ENGLISH STOP ALLOPHONES IN METRICAL THEORY

John T. Jensen

Department of Linguistics
University of Ottawa
Ottawa, Ontario KIN 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 lax 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 sonorant plus a stressed vowel unless [s] precedes (hotel vs agatonish), before a sonorant plus a stressed vowel unless [s] precedes or [t] is followed by [1] (apply vs display, Atlantisia). 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 happy, hefty, 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 Atkins, 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).

\[ (1) \ C + [-tense] / ...[-cons] \]

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, [po] [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-
In (2), both [p] and the first [c] are foot initial, and so get aspirated, while the second [c] laxes and flaps, as in Kiparsky's treatment. This captures the essence of Kiparsky's proposal, and resolves Kiparsky's problem with akoko and hasty.

Subsequent studies have confirmed the role of the foot in segmental phonology as well as stress systems. Prince [12] states rules for gradation and overlapping in Estonian partly in terms of foot structure. Similarly, Hayes [1] discusses certain segmental processes in Yidiny, an Australian language, in terms of foot conditioning, thereby obviating the need to refer to the odd-numbered syllables in a word. Even for syllables with a voiceless onset, however, 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.)

The acoustic record provides no direct evidence of syllables and their boundaries. The syllable is an abstract unit which makes it possible to provide more insightful statements of certain phonological phenomena. Hence it is useful to assume the metrical representation of Kiparsky [9], in which the syllable has the same type of s-v label

In this representation, syllables are constituents that are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, r, glide, vowels, and stressed vowels. English imposes language-specific constraints on onsets and rimes. For example, Kiparsky [3] notes that English has a sequence of two stops (including nasals), and the rime is limited to the monosyllabic foot [V(T)][(Rv)][(T)][(V)], where the 'coronal' position may exceptionally contain clusters (where [Pvoiceless stop] + [liquid] or [glide]).

The sonority hierarchy predicts V(T)(Rv)(T)(V), the onset maximization principle predicts V(T)(Rv)D(T).

Davidson-Maxwell [2] investigated this question experimentally. He measured the degree of aspiration in words like deepee and compared it to that of words like ploo (aspirated) and ploo (unaspirated). Measurements revealed that the stops in words like deepee are normally unaspirated, supporting the syllabification V(T)(Rv)D(T). The only exception were in words that contained a 'prefix with a -' followed by an intuitively transparent morpheme boundary, e.g. mis-also miscommunicate, whereas the stop was aspirated. It supports a syllable division coinciding with the morpheme boundary, i.e. V(T)(Rv)D(T) 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 other evidence for the sonority hierarchy. In addition, we might expect to find similar effects from a preceding aspiration also. Furthermore, we expect to find a greater degree of aspiration (3.0 and 3.05-msec respectively) than between /p/ and /t/ at 0.95 and 0.95, as expected.

In this representation, syllables are constituents that are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, r, glide, vowels, and stressed vowels. English imposes language-specific constraints on onsets and rimes. For example, Kiparsky [3] notes that English has a sequence of two stops (including nasals), and the rime is limited to the monosyllabic foot [V(T)][(Rv)][(T)][(V)], where the 'coronal' position may exceptionally contain clusters (where [Pvoiceless stop] + [liquid] or [glide]).

The sonority hierarchy predicts V(T)(Rv)(T)(V), the onset maximization principle predicts V(T)(Rv)D(T).

Davidson-Maxwell [2] investigated this question experimentally. He measured the degree of aspiration in words like deepee and compared it to that of words like ploo (aspirated) and ploo (unaspirated). Measurements revealed that the stops in words like deepee are normally unaspirated, supporting the syllabification V(T)(Rv)D(T). The only exception were in words that contained a 'prefix with a -' followed by an intuitively transparent morpheme boundary, e.g. mis-also miscommunicate, whereas the stop was aspirated. It supports a syllable division coinciding with the morpheme boundary, i.e. V(T)(Rv)D(T) 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 other evidence for the sonority hierarchy. In addition, we might expect to find similar effects from a preceding aspiration also. Furthermore, we expect to find a greater degree of aspiration (3.0 and 3.05-msec respectively) than between /p/ and /t/ at 0.95 and 0.95, as expected.

In this representation, syllables are constituents that are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, r, glide, vowels, and stressed vowels. English imposes language-specific constraints on onsets and rimes. For example, Kiparsky [3] notes that English has a sequence of two stops (including nasals), and the rime is limited to the monosyllabic foot [V(T)][(Rv)][(T)][(V)], where the 'coronal' position may exceptionally contain clusters (where [Pvoiceless stop] + [liquid] or [glide]).

The sonority hierarchy predicts V(T)(Rv)(T)(V), the onset maximization principle predicts V(T)(Rv)D(T).

Davidson-Maxwell [2] investigated this question experimentally. He measured the degree of aspiration in words like deepee and compared it to that of words like ploo (aspirated) and ploo (unaspirated). Measurements revealed that the stops in words like deepee are normally unaspirated, supporting the syllabification V(T)(Rv)D(T). The only exception were in words that contained a 'prefix with a -' followed by an intuitively transparent morpheme boundary, e.g. mis-also miscommunicate, whereas the stop was aspirated. It supports a syllable division coinciding with the morpheme boundary, i.e. V(T)(Rv)D(T) 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 other evidence for the sonority hierarchy. In addition, we might expect to find similar effects from a preceding aspiration also. Furthermore, we expect to find a greater degree of aspiration (3.0 and 3.05-msec respectively) than between /p/ and /t/ at 0.95 and 0.95, as expected.

In this representation, syllables are constituents that are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, r, glide, vowels, and stressed vowels. English imposes language-specific constraints on onsets and rimes. For example, Kiparsky [3] notes that English has a sequence of two stops (including nasals), and the rime is limited to the monosyllabic foot [V(T)][(Rv)][(T)][(V)], where the 'coronal' position may exceptionally contain clusters (where [Pvoiceless stop] + [liquid] or [glide]).

The sonority hierarchy predicts V(T)(Rv)(T)(V), the onset maximization principle predicts V(T)(Rv)D(T).

Davidson-Maxwell [2] investigated this question experimentally. He measured the degree of aspiration in words like deepee and compared it to that of words like ploo (aspirated) and ploo (unaspirated). Measurements revealed that the stops in words like deepee are normally unaspirated, supporting the syllabification V(T)(Rv)D(T). The only exception were in words that contained a 'prefix with a -' followed by an intuitively transparent morpheme boundary, e.g. mis-also miscommunicate, whereas the stop was aspirated. It supports a syllable division coinciding with the morpheme boundary, i.e. V(T)(Rv)D(T) 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 other evidence for the sonority hierarchy. In addition, we might expect to find similar effects from a preceding aspiration also. Furthermore, we expect to find a greater degree of aspiration (3.0 and 3.05-msec respectively) than between /p/ and /t/ at 0.95 and 0.95, as expected.

In this representation, syllables are constituents that are required to observe the sonority hierarchy, according to which segments are ordered (from weakest to strongest) as stops, fricatives, nasals, r, glide, vowels, and stressed vowels. English imposes language-specific constraints on onsets and rimes. For example, Kiparsky [3] notes that English has a sequence of two stops (including nasals), and the rime is limited to the monosyllabic foot [V(T)][(Rv)][(T)][(V)], where the 'coronal' position may exceptionally contain clusters (where [Pvoiceless stop] + [liquid] or [glide]).

The sonority hierarchy predicts V(T)(Rv)(T)(V), the onset maximization principle predicts V(T)(Rv)D(T).

Davidson-Maxwell [2] investigated this question experimentally. He measured the degree of aspiration in words like deepee and compared it to that of words like ploo (aspirated) and ploo (unaspirated). Measurements revealed that the stops in words like deepee are normally unaspirated, supporting the syllabification V(T)(Rv)D(T). The only exception were in words that contained a 'prefix with a -' followed by an intuitively transparent morpheme boundary, e.g. mis-also miscommunicate, whereas the stop was aspirated. It supports a syllable division coinciding with the morpheme boundary, i.e. V(T)(Rv)D(T) 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 other evidence for the sonority hierarchy. In addition, we might expect to find similar effects from a preceding aspiration also. Furthermore, we expect to find a greater degree of aspiration (3.0 and 3.05-msec respectively) than between /p/ and /t/ at 0.95 and 0.95, as expected.
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 lax in 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 -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 [po'ʔe] '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 [ɡe'ʔe] by her account, with a flap. Kahn notes an alternate pronunciation [ɡe'ʔe] or [ɡe'ʔe], 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 lax 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 Strong-Jensen, Helen Goodluck, and Phil Hauptman and his Introduction to Language class for help in preparing this paper. All errors are my own.

**REFERENCES**