suffix that attracts stress rightwards is hierarchically "inside" the accented prefix. This suggests that the pattern is not sensitive to the relative scope of the prefix and suffix. The bracketing is illustrated in (8).

(8) a. [nées [sepée [[páay] núu]]]

b. not *[[nées [sepée [páay]]] núu]

In Nez Perce, accented prefixes compete with accented suffixes: the rule is that (non-final) suffixes win. The suffix need not be higher than the prefix. The pattern is not sensitive to relative height, but it is sensitive to the phonological content of both (i.e. to lexical accent).[7,8]

2 A typological gap: simultaneous height and content

We know of no stress sytem that combines the morphological sensitivities of Chamorro and Nez Perce. We propose that this is a principled gap, and that it is impossible for a language to compare the relative height of prefixes and suffixes while also comparing phonological properties of those affixes. This means there can be no stress pattern like (9).

(9) *Winner by Height, If Special:*

a. [ti [[[tú be] tá] to]] → titubetáto

b. [ti [[rú [be tá]] to]] → tirúbetato

c. [ti [[[tú be] sa] to]] → titúbesato

In the *Winner by Height, If Special* (WHIS) pattern, two lexical accents compete to bear primary stress, one on a prefix, and the other on a suffix. As in Nez Perce, if one of the two is not lexically accented, then the other will win. As in Chamorro, the winner is the one that is structurally outermost (unlike in Chamorro, only when both are accented).

WHIS is a prediction of the phonological cycle, and its apparent nonexistence is therefore significant. In what follows, we suggest that the nonexistence of WHIS lends support to proposals that phonology is limited to (a subset of) finite-state computations.

3 Phonology is finite-state

A finite state automaton is a machine that reads a string and uses its contents to determine a path through a finite number of states. When the end of the string is reached, it halts, and either accepts the string, if it is in one of a special set of accepting states, or rejects it, if not. In this way, finite state devices define sets of strings.

In the earliest days of generative grammar, syntax was shown to go beyond finite state computations (Chomsky 1957). Chomsky's argument against finite state syntactic theories was that they would fail to generate unbounded center embeddings in relative clauses: *the* *man* [*the dog bit*] *ran*, *the man* [*the dog* [*the cat was afraid of*] *bit*] *ran*, …. Because a finite state device cannot pair up non-adjacent subjects and verbs, it cannot describe this type of dependency. The same argument applies to the word-internal hierarchical structures of morphology, particularly in polysynthetic and agglutinative languages.[9]

In contrast to syntax and morphology, phonological patterns do not show any evidence of going beyond finite state power. All phonological theories on the market yield grammars that can be translated into finite state devices, once we restrict our attention to the parts of the theories that are actually used in analyses of natural language (Johnson 1972, Kaplan and Kay 1994, Frank and Satta 1998, Graf 2010). Indeed, this difference has been used to argue that phonology and syntax are distinct modules (Heinz 2011, Berwick et al. 2011).

In phonology we are interested in mappings from possible underlying forms, including relevant morphological information, to surface strings (by "mapping" we mean an entire set of legitimate pairs, not just one). The theory of phonology should predict which mappings have possible grammars (computations), and which do not.

One way to use finite state automata to deal with mappings is to see pairs of strings as strings of pairs, as in (10). To encode morphological information, we can add elements to the string that will eventually get deleted, for example open and close brackets (the [and] symbols), as in (10). This enrichment is necessary in order to move from computations that just deal with surface properties, like "stress the penultimate vowel," to ones that deal with underlying properties as well, like the the pattern seen in Nez Perce: "stress the rightmost lexically accented suffix, otherwise stress the leftmost lexical accent"—proceeding right-to-left, the rightmost accented suffix is the rightmost accent that has a] symbol to its left.[10]

(10) a. σ́ [σ́ [σ́]] → **σ́**σσ (see 6)

σ́	[	σ́	[	σ́	]	]
σ́		σ		σ		

b. σ́ [σ́ [[σ́] σ́]] → σσσ**σ́** (see 7)

σ́	[	σ́	[	[	σ́	]	σ́	]	]
σ		σ			σ		**σ́**		

But just as it is impossible for a finite state automaton to match subjects to verbs, or count them, it is impossible to make sure there are as many [symbols as] symbols in the input. So, finite state automata cannot identify arbitrary constituents.

For example, a finite state mapping could never insert n syllables, where n is the depth of nesting, across unbounded levels of nesting. A finite state mapping could also never change [t] to [s] whenever a higher constituent contained [i]—the dependency between the underlying–surface pair (t → s) and [i] could be, for example, between a prefix and a suffix. Computing this function implies identifying the constituent containing [t].

(11)

[	t	a	[	t	i	]	] ,
	t	a		t	i		

[	[	t	a	]	t	i	] ,
		s	a		t	i	

[	[	t	a	[	t	i	]	]	t	u	]
		t	a		t	i			t	u	

WHIS, rather than changing [t] to [s] whenever a higher constituent contains [i], makes an underlying accent unstressed or secondary stress whenever a higher constituent contains an underlying accent. In the next section, we give a more detailed informal proof that WHIS is not finite state, and claim that this accounts for the fact that WHIS is unattested.

4 WHIS is not finite-state

In proving that a computation is not finite-state, the key is to think of the input string as two parts, w followed by v. Whether wv is accepted or rejected depends on what state the hypothetical automaton ends in, which in turn depends on what state it is in after reading

w, what state it then goes to after reading v, and nothing else. There are infinitely many different possible values of w, but, since there are only finitely many states, eventually, the computation must loop back on itself, going back into a previous state.

One way of proving that a computation is not finite state is to reason about v as a function of w. The Myhill–Nerode Theorem (Nerode 1958) states that, given an arbitrary string w, and the set of acceptable strings L according to a finite-state computation, w falls into one of a finite set of equivalence classes, defined by what strings v make $wv \in L$.

For example, suppose a computation accepts all strings of zero or more a's and b's where all the b's must follow all the a's: *ab*, *aaab*, *abbb*, *aa*, *bbb*, … (a^*b^*). This gives three equivalence classes. If w is all a's, it can be extended by v consisting of zero or more a's followed by zero or more b's. All $w \in a^*$ are equivalent in that they are extendable by $v \in a^*b^*$. If w is a sequence of a's followed by one or more b's, it is extendable by v having zero or more b's. All $w \in a^*bb^*$ are extendable by $v \in b^*$. Finally, any other w is not extendable in L. The set of extensions for $w \notin a^* \cup a^*bb^*$ is $\{\}$. These three disjoint sets of possible w make up the whole set of possible strings, thus, the computation is finite state.

In contrast, consider the computation accepting only matched pairs (ab, $aabb$, $aaabbb$, …). The possible extensions of $w = aa$ are $v =$ *bb*, *abbb*, *aabbbb*, …. Not one of these is a possible extension of $w = a$, or $w = aaa$, or $w = aaaa$, or anything else. The total number of a and b must always be equal, and so each w defines its own unique class of extensions. There are infinitely many possible w's, thus this computation is not finite state.

When the symbols are pairs, the problem is the same. In the case of WHIS, we can reduce the problem to the distribution of one pair, (Ṕ→**Ṕ**): accent realized as surface stress in a prefix. WHIS is the pattern where (Ṕ→**Ṕ**) is possible only if there is no higher accented suffix Ś; otherwise we only get (Ṕ→P).[11]

We factor out the question of whether the brackets are properly matched in the input: we assume the morphosyntax yields only licit structures. The question is, instead, whether there is any computation which will give the right pattern, *when* its input is restricted to only legitimate morphological structures.

Consider w with an accented prefix realized as stressed, two nodes higher than the root:

(12) $\ldots (\acute{\mathrm{P}} \rightarrow \acute{\mathbf{P}}) \ldots [\ldots [\sqrt{\mathrm{Root}}$

An extension containing Ś→S will only be legitimate if Ś is lower than Ṕ. That means that any v of the following form will be a valid extension:

(13) $\ldots] \ldots (\acute{\mathrm{S}} \rightarrow \mathrm{S}) \ldots]\ \ldots$, with no following Ś→S or Ś→**Ś**

On the other hand, a v of the following form will never be a valid extension:

(14) $\ldots]\ \ldots]\ \ldots (\acute{\mathrm{S}} \rightarrow \mathrm{S}) \ldots$

Now consider a w with the prefix one node higher than in (12):

(15) $\ldots (\acute{\mathrm{P}} \rightarrow \acute{\mathbf{P}}) \ldots [\ldots [\ldots [\sqrt{\mathrm{Root}}$

Now, an extension of the form (14) will always be valid. Thus the set of valid extensions is different. It is easy to show that it is different at each level we put P at. Thus the computation is not finite state. Thus, if WHIS existed, phonology could not be finite state.

The problem is not solved by the fact that proper bracketing is the domain of the morphology. This would only be a solution if the morphology could filter out the extensions that distinguish the different classes of w, so that the phonological computation would not have to deal with classifying them. We will see an example of this when we look at Chamorro. However, as far as the morphology is concerned, (14) is a legitimate extension of both (12) and (15); it is up to the phonology to decide which to accept.

5 Chamorro is finite-state

Chamorro is different from WHIS in one crucial way. Since it is only the presence or absence of a higher suffix that matters to whether we get (Ṕ→**Ṕ**) or (Ṕ→P), it does not matter what becomes of the bracketing to the right of Ṕ. For such an analysis to work, we need two further assumptions. The first is that the morphology gives us binary branching, so that two affixes cannot have the same morphological prominence. The second is that only non-null morphemes introduce brackets that will be passed on to the phonology, so that no vacuous cycles are seen by the phonology.[12] Given these two assumptions, when there is an overt suffix structurally higher than (Ṕ→P), we will find the sequence [[somewhere to the left of (Ṕ→P): binary branching guarantees that any higher suffix will generate such a sequence, and the lack of vacuous structure guarantees that [[is not inserted with an empty prefix. We can construct a finite state machine relying on the fact that the sequence …[[…(Ṕ→**Ṕ**) is always illicit.

6 Phonological processes are subsequential

In this section, we show that there is another computationally interesting reason for WHIS to be impossible.

The finite state property holds of not only individual phonological processes, but entire phonological grammars. Tighter restrictions have been discovered to hold empirically on individual processes. Chandlee and Heinz (2012) proposed that phonological processes belong to a subset of the finite state computations called the subsequential functions.[13] The details of subsequential functions are not relevant here; the idea is that they cannot use arbitrary lookahead to see whether a particular change (a→b) is licensed.

Here is a pattern which is finite state, but not subsequential: given an accented prefix and an accented suffix, the prefix is stressed if it is higher, *or* if the suffix is not rightmost.

(16) *Gentler Winner-by-Height-If-Special:*

a. [tú [be tá]] → túbeta

b. [[tú be] tá] → tubetá

c. [ti [[tú [be tá]] to]] → titúbetato

d. [ti [[[tú be] tá] to]] → titúbetato

Gentler Winner-by-Height-If-Special (G-WHIS) is sensitive to relative scope, but it is finite state.[14] As in Chamorro, we can check if an accented prefix has a pair of empty brackets to its left to determine if it there is a higher suffix. We also check to see if the rightmost suffix is accented. The composition of any two functions represented by finite-state automata is also finite state. Thus, we compose the two and get a finite state automaton for G-WHIS.

On the other hand, G-WHIS is neither left subsequential nor right subsequential. To show that a function is not left subsequential, one needs to show that the condition for a non-faithful input–output pair, (a change), includes some *finitely close* element on the right (or that it is only triggered on the left). Once a hypothetical left subsequential transducer for G-WHIS reached [[, it would be unable to determine whether to license a (Ṕ→**Ṕ**) or a (Ṕ→P) until it reached the last suffix, arbitrarily far away. Thus G-WHIS is not left subsequential. It is also not right subsequential, because the computation over the reversal is not left subsequential. Consider the string in reverse: an initial (originally final) Ś will be realized as (Ś→**Ś**) or (Ś→S), depending on whether there is a Ṕ[[some unspecified distance away. Since G-WHIS is neither left nor right subsequential, it is not subsequential.

Chamorro is both left and right subsequential, since only the local environment [[matters to the realization of Ṕ. Nez Perce is not left subsequential, but it is right subsequential (unlike in G-WHIS, Ś does not depend on Ṕ in any way for its realization; the rightmost non word-final accented suffix is always stressed). Both Chamorro and Nez Perce stress placement conform to the generalization that phonological processes are subsequential.

7 Conclusion

The facts we have presented do not bear directly on *whether* there is serial feeding between cyclic domains. Rather, we have argued against a theory (i) with cyclic feeding, in which (ii) the domains passed to the phonology are isomorphic to true morphosyntactic constituents.

Stress assignment in Chamorro appears to use the full power of the phonological cycle, If this were the right analysis, determining the boundary between the two affix domains would have the net effect of locating where each constituent begins and ends. But as we have seen, with the right morphological representation, Chamorro can be analyzed making use of only local information. This is important, because unrestricted use of constituency in this way (as predicted by cyclic theories) exceeds the established outer computational limits of phonology. If the WHIS pattern is attested, however, it would indeed require that the net computation of phonology be non finite state. In that case the generalization that phonology is finite state would have to be re-evaluated.

Notes

[1]All Chamorro generalizations and examples are drawn from Chung 1983.

[2]Chung's data indicates that existing stresses are demoted to secondary. This does not affect our argument, and so we will leave secondary stresses out.

[3]All Nez Perce generalizations and examples are drawn from Crook 1999.

[4]Nez Perce nouns show a different pattern: in the absence of any lexical accent, stress on nouns is required to surface on the noun root, even if that results in stress to the left of the penultimate syllable.

[5]That the directional suffix is below the scope of the plural object marker is also illustrated by the fact that the addition of *(n)úu* changes which argument is identified as plural (see Crook 1999, p. 480).

[6]See Crook 1999 for evidence that this stress shift does not occur with unaccented suffixes.

[7]Lexical accent in Nez Perce is unpredictably associated with individual syllables, rather than with entire morphemes, and so is most reasonably treated as part of the phonological

representation, rather than, for example, in terms of morphosyntactic features or diacritics.

[8]As far as we know, there are no accented suffixes in Nez Perce that structurally outscope an accented prefix, and so we cannot show any case in which the winning suffix is structurally higher than the prefix. This leaves the pattern open to an alternate, structure-sensitive analysis, in which accented suffixes get stress only when they are structurally *lower* than an accented prefix. This is an implausible generalization under a cyclic theory, but a system like this would be a counterexample to the typological generalization we introduce below.

[9]This is clearly true if morphological structure simply is syntactic structure, as in theories like Distributed Morphology (Halle and Marantz, 1993, et seq.), but the empirical arguments that morphological structures can have center embeddings, for example from Bar-Hillel et al. 1961 and Carden 1983, apply even if morphological and syntactic structures are different. Savitch (1993) presents a more general argument in favour of assuming that linguistic patterns are formally unbounded. Below, we assume that the Chamorro pattern holds for unboundedly many levels of nesting. Since the prefixes and suffixes in question belong to a closed class, some division into affix strata could be found, yielding a bounded version of the pattern. If it could be argued that there were no need for unbounded morphological nesting at all, then it would be easy to show that a bounded version of the WHIS pattern, which we claim is typologically absent, would be finite state. However, it would nevertheless be impossible for the same reason as the G-WHIS pattern we discuss below: it would not be subsequential.

[10]In our proof, we show phonological representations as unanalyzed segments, with morphological constituency added as the only other structure. This is a simplification for the sake of presentation; the result holds even if we assume autosegmental representations, or other phonological representations that are graphs rather than just sequences. Our conclusion is general: if phonology is finite state, then, whatever representations phonology deals with, they do not encode the true morphological constituency.

[11]Similarly, (Ś→**Ś**) is only possible if there is no higher Ṕ, and, otherwise, we can only get (Ś→S); but the dependency of the prefix on the suffix turns out to be enough.

[12]This distinguishes between phonologically vacuous morphemes and higher affixes introduced at the opposite edge of a word: only the latter results in a [[(or]]) sequence. Null morphemes—morphemes that never have phonological content—should also be distinguished from morphemes that are coincidentally rendered silent by some phonological process, and possibly also from morphemes with a zero allomorph in only some environments. Here we adopt the strong position that phonological alternations are never sensitive to phonologically vacuous cycles, though morphological alternations can be, as in Pinker's (1999) proposal that the difference in the English past tense between [flajd] (regular, as in baseball "flyed out") and [flu] (irregular) arises due to [flayd] containing an extra layer of vacuous structure, (a reverbalized nominalization), without this being unworkable. Under no standard analysis is this alternation purely phonological. It is universally attributed, instead, to morphological properties of the two environments: either separate lexical entries, as in lexicalist approaches, or locality constraints on root-triggered readjustment rules in a theory like Distributed Morphology. This position may be too strong, however: Marvin's (2002) account of stress preservation in English, for example, relies on phonology being sensitive to vacuous morphological structure. There, the difference between *trànsp*[ə]*rtátion*, with an unstressed vowel subject to reduction, and *còndènsátion*, with a non-primary stress hypothesized to be preserved from an earlier cycle, is driven by there being an extra, segmentally vacuous cycle in *condensation*. If this analysis is correct, at least some vacuous cycles must be visible in the phonology. Our analysis of Chamorro could be maintained in this case, if vacuous cycles triggered stress shift just as higher suffixes do. Given the data available, we cannot determine whether any vacuous cycles interact with the relevant stress pattern.

[13]See also Heinz and Lai 2013, Chandlee 2014. For details on subsequential functions,

see Schützenberger 1977, Mohri 1997.

[14]Notice this is different from Nez Perce. Nez Perce is sensitive to whether the suffix is rightmost, but it is insensitive to the relative scope of prefixes and suffixes.

References

Bar-Hillel, Yehoshua, Micha Perles, and Eliahu Shamir. 1961. On formal properties of simple phrase structure grammars. *STUF-Language Typology and Universals* 14:143–172.

Bermúdez-Otero, Ricardo. 2011. Cyclicity. In *The Blackwell companion to phonology*, 2019–2048. Oxford: Blackwell Publishers.

Berwick, Robert C., Kazuo Okanoya, Gabriel J.L. Beckers, and Johan J. Bolhuis. 2011. Songs to syntax: the linguistics of birdsong. *Trends in Cognitive Sciences* 15:113–121.

Carden, Guy. 1983. The non-finite-state-ness of the word formation component. *Linguistic Inquiry* 537–541.

Chandlee, Jane. 2014. Strictly local phonological processes. Doctoral Dissertation, University of Delaware, Newark.

Chandlee, Jane, and Jeffrey Heinz. 2012. Bounded copying is subsequential: Implications for metathesis and reduplication. In *Proceedings of the Twelfth Meeting of the Special Interest Group on Computational Morphology and Phonology*, 42–51. Association for Computational Linguistics.

Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. Harper and Row.

Chomsky, Noam, Morris Halle, and Fred Lukoff. 1956. On accent and juncture in English. In *For Roman Jakobson*, ed. Morris Halle, 65–80. The Hague: Mouton.

Chung, Sandra. 1983. Transderivational relationships in Chamorro phonology. *Language* 59:35–66.

Crook, Harold. 1999. The phonology and morphology of Nez Perce stress. Doctoral Dissertation, University of California, Los Angeles.

Frank, Robert, and Giorgio Satta. 1998. Optimality theory and the generative complexity of constraint violability. *Computational Linguistics* 24:307–315.

Graf, Thomas. 2010. Comparing incomparable frameworks: A model theoretic approach to phonology. *University of Pennsylvania Working Papers in Linguistics* 16.

Halle, Morris, and Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In *The view from building 20*, ed. Kenneth Hale and Samuel Jay Keyser. Cambridge, MA: MIT Press.

Heinz, Jeffrey. 2011. Computational phonology - Part I: foundations. *Language and Linguistics Compass* 5:140–152.

Heinz, Jeffrey, and Regine Lai. 2013. Vowel harmony and subsequentiality. In *13th Meeting on the Mathematics of Language*, 52. Sofia: Association for Computational Linguistics.

Johnson, C. Douglas. 1972. *Formal aspects of phonological description*. The Hague: Mouton.

Kaplan, Ronald M., and Martin Kay. 1994. Regular models of phonological rule systems. *Computational Linguistics* 20:331–378.

Kean, Mary-Louise. 1974. The strict cycle in phonology. *Linguistic Inquiry* 5:179–203.

Kiparsky, Paul. 2000. Opacity and cyclicity. *Linguistic Review* 17:351–366.

Marvin, Tatjana. 2002. Topics in the stress and syntax of words. PhD thesis, Massachusetts Institute of Technology, Cambridge.

Mascaró, Joan. 1976. Catalan phonology and the phonological cycle. PhD thesis, Massachusetts Institute of Technology, Cambridge.

Mohri, Mehryar. 1997. Finite-state transducers in language and speech processing. *Computational Linguistics* 23.

Nerode, A. 1958. Linear automaton transformations. *Proceedings of the American Mathematical Society* 9:541.

Piggott, Glyne, and Heather Newell. 2005. Syllabification, stress and derivation by phase in Ojibwa. *McGill Working Papers in Linguistics* 20.

Pinker, Steven. 1999. *Words and rules: The ingredients of language*. New York: HarperCollins.

Savitch, Walter J. 1993. Why it might pay to assume that languages are infinite. *Annals of Mathematics and Artificial Intelligence* 8:17–25.

Scheer, Tobias. 2010. *A guide to morphosyntax-phonology interface theories: How extra-phonological information is treated in phonology since Trubetzkoy's Grenzsignale*. Berlin: Walter de Gruyter.

Schützenberger, M.P. 1977. Sur une variante des fonctions sequentielles. *Theoretical Computer Science* 4:47–57.