# ENGLISH STOP ALLOPHONES IN METRICAL THEORY

John T. Jensen

Department of Linguistics
University of Ottawa
Ottawa, Ontario K1N 6N5 Canada

ABSTRACT

The foot, a prosodic unit containing one stressed syllable, is the domain for determining the allophones of stops in English. Aspiration is restricted to foot-initial position. Consonants are laxed within a foot after a nonconsonantal segment and lax voiceless stops are glottalized in syllable codas; lax alveolar stops are flapped syllable initially. Some revisions to the rules establishing feet are proposed. Because the metrical grid provides no constituents, it is not adequate for predicting the distribution of stop allophones in English.

## INTRODUCTION

In contemporary phonology there is general agreement that representations need to be enriched with prosodic organization, including such units as the syllable and the foot. This view contrasts sharply with the practice of early generative phonology [1], where phonological representations consisted entirely of strings of segments and boundaries. The original motivation for metrical theory was to offer a more natural account of stress systems [11], but it soon became apparent that prosodic organization also allows for the correct description of certain segmental processes as well. Aspiration of voiceless stops in English, for example, occurs in a variety of disparate environments. Selkirk [15] lists word-initial position *(Toronto)*, before a stressed vowel unless [s] precedes *(hotel* vs *astonish)*, before a sonorant plus a stressed vowel unless [s] precedes or [t] is followed by [l] *(apply* vs *display, Atlantic)*. Such a process is difficult to describe in purely segmental terms, and indeed no systematic account of stop allophony appears in *The Sound Pattern of English* [1]. Selkirk observes correctly that aspiration occurs only in syllable-initial position, which partially accounts for these observations. In order to account for the nonaspiration of the underlined stops in words like *happy, hefty*, Selkirk proposes language-particular resyllabification rules that attract consonants leftward out of stressless syllables, giving *happ.y, heft.y*, thereby removing these stops from the domain of aspiration. While I find these resyllabifications counterintuitive, there is an empirical argument against this analysis. Selkirk's resyllabification rules are subject to a structure-preservation principle that requires derived syllables to conform to the canonical syllable patterns of the language. In a word like *At.kins*, resyllabification to **Atk.ins* is impossible, since English syllables never end in *-tk*. This predicts that [k] of *Atkins* should aspirate, which it does not, any more than the [t] of *actor*, where *act.or* would be a possible resyllabification.

A second syllable-based approach to English stop allophones is that of Kahn [8], which is couched in terms of autosegmental phonology. Instead of resyllabification, Kahn allows consonants to be ambisyllabic, i.e., part of both the preceding and following syllables. This would be the case of [p] in *happy* and [t] in *hefty*, for example. Kahn's rule aspirates voiceless stops that are syllable initial but not syllable final (i.e., not ambisyllabic) and thus achieves the same effect as Selkirk, and runs into the same difficulty with *Atkins*. Since [k] here can't be ambisyllabic, he wrongly predicts that it should aspirate. (In fact, he claims that it does aspirate in slow speech, but I find this possible only in very careful speech where both syllables are stressed.)

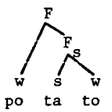
Kiparsky [9] was the first to propose that the stress feet of Liberman and Prince [11] could also be considered the domain of certain segmental processes. Instead of resyllabification or ambisyllabification, Kiparsky proposed rule (1) (modified).

(1) C → [-tense] / ...[-cons]____ within a foot

Kiparsky restricts aspiration to tense voiceless stops at the beginning of a syllable, thus accounting for *happy*. But Kiparsky predicts aspiration on the second syllables of *hefty, Atkins*, where the stops [t] and [k] are unaffected by rule (1), since they are preceded by [+consonantal] [f] and [t] respectively. Hammond [4] notices such problems with the foot-based analysis, and advocates a return to Kahn's ambisyllabic approach. I propose to retain the foot-based approach, but to restrict aspiration to foot-initial position. Some modification of Kiparsky's system is needed anyway. Working within the original metrical framework [11], Kiparsky retained the feature [±stress] and with it the possibility of stressless feet. He analyzes *potato* as two feet, the first unstressed, $[_{F}$po$][_{F}$tato], predicting aspiration on the foot-initial [p] and [t] and flapping (via laxing) of the second [t]. Since then, metrical theory has rejected the feature [stress], holding that stress is the property of being the strongest syllable in a foot [14]. If [po] of *potato* is not a foot, and aspiration is limited to foot-initial position, how does the [p] come to be aspirated? Hayes [5, 6] proposes that stray syllables (i.e. those not associated with any foot) are adjoined to an adjacent foot. If we assume that ad-

junction creates nested feet, we get the representation in (2) (where w=weak, s=strong, F=foot; s and w must always appear as sisters to each other.

(2)




In (2), both [p] and the first [t] are foot initial, and so get aspirated, while the second [t] laxes and flaps, as in Kiparsky's treatment. This captures the essence of Kiparsky's proposal, and resolves Kiparsky's problem with *Atkins* and *hefty*.

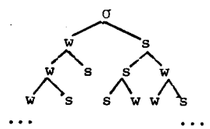
Subsequent studies have confirmed the role of the foot in segmental phonology as well as stress systems. Prince [12] states rules for gradation and overlength in Estonian partly in terms of foot structure. Similarly, Hayes [7] discusses certain segmental processes in Yidiny, an Australian language, in terms of foot conditioning, thereby obviating the necessity for phonological rules to refer to the odd-numbered syllables in a word. Even for stress systems, constituency is necessary. Halle and Vergnaud [3] cite a number of studies showing that deletion of a (potentially) stressed vowel in many languages results in a stress shift to an adjacent syllable within the foot.

In the remainder of this paper, I will first review the properties of syllables, propose some modifications to Hayes's rules of foot construction, then show the role of the foot in English stop allophones, making crucial reference to rule (1).

SYLLABLES

The acoustic record provides no direct evidence of syllables and their boundaries. The syllable is an abstract unit which makes it possible to provide a more insightful statement of certain phonological processes. Among competing approaches, we assume the metrical representation of Kiparsky [9], in which the syllable has the same type of s-w labelling as the foot, as in (3).

(3)




In this representation, sister constituents are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, l, r, glides, vowels. In addition, English imposes language-specific constraints on onsets and rimes. For example, a syllable cannot begin with a sequence of two stops (including nasals), and the rime is limited to the sequence V([+sonorant])(C)([+coronal]), where the 'coronal' position may exceptionally contain [st] or [sθ] as in *next, sixth*. An additional [s], [z], [t], or [d] may follow if it is inflectional, e.g. *sixths*. Even though contrary to the sonority hierarchy, [s] plus voiceless stop can occur in the onset, and also the sequence [s] plus voiceless stop plus liquid (but not **stl-*).

The question of dividing between syllables is a more difficult one. There is no difficulty with *At.kins*, which can only be syllabified as shown. Kiparsky proposed that, in English, the onset is maximized when two or more divisions are possible at the boundary between two syllables. In this respect English contrasts with Finnish and Estonian, where the coda is maximized, and where, in general, the onset is limited to a single consonant. This raises an interesting question: what happens to VsT(R)V clusters (where T=voiceless stop, R=liquid or glide)? The sonority hierarchy predicts Vs.T(R)V; the onset maximization principle predicts V.sT(R)V.

Davidsen-Nielsen [2] investigated this question experimentally. He measured the degree of aspiration in words like *despise* and compared it to that of words like *pin* (aspirated) and *spin* (unaspirated). Measurements revealed that the stops in words like *despise* are normally unaspirated, supporting the syllabification V.sT(R)V. The only exceptions were in words that contained "a prefix with *-s* followed by an intuitively transparent morpheme boundary, e.g. *miscalculate, discourteous,*" where the stops are aspirated, thus supporting a syllable division coinciding with the morpheme boundary, i.e. Vs.T(R)V in these words.

In a sense, or course, this argument is circular. The syllable boundary is inferred from the degree of aspiration on the stop, while the rule for aspiration is assumed to affect only syllable-initial stops. (The stops in question are also foot initial, and so consistent with our hypothesis also.) However, this conclusion is independently supported by the stress pattern of these words. The prefixes *mis-* and *dis-* exhibit secondary stress, and we might expect to find similar effects from a preceding stressed syllable, even if it doesn't constitute a morpheme. In Davidsen-Nielsen's material, *gèstátion* and *fàstídious* have a somewhat greater average degree of aspiration (3.0 and 3.09 csec respectively) than *bestow* and *establish* (2.5 and 2.41 csec, respectively. To test this further, I recorded two speakers of North American English in words containing *s*-stop clusters, both with and without stress on the preceding syllable. Results were analyzed using a Mingograf 804 connected to a Kay Elemetrics Visipitch 6087 and also with a Kay Elemetrics Sonagraph 7800. The results are given in Table 1.

| | | JJ | ML |
|---|---|---|---|
| a. | ìnfèstátion | 2 | 2.5 |
| | èlàstícity | 5 | 3 |
| | plàstícity | 3 | 2 |
| | pèstíferous | 2 | 2 |
| | óstensible | 2 | 2 |
| | MEAN | 2.8 | 2.3 |
| b. | askánce | 1 | 2 |
| | órchestra | 2 | 2 |
| | astónish | 1.5 | 1 |
| | pedestal | 1 | 2 |
| | sustáin | 1 | 1.5 |
| | MEAN | 1.3 | 1.7 |

Table 1. Duration of release stage (in csec.) of medial stops (underlined) in words with (a) and without (b) secondary stress on the preceding syllable.





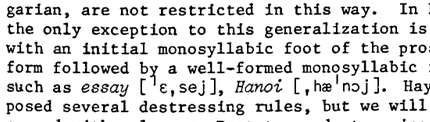


While these results are not conclusive, there is somewhat more aspiration in words where a secondary stress precedes the cluster in question than when the preceding syllable is unstressed. This supports the syllabification Vs.T(R)V for the words of Table 1(b).

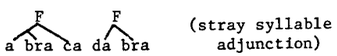
FEET AND ASPIRATION

For the core system of English stress, Hayes [5] proposed left-dominant maximally binary feet constructed right to left across a word. Ternary feet (i.e. with three syllables) can arise only by adjoining a stray syllable to a binary foot. Syllables become stray either by being made extrametrical or as a result of destressing. Before foot assignment, the final consonant of a word, the final suffix of an adjective, and the final syllable of a noun are extrametrical. The rightmost foot may be binary only if its second syllable ends in a short (lax) vowel. Conversely, a monosyllabic foot must contain a long vowel, a diphthong, or at least one final consonant. This accounts for the familiar observation that a monosyllabic (stressed) word cannot end in a 'checked' vowel; i.e. *bee* [biː] and *bit* [bɪt] are possible (and actual) words, while **bi* [bɪ] is an impossible word.. Prince [12] claims that Estonian is subject to the same constraint on possible feet. Other languages, e.g. French and Hungarian, are not restricted in this way. In English, the only exception to this generalization is words with an initial monosyllabic foot of the proscribed form followed by a well-formed monosyllabic foot, such as *essay* [ˈɛˌseɪ], *Hanoi* [ˌhæˈnɔɪ]. Hayes proposed several destressing rules, but we will be concerned with only one: Poststress destressing, which removes a binary foot whose first syllable is open and which immediately follows a monosyllabic foot. Hayes appeals to this rule in his derivation of words like *abracadabra* (4).

(4) a. b.




c.



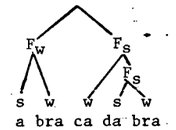

Speakers find the division (4c) counterintuitive. To test this, I asked a group of 28 native English-speaking first-year undergraduate linguistics students to divide words of this type "into two parts, according to the pronunciation." None of the subjects knew the purpose of the test beforehand. Control words were inserted into the list to prevent extraneous strategies from being used. Subjects had a printed list of words and were asked to indicate a single division between letters in each as they were pronounced by the author, with only a short interval between tokens. The results are shown in Table 2.

With two unstressed syllables flanked by two stressed syllables, the preferred pattern seems to be to join the first unstressed syllable to the preceding foot and the second to the following foot, as long as the second unstressed syllable is open. This produces the intuitively correct structure (5) from (4b).

| | 1/ | | 2/ | | 3/ | | other | |
|---|---|---|---|---|---|---|---|---|
| abracadabra | 1 | *3.6%* | 23 | *82.1%* | 3 | *10.7%* | 1 | *3.6%* |
| Navratilova | 5 | *17.8%* | 22 | *78.6%* | 1 | *3.6%* | 0 | |
| Winnepesaukee | 1 | *3.6%* | 14 | *50.0%* | 11 | *39.3%* | 2 | *7.1%* |
| Tippecanoe | 4 | *14.3%* | 24 | *85.7%* | 0 | | 0 | |
| Luxipallila | 3 | *10.7%* | 22 | *78.6%* | 1 | *3.6%* | 2 | *7.1%* |
| Nebuchadnezzar | 2 | *7.1%* | 5 | *17.8%* | 17 | *60.7%* | 4 | *14.3%* |
| Kilimanjaro | 1 | *3.6%* | 17 | *60.7%* | 10 | *35.7%* | 0 | |

Table 2. Division of words into two parts. Number *(and percentage)* of responses. Column headings: 1/ indicates a division after the first syllable, 2/ after the second, 3/ after the third. Other includes no division or more than one division.

(5)




If the second unfooted syllable is closed, it joins the foot on the left: 60.7% of subjects preferred the division [Nebuchad][nezzar]. Prince suggests that this is because *chadnezzar* (with the first syllable unstressed) is not a possible word type in our terms, a possible foot [13]. On the other hand, *Kilimanjaro* may go both ways, [Kili][manjaro] or [Kiliman][jaro], since both divisions into two parts give two possible phonological words, or feet.

Prince uses facts such as these to argue against the foot as a legitimate phonological unit. Because the metrical theory of stress uses only a small fraction of the types of tree structure that the theory allows in principle, he proposes eliminating the trees and displaying relative prominence in terms of a grid, in which column height correlates with greater prominence. Such a representation has no constituents, and thus no way of capturing the segmental processes that we have seen depend on these constituents. Prince notes the virtually obligatory aspiration of the [t] of *Navratilova*, unexpected if it is metrically structured as in (4c). However, rather than discard foot theory, the answer lies in modifying it so that it will produce structures like (5), where aspiration of [k] *(abracadabra)* and [t] *(Navratilova)* is expected, under the hypothesis that aspiration of tense voiceless stops occurs only in foot-initial position.

As with syllabification, we sought instrumental verification of the proposed division into feet (5). Table 3 gives the duration of the release stage of the stops at the beginning of the third syllables of the words of Table 2 (except for *Kilimanjaro*, which has no stop in that position). Speakers and equipment are the same as in Table 1.

| | JJ | ML |
|---|---|---|
| abracadabra | 4 | 4 |
| Navratilova | 7 | 6 |
| Winnepesaukee | 5 | 2 |
| Tippecanoe | 5 | 5 |
| Luxipallila | 2 | 5 |
| Nebuchadnezzar | 6 | 4 |
| MEAN | 4.83 | 4.33 |

Table 3. Duration of release stage (in csec.) of medial stops (underlined) in potential foot-initial position

These results are consistent with the hypothesis that (5) represents the correct foot structure, on the assumption that only foot-initial voiceless stops are aspirated.

## GLOTTALIZATION

Glottalization of stops is manifested differently in various English dialects. Cockney is notorious for the extent to which glottalization appears between vowels. In RP and North American dialects, glottalization is restricted to voiceless stops in syllable codas laxed by rule (1). Examples are *octave*, *atlas*, *at* Lynne's. The only case where voiceless stops are glottalized in syllable-initial position is before syllabic [n], as in *kitten*. Nonrhotic speakers (e.g. RP) can also have glottalized [t] in words like *pattern*, where *r*-loss makes the [n] syllabic; North American speakers, with syllabic [r] in such words, have the expected flap. It is notable that Cockney speakers use glottalized stops (or [ʔ]) where North American speakers have the flap. In my analysis, this results from the lack of the flapping rule in British dialects, coupled with the extension of the glottalization rule to lax voiceless stops in all positions, and is especially noteworthy when it affects labial and velar stops, as in [pəjʔə]'paper'.

In Selkirk's account, both flapped alveolars and glottalized stops are in syllable-final position as a result of her resyllabification rule. She therefore resorts to a feature [±release], claiming that alveolar stops are flapped in syllable-final position when they are released, generally before a vowel. Unreleased voiceless stops are glottalized. This runs into two difficulties, only one of which she discusses. Phrases like *get off* can only be pronounced [gɛɾɔf] by her account, with a flap. Kahn notes an alternate pronunciation [gɛtʾɔf] or [gɛʔɔf], both impossible under Selkirk's analysis, since stops are obligatorily released before vowels and thus never glottalized there. She proposes that [ʔ] is inserted before certain initial vowels under emphasis. This makes [t] unreleased, since it is followed by a nonvowel. While this works for the North American dialects she is discussing, it won't account for the Cockney facts just mentioned. The medial stop in *paper* is followed by a vowel, and there is no possibility of inserting [ʔ] under emphasis. In any case, nonrelease is not generally associated with glottalization, as many languages have phonemic released glottalized stops (e.g. Georgian). We conclude that it is more natural to describe the difference between glottalized and flapped allophones in English in terms of syllable position and dispense with the feature [release].

## FLAPPING

In North American English, but not in most forms of British English, alveolar stops [t], [d], [n] are flapped within words before stressless vowels, and often between words regardless of stress. So, the second [t] of *potato* is flapped, as is the [t] in *met Ann*, although this [t] can also be glottalized. The difference depends on the syllabic status of [t] here. Kiparsky proposes a rule that flaps alveolar stops in syllable-initial position if they are lax (by rule (1)). Since [t] in *met Ann* is lax, it will flap only if it is resyllabified with the following vowel; otherwise it is glottalized. We assume that resyllabification is optional at word boundaries. Notice that, in phrases, it doesn't matter that the following vowel is stressed. What matters is that the [t] of *met* is laxed within its foot before it is syntactically concatenated with *Ann*.

## SUMMARY

Laxing (1): Consonants become lax after a nonconsonantal segment within a foot.

Aspiration: Tense voiceless stops are aspirated at the beginning of a foot.

Glottalization: Lax voiceless stops are glottalized in the syllable coda. (Generalized in Cockney to all positions).

Flapping: Lax alveolar stops (including [n] are flapped in the syllable onset (North American only).

## ACKNOWLEDGEMENTS

I wish to thank Melanie Lukach, Régis Foré, Margaret Stong-Jensen, Helen Goodluck, and Phil Hauptman and his Introduction to Language class for help in preparing this paper. All errors are my own.

## REFERENCES

[1] N. Chomsky, M. Halle, *The Sound Pattern of English*, Harper and Row, 1968.
[2] N. Davidsen-Nielsen, "Syllabification in English words with medial *sp, st, sk*," *Journal of Phonetics* 2, 15-45, 1974.
[3] M. Halle, J.-R. Vergnaud, "Stress and the cycle," *Linguistic Inquiry* 18, 45-84, 1987.
[4] M. Hammond, "Foot domain rules and metrical locality," *West Coast Conference on Formal Linguistics* 1, 207-218, 1982.
[5] B. Hayes, *A Metrical Theory of Stress Rules*, MIT dissertation (IULC), 1981.
[6] B. Hayes, "Extrametricality and English stress," *Linguistic Inquiry* 13, 227-276, 1982.
[7] B. Hayes, "Metrical structure as the organizing principle of Yidiny phonology," *The Structure of Phonological Representations (Part I)* 97-110, Foris, 1982.
[8] D. Kahn, *Syllable-based Generalizations in English Phonology*, MIT dissertation (IULC), 1976.
[9] P. Kiparsky, "Metrical structure assignment is cyclic," *Linguistic Inquiry* 10, 421-441, 1979.
[10] P. Kiparsky, "From cyclic phonology to lexical phonology," *The Structure of Phonological Representations (Part I)* 131-175, Foris, 1982.
[11] M. Liberman, A. Prince, "On stress and linguistic rhythm," *Linguistic Inquiry* 8, 249-336, 1977.
[12] A. Prince, "A metrical theory for Estonian quantity," *Linguistic Inquiry* 11, 511-562, 1980.
[13] A. Prince, "Relating to the grid," *Linguistic Inquiry* 14, 19-100, 1983.
[14] E. Selkirk, "The role of prosodic categories in English word stress," *Linguistic Inquiry* 11, 563-605, 1980.
[15] E. Selkirk, "The syllable," *The Structure of Phonological Representations (Part II)*, 337-383, 1982.