ESCOL 1993

Alignment and Prosodic Heads

Janet Pierrehumbert
Northwestern University.

I) Introduction

Speech is produced and understood in time. One of the tasks of phonological theory is specifying how the elements of sound structure are temporally sequenced and coordinated. This is the task of autosegmental and metrical phonology. These historically had somewhat separate concerns, with autosegmental phonology treating issues of temporal alignment and metrical phonology treating grouping and prominence. But recently, an integrated picture has emerged, in which all levels of prosodic structure are implicated in controlling what phonological contrasts occur , when they are produced, and how they relate to the morphological structure.

This paper examines a number of post-lexical phenomena involving alignment to a head prosodic position: The Evanston variant of the Name Game; a prevalent type of blends; rhyming; and pitch accent assignment in English. The significance of these phenomena for current theory is assessed. Pitch accent assignment, a historically intractable problem, is found to yield to a constraint based approach. The analysis builds on the general ability of constraint-based theories to automatically supply prosodic structure needed to ensure well-formedness, and specifically on the central claim of Optimality Theory (Prince and Smolensky, 1993 MS) that some constraints take priority over others. It adapts the concept of relativized minimality as proposed in prosodic circumscription theory (McCarthy and Prince, 1990) to a new use. Incorporating relativized minimality into the constraint based approach is found to require the use of meta-rules (Gazdar et al. 1985).

The starting point for the discussion here will be McCarthy and Prince (in press), which offers a generalized treatment of alignment in phonology proper and in prosodic morphology. This paper is reviewed in Section II. Section III deals with the issue of how phonological information is combined after an alignment is established. Then, examples of alignment at head position are presented. Formalization of the regularities is found to require a modification of McCarthy and Prince's treatment of heads; heads are argued to function as locations, not as prosodic categories. Building on this foundation, the main regularities of pitch accent assignment are presented and analyzed.

II) Generalized Alignment

McCarthy and Prince (in press) provides an integrated treatment of many alignment phenomena occurring within lexical items. It is based on alignment functions of format [1].

[1] Align (Cat1, Edge1, Cat2, Edge2)

In [1], ``Cat1'' and ``Cat2'' are either morphological categories (such as ``Stem'' or ``the affix /um/''), or phonological categories (e.g. ``Syllable'', ``Foot'', or ``Prosodic Word''). [1] is read as ``For all Cat1, there exists Cat2 such that Edge1 of Cat1 and Edge2 of Cat2 coincide), or, more informally ``Anytime you see Cat1, you also see a Cat2 such that they are ligned up at the specified edges''. The various alignment

functions found in the grammar of a language are interpreted softly under Optimality Theory (see also Prince and Smolensky, MS). That is, a constraint can be violated in deference to a higher ranked constraint, but constraint violations are minimal. Constraints which are never violated are handled by a technical mechanism whose nature will not be critical here; they will be referred to in the rest of this paper as "absolute constraints".

The Tagalog affix -um provides an example of this approach. This affix appears as either a prefix or an infix, depending on whether the stem begins in a vowel or in a consonant.

[2] um-aral ``teach"
s-um-ulat ``write"
gr-um-adwet ``graduate"

This behavior is formalized as in [3]

[3] Align([um], L, Stem, L)

Roughly ``All instances of -um are aligned as far to the left of some stem as is consistent with all more important phonotactic constraints".In this example, the constraint against syllables with codas induces infixation. Note that not all stems have an -um; hence the universal quantifier on the first argument and the existential quantifier on the third argument.

A previous treatment of Tagalog -um is offered in McCarthy and Prince (1990, p. 227) within the framework of prosodic circumscription. Comparing the two treatments brings out the advances made in McCarthy and Prince (in press). In prosodic circumscription theory, a morphological operation (in this case, affixation of -um) applies to a base under factoring imposed by identification of a phonological constituent. That is, the circumscription divides the base into two parts and the morphological operation applies to one of the parts, which are then reassembled. For example, to derive the form ``s-um-ulat", the stem ``sulat" is divided into ``s * ulat", and -um is prefixed to the second part before the first part is reattached.

In general, the circumscription function has form [4], with three arguments.

[4] PHI(C, E, B)

B is the (morphological) base, C is a prosodic constituent type, and E is an edge at which a constituent of type C is identified. One part of the factoring is the material dominated by the C parsed out, and the other is the remainder of the base.

For the Tagalog affix -um, PHI is invoked by a morphological operator, in this case ``the operation of prefixing -um to a base". Note that the operator which invokes the circumscription implicitly provides a fourth argument beyond the three found in [4], namely the affix itself. E is the left edge, since the affix is found near the beginning of the form, and the parsed out constituent is a consonant, if any. 1)

Both the previous approach and the new one share an important insight about the status of elements of prosodic structure, such as syllables, feet, and prosodic words. These elements appear in argument positions in the alignment or circumscription function. That is, they are treated as things, on a par with other things such as tones, phonemes, or features; the structure imputed to a form is distinguished from the means used to compute it. And indeed the means used to compute the structure leave no trace apart from the structure they actually assign.

The Stem in [3] corresponds to the base B in [4]. Thus, in Generalized Alignment, the existential argument Cat2 takes over the role of the base in prosodic

circumscription. The affix triggering a circumscription appears as Cat1 (the universal) in Generalized Alignment. However, prosodic circumscription theory did not use phonological categories as bases, and provided only preliminary speculations about phonological categories as operators. One reason that Generalized Alignment is ``generalized'' is that it treats alignment of both phonological and morphological elements in a uniform framework.

A further source of generality in ``Generalized Alignment'' may be traced to that fact that in describing prosodic systems, it makes heavy use of pairs of constraints in which the values of the Cat1 and the Cat2 arguments are switched. For example, [5] would describe a prosodic system in which every prosodic word has a foot at its right edge, and feet can occur only at the right edge of a prosodic word, leaving the rest of the word, if any, unmetrified.

[5] Align (PrWd, R, Ft, R)
[6] Align (Ft, R, PrWd, R)

(In McCarthy and Prince, in press, the resources of Optimality Theory are exploited in order to permit constraint pairs such as [5] to account for more complex and interesting prosodic regularities.) In prosodic circumscription theory, the base is always treated as existential and the affix as universally quantified. [5] in effect represents the possibility of reciprocal circumscription. This possibility was in no way formally precluded in the earlier work, but a procedural outlook evidently led to its not being explored.

Nothing in [3] corresponds to argument C of [4]. This argument has been deliberately omitted because of one of the main points of the new analysis. The location of -um is not idiosyncratic, but is on the contrary predictable from considerations of syllable structure. Its placement derives from its -VC form, and a CV affix showing the same behavior (infixation after the initial onset of a consonant initial stem) is predicted to be impossible. The paper pursues the hypothesis that C will prove to be predictable in all cases, though it does not offer accounts of some of the most complex phenomena previously analyzed, most notably in Arabic and in Choctaw. (McCarthy, 1992; Lombardi and McCarthy, 1991).

In one respect which will prove crucial here, Generalized Alignment represents a significant loss of generality in comparison to prosodic circumscription theory. Prosodic circumscription provides a certain support for constraints on relations of pairs or sets of words. This can be seen from McCarthy (1992), in which some measures of the Arabic verbs are used as the bases for computing forms of other measures and of nominalizations. The alignment functions of Generalized Alignment could, in principle, support evaluation of the relationship of words to each other. Indeed this adaptation is an obvious limiting case of a morphological category; if the affix -um can be a (rather specific) morphological category, why can't a full lexical item be a morphological category? However, the theory as a whole does not support this line of reasoning because it is embedded in Optimality Theory. In Optimality Theory, a function GEN provides all candidate forms of a given lexical item, and an evaluator selects the best by running down its list of prioritized constraints, and determining which are met in which of the candidates. That is, the evaluator checks a list of alternative candidates for expression of a single lexical item. The architecture thus posits a strong separation between constraints and lexical items. Furthermore, there appears to be no provision for constraints which are specific to several related lexical items.

These two factors place Optimality Theory in constrast with Declarative Phonology (see Coleman 1991, Scobbie 1991, and Broe, 1993). In Declarative Phonology, the distinction between lexical items and constraints is minimized; both are pieces of phonological structure, with constraints being simply less specific than

individual lexical items. Lexical Phonology can also be viewed in this light. The intermediate representations created by the cycle provide a mechanism for relating the derived properties of complex forms to derived properties of their simpler morphological relatives. Insofar as the critical properties of the simpler forms actually show up on their surface realizations, the regularities handled by this mechanism do not necessarily provide any support for a derivational approach with intermediate levels of representation. On the contrary, such regularities could equally be handled by a constraint-based approach operating over sets of morphologically related words. Burzio's (1991, MS) treatment of stress is in this direction. Spreading activation models also have the potential for describing many cyclic phenomena, because of the reciprocal activation of lexical items which share phonological properties and/or elements of meaning.

In the rest of this paper, the alignment functions of Generalized Alignment will be adapted to describe a number of post-lexical phenomena for which they were not originarily designed. The description will also rely on the broad insight of Optimality Theory that some phonological constraints are more rigously enforced than others. However the analysis does not rely on the specific architecture of Optimality Theory. In fact, several modifications of Optimality Theory are proposed. Specifically, the analyses will require simultaneous evaluation of multiple lexical items; adaptation of the concept of relativized minimality from prosodic circumscription theory; and packaging of constraints using meta-rules in the style of Gazdar et al. (1985).

III) Combining Information

When an alignment between morphological or phonological entities is established, the representations must be combined. In Generalized Alignment, the categories may occur either in sequence or simultaneously. For example, the phonemes of the Tagalog affix -um are sequenced with respect to the phonemes of the stem. A metrical foot, in contrast, will be produced simultaneously with the material to which it is aligned. Whether the combination is sequential or simultaneous can be derived from general principles of well-formedness. A typology of cases handled by the theory may then be laid out as follows; each table entry gives an example of the type of case covered in that category.

[7]	Simultaneous	Sequential
Morpheme-Morpheme		Affixation
Prosody-Morpheme/ Morpheme/Prosody	Alignment of PrWd with Root	Affixation at head foot
Prosody-Prosody	Footifying PrWd	Iteration of parsing

Generalized Alignment offers no examples of morpheme-to-morpheme alignment with the morphemes produced simultaneously, despite the fact that Prince and Smolensky (MS) provides groundwork. Cases of simultaneous morphemes would include not only the case of tonal morphemes assigned to phrases (which will be discussed below), but also morphologically distinctive umlaut or nasalization. However the extension of the approach is obvious and will be presupposed in the discussion that follows. In brief, alignment is established exactly as in other cases, and the morphemes are produced simultaneously just in cases in which independent principles of well-formedness permit this.

For post-lexical processes, a strong case can be made for yet a third way of combining the information. As argued by Shattuck-Hufnagel (1979) speech error data indicate the existence of a buffer, representing the plan for an actual utterance, into which the phonological content of the words in the utterance is transferred.

into which the phonological content of the words in the utterance is transferred. Without such a buffer, it is extremely difficult to speech errors involving transpositions (spoonerisms) Such a model is further supported by other data on speech production discussed in Levelt (1989) and Sternberg et al (1980).

The existence of a post-lexical buffer is also supported by the linguistic evidence, since the phrasal prosody must be constructed somewhere. Since the phrasal prosody constructed over word tokens does not permanently impact the lexical representations of the word types, it is clear that this construction must take place outside of the lexicon.

In the presentation of examples which follows, the existence of a post-lexical buffer will be exploited in two ways. First, we suggest it is the basis for word games and blends such as

[8] brown + green --> breen

Following Kubozono (1990), we view such blends as arising when one lexical item is interrupted by another. In deliberately formed blends, the result may of course be lexicalized as a new item. Second, the assumption that the buffer is the location for the construction of phrasal prosody means that words in the lexicon must be aligned with prosodic positions in the buffer in the course of compiling an utterance plan. This alignment in effect means that words are parsed twice; each word type has its lexical prosody, and in the course of being aligned with the phrasal prosody, each word token is also assigned the exact prosody it has in any given utterance. Of particular interest in this regard is the suggestion in Sternberg et al. (1980) that the units of the buffer are metrical feet; examination of the materials in their experiments shows that the number of units could equally well be taken to correspond to the number of pitch accents. The parse needed to establish pitch accent locations will be the most important example discussed below.

Some of the post-lexical phenomena discussed below are closely analogous to lexical phenomena involving prespecification or over-writing. (see Steriade, 1988). We do not intend to imply that the mechanisms of aligning items with a buffer, or interrupting one item with another, are confined to the post-lexical phonology, but merely that in view of the psycholinguistic evidence the case for these mechanisms in the post-lexical phonology appears to be particularly strong. As the theory of Lexical Phonology has made plain, lexical rules often grammaticalize post-lexical allophony. It would hardly be surprising if the resemblance between lexical and post-lexical processes went beyond particular segmental rules to general cognitive mechanisms.

IV) Examples of Alignment at Head Position

This section introduces a number of phenomena involving alignment at head position.

The ``Name Game" (Perry Ellis, 1967) is still played by children around the country. [9] provides an example of how the game is played in Evanston.

[9] Jennifer, bobennifer, fa fana fana fo fennifer ...

In order to determine the rule for the game, two informants, ages 6 and 4, were interviewed in an open response format. That this, they were asked to provide game forms for assorted names, and their responses were recorded without comment. The critical examples are as in [10], with a vertical slash used to indicate where each name is being partitioned to construct the game form.

[10] Fr | omkin, bobomkin ...
McC | arthy, bobarthy ...
Mon | adnock, bobadnock ...
Pierreh | umbert, bobumbert ...
Alex | andra, bobandra ...
Massach | usetts, bobyusetts ...

The base form is partitioned at the head vowel regardless of where it is in the word. This pattern contrasts with that reported in Hammond (1990), in which only trochaic stems could be used in (a different version of) the Name Game. In the Evanston version, the partition of the base form may have a nonconstituent for the first part, the second part, or both: for example, neither ``Pierreh" nor ``umbert" is a constituent under any theory.

This game could not be described using prosodic circumscription, since (in the absence of a significant extension) prosodic circumscription would require identification of a constituent at an edge. However, using the approach of Generalized Alignment, the game may be provisionally described as: All [bob]s are aligned at their right edge with the left edge of the head vowel of a prosodic word. Combination of the two items is achieved by the interrupted word method.

As argued in Pierrehumbert (1980), and Beckman and Pierrehumbert (1986), intonation contours in English consist of one or more pitch accents plus boundary tones marking phrasal edges. Each pitch accent is aligned with the most prominent syllable of its domain. For example, when a single pitch accent is assigned to a compound, its location is the main stress of the compound:

[11] Brick warehouse (Warehouse made of bricks)
H*

[12] Brick warehouse (Warehouse for bricks)
H*

In this theory, accents consist of either one or two tones. Bitonal accents have one tone marked as the head tone, and the position of the head tone is distinctive. Using a * to indicate the head tone, this yields a total of four bitonal accents: L+H*, L*+H, H*+L, and H+L*. An experiment reported in Pierrehumbert and Steele (1990) demonstrates that L+H* and L*+H are categorically distinct. That is, there is a categorical distinction between [13] (which contains a relatively early rise) and [14] (in which a delayed rise of the same general shape is assigned to the same syllable as in [13]).

[13] Only a millionaire.
|
L+H*

[14] Only a millionaire.
|
L*+H

A formulation of the meaning difference between [13] and [14] is provided in Pierrehumbert and Hirschberg (1990); intuitively, [13] would convey the speaker's commitment to the proposition, whereas [14] would signal lack of commitment, which could arise from incredulity or other reasons.

Deferring until Section VI the issue of what the phonological domain for a pitch accent actually is, we note that [13] and [14] display head-head alignment. That is, the head of the pitch accent is aligned with the prosodic head of the associated text.

In blends, the production of one lexical item is interrupted by another. Kubozono reports tallies on the switch points for the blends in the 1934 Wentworth corpus. For monosyllabic blends, 77% switched at the boundary between the onset and the peak. That is, the predominant switch point is at the head of the syllable. Kubozono's own interpretation that the switch point is after the onset constituent is unavailable in moraic theory, which has no onset constituent. It is further undercut by experimental results reported in Pierrehumbert and Dhingra-Nair (in prep.). Pursuing this issue, all the polysyllabic blends reported in the corpus of speech errors in Fromkin (1973) were tallied according to where the switch point occurred. In many, the switch point is of course ambiguous; for example in [15], the switch point could be either before or after the /l/.

[15] splinters + blisters --> splisters

However, approximately half the blends were consistent with postulation of a switch point at head position. Although the head position is by no means the only possible switch point, this would be a disproportionate number in view of the large number of phonemes in polysyllabic words.

In poetry, monosyllabic words and other words with final stress rhyme exactly if they match from the beginning of the vowel to the end of the word. This type of rhyme is referred to as a ``masculine rhyme"; in a feminine rhyme, words with nonfinal stress match from the stressed vowel to the end of the word, as in the following example from Byron:

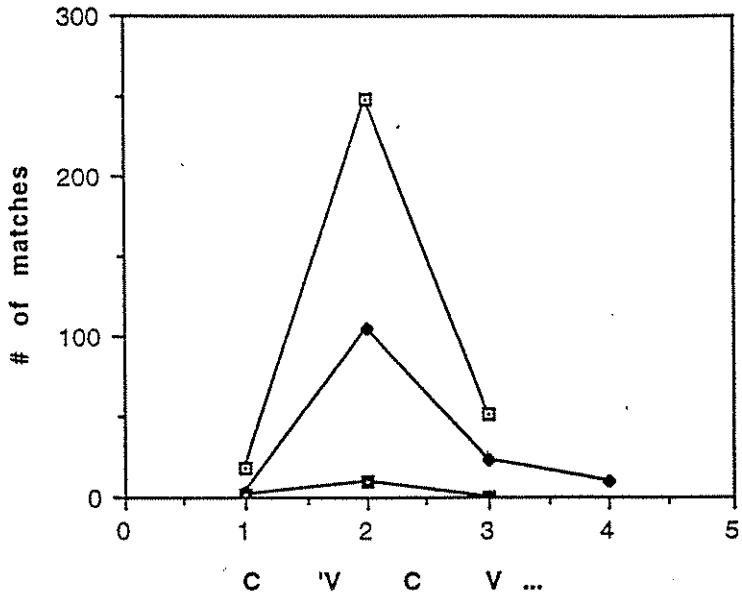
[16] ... Adeline
... so sad a line.

(In [16], note how the intensifier has been used to force the phrasal stress onto the antipenultimate syllable, indicating that rhyme is defined over line endings and not merely over word endings.) In inexact rhymes or near-rhymes, the correspondant vowels are not actually identical, though typically they are related in quality. For example, ``loud" and ``Maude' constitute a near-rhyme, since they have similar but not identical qualities for the stressed vowel, but match exactly over the post-vocalic region (which in this case is just a single consonant.)

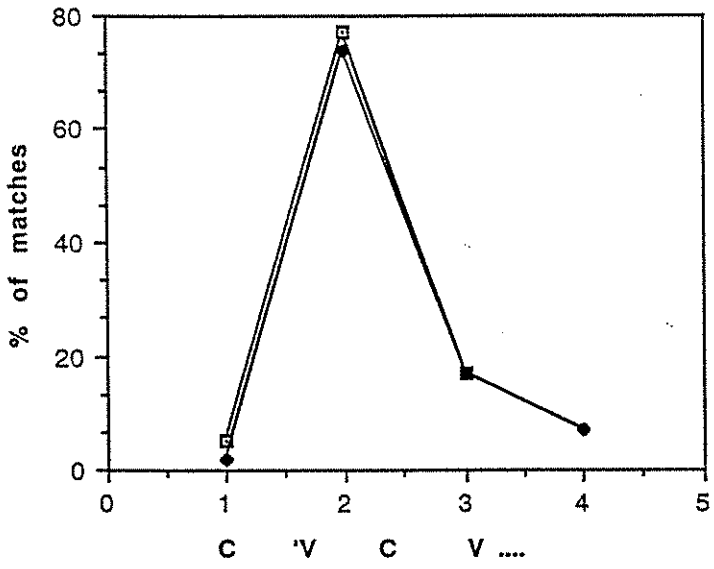
In order to evaluate the relationship of masculine rhyme to feminine rhyme, all the rhymes (over 400) in Canto XIV of Byron's Don Juan were coded according to where the region of exact match between rhymed words began. In a few cases, the match actually begins before the stressed vowel, e.g. ``approach, reproach". Near-rhymes were coded as mismatched on the stressed vowel, in other words matching from the following consonant onwards. Byron's work was selected for this coding because he is an extremely promiscuous rhymer whose comical style includes many feminine rhymes.

The results are displayed in [17] and [18] . [17] shows rhymes coded according to the total number of examples of each type; open boxes stand for masculine rhymes, and diamonds and filled squares stand for feminine rhymes involving (respectively) penultimate and antipenultimate stress. In [18] the two common types have been replotted by percentage of cases of each type in which the exact match starts at each given point.

[17]



[18]



Examination of these graphs leads to the following observations. First, matches beginning earlier than the stressed vowel are very infrequent. Second, feminine rhymes are less common than masculine rhymes; this is not surprising since they involve statistically less likely matches of longer material than masculine rhymes, and are therefore in some sense more difficult. Despite this increased difficulty, however, feminine rhymes are no more likely to be imperfect than masculine rhymes; the proportion of near-rhymes is almost identical for the two

cases. In short, feminine rhymes are not merely broadly analogous to masculine rhymes, but are literally the same thing, as exhibited by a detailed pattern of variability. This means that rhymes cannot be characterized as involving the match of a phonological constituent of any type. Masculine rhymes might be characterized in terms of the rhyme constituent of a syllable (though even this characterization runs afoul of cases involving extrasyllabic final consonants.) However the matching region in a feminine rhyme cross-cuts the syllable structure, and is not a phonological constituent in any theory. Therefore, rhymes must be characterized in terms of an alignment of words (or lines) at head position; a match is enforced from the alignment point onwards.

V) Formalizing Alignment at Head Position

In McCarthy and Prince (in press), heads are treated as prosodic categories. For example, constraint [19] is used to describe the behavior of the Ulwa affix [ka], which is placed after the head foot.

[19] Align([ka], L , Ft', R)

However, this formulation does not bear scrutiny, and we propose instead to treat headedness as a location, analogous to the left and right edges. There are three reasons for this proposal.

First, as shown in Liberman and Prince (1977) and discussed at length in Pierrehumbert and Beckman (1988), the phonologically available locations in any phonological structure are the edges (or the peripheral locations) and the head, if defined. ``Prosody" covers all aspects of temporal sequencing, grouping and prominence; reference to the left or right edge is a reference to a prosodic location and should be kept on a par with any reference to the other privileged prosodic location, namely the head.

More specifically, although the arguments L and R appear unadorned in the second and fourth places of the Align, it is important to understand that they are actually functions. That is, [20] actually means [21].

[20] Align([um], L, Stem, L)
[21] Align([um], Left([um]), Stem, Right(Stem))

Even more explicitly, we may say that in

[22] Align(Cat1, Edge1, Cat2, Edge2)

Edge1 is a function which returns the location of the edge of type Edge1 in every token of type Cat1. Similarly, Edge2 returns the location of the edge of type Edge2 in a token of type Cat2. This may be compared to the designated terminal element in Liberman and Prince(1977); though it too has the appearance of a thing, it is found by recursively evaluating headedness in the prosodic structure, and the recursive evaluation is a function.

There are also more substantive reasons for this proposal. Consider first that in blending, the head functions like an edge. Compare:

[23] Dem | ocrat + Rep | ublican --> Demublican
sm | oke + f | og --> smog

[24] Derriere | + Der | ida --> Derrierida
Cal | ifornia + M | exico --> Calexico
Pierre | humbert and | Hirschberg --> Pierrehirschberg

In [23], the relevant location in the right hand member of the blend is Left(dte), that is the left edge of the head vowel. But the right edge of the left hand member has no coherent definition: Right(????). Rather, we wish to say that the first member leaves off at exactly the same location where the second member begins, namely Left(dte). This follows from the fact that all relevant internal locations act like initial and final edges in blending, as illustrated in [24].

Secondly, consider the case of head-head alignment as exemplified by pitch accent alignment. Adopting the original format of Generalized Alignment would yield [25] as a function controlling the alignment of L*+H.

[25] Align(Head[L*+H], Edge1, Head(Prosodic domain), Edge2)

But what are Edge1 and Edge2? The format requires a selection between L and R, but there is no basis for choosing since the two heads are produced simultaneously. That is, the original format does not extend coherently to this situation. It would be better to say something like:

[26] Align([L*+H], Head, Prosodic Domain, Head)

Taking this approach, the other examples introduced in Section IV are taken to involve alignments as follows. In the Evanston form of the Name Game, the right edge of "bob" aligns with the left edge of the head vowel, or dte, of the name (likewise for the right edge of "fa fana fana fo f"). Taking the name to be a prosodic word gives:

[27] Align([bob], R, PrWd, L(dte))

With the alignment established, the phonological forms are combined by switching from the "bob" to the PrWd at the specified alignment point.

Rhymes involve alignment of words or lines, at head position, as in [28].

[28] Align(Word, dte, Word, dte)

Note, however, that in the case of rhyming, the reading of this function with universal and existential quantifiers is perhaps somewhat forced. That is, it is not the case that for every word there exists another word which rhymes with it. Some words simply have no rhymes, because the region from the head to the end is unique. The universal-existential reading can be enforced only by confining the quantifiers to particular verse forms. Thus, one might say: "In rhymed couplets, for every word ending an even numbered line, there exists a word ending the next line which is lined up at dte position". However, since limericks have an AA BB A rhyme scheme, they would require a different statement. Obviously something is being lost; we want to talk about rhyme *in general*.. One way to do this is note that [28] can be responsible only for the alignment, not for the adequacy of the match; it says nothing about what happens after the head. As such, it can be read with a double universal "For any word, pick any other word and line them up at head position. Then you can find out if they rhyme."

VI) Pitch Accent Assignment in English

One of the earliest and most intractable problems in prosodic morphology is that of pitch accent assignment in English. The problem is one in morphology because intonational elements convey an assortment of pragmatic meanings, in a manner broadly analogous to pragmatic particles. In short, intonational elements are morphemes (or possibly lexical items) whose phonological content happens to be purely tonal. This point already emerges in Liberman and Sag (1974), Sag and Liberman (1975) and Liberman (1975). It is developed in much subsequent literature including Gussenhoven (1983), and Ladd (1980). Pierrehumbert and Hirschberg (1990) advance the theory that meanings of complex intonation patterns arise compositionally from meanings identified for the various pitch accents and the boundary tones.

Pitch accent assignment in English is a prosodic problem because, as noted in Section IV, pitch accents are aligned with the prosodically strongest syllable of their domain. The position of boundary tones is also prosodically determined, but by grouping rather than prominence. The main difference between pitch accent assignment and what are considered to be standard examples of prosodic morphology is that standard examples involve lexical processes. Pitch accent is post-lexical; the well formedness of the accent placement is determined with respect to the intonation phrase, a prosodic unit which can only be formed after words are combined into sentences.

The problem has been intractable because the major regularities governing accent placement have proved difficult to express in several successive versions of autosegmental and metrical phonology, beginning with Liberman (1975). The primary observation is that accents go on the strongest prosodic positions in a phrase. Specifically, the weakest syllable which can bear an accent is the head of a foot. If some syllable in a phrase carries a pitch accent, then all syllables in the phrase which are metrically as strong or stronger also carry a pitch accent. The syllable carrying the last accent in the phrase is the most prominent element of the phrase; that is, the last accent falls on the dte of the entire phrase. There is also a less strict preference for an accent near the beginning of the phrase. Thus, the patterns in [29] are acceptable (though by no means the only acceptable patterns for the texts shown); the patterns in [30] are unacceptable. (In both [29c] and [29d], metrical grid levels below the prosodic word have been omitted for the sake of legibility .)

[29a]

```
          x
x         x
x   x     x
x  xx  x  xx
Appa lachi cola
|          |
H*         H*
```

b]

```
            x
x   x       x
x  x x  x  x x
Appa lachi cola
            |
            H*
```

c]

```
x              x                x
x              x      x   x     x
Organized on the model of eleven gallons of worms
|              |                |
L*+H           L*+H             L*+H
```

d]

```
                                                      x
    x                         x                       x
    x                x        x      x                x
    Organized on the model of eleven gallons of worms
    |                         |                       |
    L*+H                      L*+H                    L*+H
```

[30a]

```
    x               x
    x               x
    x     x         x
    x   x x   x    x x
    Appa lachi cola
          |       |
          H*      H*
```

b]

```
                  x
    x     x       x
    x  x x   x  x x
    Appa lachi cola
    |
    H*
```

c]

```
                              x
    x                         x                       x
    x                x        x      x                x
    Organized on the model of eleven gallons of worms
    |                         |                       |
    L*+H                      L*+H                    L*+H
```

[30a] is bad because the first stressed syllable is not accented even though the metrically weaker medial stressed syllable is accented. [30b] is bad because the last accent precedes the dte of the phrase. [30c] is bad because the last accent follows the dte of the phrase.

A further complicating factor is the the position of the prosodic head of the phrase -- and indeed of all of the more prominent elements -- varies with respect to the edges because of focus effects. This point may be illustrated by the following extract from Huckleberry Finn, in which the emphasis has been supplied by Twain.

[31] JIM: 'kase he say dey AINT no witchs [...] I jis' wish to goodness he was heah now -- DEN what would he say.

As a result of this variability, edges and heads are not interconvertible in the phonological representation, in contrast to cases examined in Prince (1983). That is, it is essential to support independent phonological reference to heads and to edges.

Pitch accent assignment in English presents perhaps the worst outstanding problem for the level-ordering hypothesis advanced in Selkirk (1984), Nespor and Vogel (1986) and Beckman (1986). Although a level of "accentual phrase" can be defined in some other languages such as Japanese (see e.g. Poser., 1985, Beckman, 1986, and Pierrehumbert and Beckman,1988) it is impossible to do so for English. The difficulty arises because the accentual phrase can be either bigger than a prosodic word, as in [32], or smaller than a prosodic word, as in [33].

[32] [Organized] [on the model of eleven] [gallons of worms]
H* H* H*

[33] [Consti][tutional]
H* H*

The fact that "constitutional" remains a single prosodic word even when it has two pitch accents is shown by the lack of either phonological tensing or phonetic lengthening on the unstressed vowel ending the first putative phrase. In short, the evidence about where in the prosodic hierarchy the accentual phrase should be inserted is contradictory, and as a result it is impossible to place it at any particular level in the hierarchy.

Prosodic promotion is generally available to strengthen prosodically weak elements if the speaker for any reason wishes to accent them. In the attested example [34], the unstressed final syllable has been promoted all the way to the main word stress so that it can carry the last and therefore strongest accent:

[34]
x
x x x x
x x x x x
Un fortu nate ly !
H* H* H* H*

In the rejoinder of [35], the light variant of the article "the", /ðə / is replaced by the heavy variant / ðɨ i y/ even though it is preconsonantal. The replacement is dictated by the contrastive accent on the article.

[35] -- A Prof. John McCarthy is speaking this evening.
-- You mean the /ðə/ John McCarthy from U. Mass.?
THE /ðɨy/ John McCarthy?

One of the phonological alternations which figured prominently in the development of metrical theory is the so-called "Rhythm Rule", illustrated by the different pronunciations of the word "Massachusetts" in [36a] and [36b]

[36 a] b]

x
x x
x x x x
Massachusetts

x x
x x x
x x x x xxx
Massachusetts miracle

Although treated in Liberman and Prince (1977) as a purely rhythmic phenomenon, more recent experiments and studies of continuous speech (Beckman et al, 1990; Ross et al, 1992) have shown that this is incorrect. In fact, the phenomenon is directly tied to accent placement. The difference between pronunciations of types [36a] and [36b] is only made in the presence of accent. This discovery is not surprising, since the citation forms which were the basis for the introspective data reported in the earlier papers were almost surely produced as single intonation

phrases with two accents. The instrumental studies further establish that forms of type [36b], with early placement of the accent, are found far more freely in prenuclear position than the introspective data had suggested. The historical background for this conception of the rhythm rule is usefully reviewed in Gussenhoven (1988).

A last observation might be viewed as an observation about the field of intonational analysis rather than as an observation about the data. This is the fact that there has been a protracted and unsatisfying dispute about whether intonation phrases may have more than one pitch accent. This dispute has taken place both within and between schools of phonology; compare, for example, the positions of Halliday (1967), Crystal (1969), Ladd (1980), Pierrehumbert (1980) and Gussenhoven (1983). However, vexed questions are often vexed exactly because they are ill-posed. In the analysis about to be described, each accent "sort of" has its own intonation phrase, and judgments about whether the phrase is present or not would be expected to be uncertain.

These varied and perplexing regularities can be integrated by bringing together two key ideas. First, we accept the joint position of both Optimality Theory and Declarative Phonology that phonological representations result from the combined consequences of constraints on their form. Specifically, we take from Optimality Theory the notion that meeting some constraints can take priority over meeting others, with the consequence that low priority constraints may be active in the grammar while still being violated in some forms. Second, we adopt and extend the concept of relativized minimality as developed in prosodic circumscription theory. Since this concept has not yet been introduced, here it is in brief.

McCarthy and Prince (1990) exploit heavily the concept of the minimal prosodic word, Wmin, which must consist of at least one foot, which in turn must be at least a bimoraic syllable. The interpretation of Wmin in prosodic morphology is variable, depending on the extent to which its minimization is specified. For example, English nickname formation takes as its base the fully minimized prosodic word (or heavy monosyllable):

[37] Patricia ---> Pat --> Patty

In contrast, in Yupik, the proximal vocative is freely monosyllabic or disyllabic, reflecting a Wmin minimized to the foot level but not all the way to the syllable level. That is, the proximal vocative form takes any otherwise established foot.

Extending this concept, we propose a minimal intonation phrase iMin; this would be the "minimal intermediate phrase" in the terminology of Pierrehumbert and Beckman (1988). iMin is now a variable which ranges over the following hierarchy of prosodic types:

[38]

Intermediate phrase	i
<Phrasal grid levels>	.
	.
Prosodic word	PrWd
Foot	Ft
Bimoraic syllable	σ
	/\
	μ μ

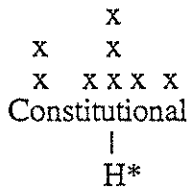
Note that since a prosodic word must have at least one foot, so must iMin. It is extremely important to be clear on a technical point which is not especially clear in

McCarthy and Prince (1990); neither Wmin nor iMin is a prosodic category, in the sense of occupying a determinate place in a level-ordering scheme. Rather, Wmin and iMin are variables which range over prosodic categories. In any particular application, a value of the variable will be selected either freely or on some principled basis. This value is what appears in the phonological representations; neither Wmin or iMin will appear as a node label in the representation of particular forms.

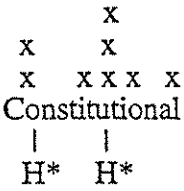
The core of the proposal for describing the alignment of pitch accents is that each intermediate phrase is exhaustively parsed into iMins under relativized minimality. One pitch accent is assigned to each iMin at head position. This reparsing (or prosodic parsing of lexical items which already have a prosodic structure assigned to them) arises from the process of aligning lexical items with prosodic positions in the post-lexical buffer. The choice of degree of minimization is phonologically free, though when other levels of linguistic description are considered, the choice can be seen to reflect discourse factors.

This description amounts to aligning pitch accents to elements prominent at any grid level at or above the foot; if a foot is established over a monosyllable, the monosyllable must either be bimoraic or be coerced into bimoraic form. Simple examples are provided in [39]; in [39a], the instantiation of iMin is PrWd. In [39b], it is the foot.

[39 a]



b]



In establishing the alignment, the following constraints are inviolable; they are met in every well-formed intermediate phrase. First , every pitch accent is aligned with the head of an iMin and every iMin has a pitch accent.

[40] Align([<pitch accent>], dte, iMin, dte)
Align (iMin, dte [<pitch accent>], dte)

Note that [40] abbreviates an effective conjunction in the Cat1 position and an effective disjunction in the Cat2 position. The second constraint in the pair actually means:

[41] Align(iMin, dte, [H*] or [L*] or [L*+H] ..etc, dte)

Second, actual (as opposed to relativized) intermediate phrases have their strongest prominence at the last accent, that is, on the last iMin. This is an example of an End Rule (cf Prince, 1983; Prince and Smolensky, MS). However, standard examples of an End Rule involve promoting the head of the rightmost or leftmost instance of a prosodic category to be head of the immediately superior level in the prosodic hierarchy. For example, in the classic understanding of the End Rule governing word stress in English, the last foot is made the head of the immediately superior category, the PrWd. The iMins in i, however, are not necessarily the immediate daughters of i, and the End Rule can therefore promote the head of the last iMin by an arbitrary number of grid levels. Formulating the constraint has a delicate dependence on the technical treatment of peripherality and headedness. As pointed out in footnote 44 of McCarthy and Prince (in press), the paper does not

actually supply such a technical treatment. Details of the End Rule for iMins in i will be deferred to a future paper.

Three violable constraints also figure in the patterns of pitch accent placement described. First, head prosodic words in i like to have a pitch accent near the right edge:

[41] Align(i, R(HeadPrWd), [<pitch accent>], R?)

Second, iMins like to have an accent near the left edge:

[42] Align(iMin, L , [<pitch accent>], L?)

[41] and [42] together are responsible for the metrical alternation found in words like "Massachusetts"; the fact that "Massachusetts" is pronounced with more prominence on the rightmost foot when in nuclear position reflects the precedence of [41] over [42].

Third, some type of scansion constraint appears to be required to explain the preference for nice alternating patterns in extended materials.

The exhaustive parse of i into iMins, with a uniform choice of iMin, successfully accounts for the fact that if any syllable in a phrase is accented, all syllables which are equally strong or stronger are also accented. It does this without positing an accentual phrase, which as shown above is highly problematic for English. Furthermore, the analysis captures the observation that accents seem to define phrases even when other accoutrements of phrasing are not present. This follows from the fact that a kind of virtual phrasing is set up, which does not however appear as a type of node or grouping in the phonological representation. The fact that syllables can be prosodically promoted to an arbitrary extent to bear either prenuclear or accent follows from the ability of Optimality Theory or any other constraint based approach to supply missing prosodic structure in order to ensure well-formedness. Similarly, the possibility of expanding a light syllable to heavy in order to support an accent can follow from a constraint of the FILL type, as posited in Optimality Theory to cover prosodically induced epenthesis.

The analysis critically depends on a uniform instantiation of iMin, both during the iteration which parses i into iMin, and across all other constraints contributing to the well-formedness of the phrasal prosody and intonation. Technically speaking, this uniformity cannot be taken for granted; for example, when PrWds are parsed into feet, the feet need not in general be of the same type. Rather, any feet which are permissible in the language may be assembled back-to-back to parse the word. In fact, Optimality Theory has no way of enforcing a uniform instantiation; all constraints are formally self-contained, and the only relationship defined over them is the prioritization with which they are enforced.

In order to formalize the analysis, it is accordingly necessary to strengthen the support for defining relationships over constraints. The most straightforward way to do this is to exploit the concept of meta-rules, presented in Langendoen (1976) and greatly developed in GPSG (Gazdar et al, 1985). That is, constraints involving iMin or any other variable are not taken to be constraints per se, but rather schemata for specifying constraints. Instantiating the variable with some particular value yields the actual constraint. To ensure a uniform interpretation of iMin across several constraints, the constraints in question can then simply be packaged using logical AND:

[43] Align ([<pitch accent>], dte, iMin, dte)
AND Align(iMin, dte, <pitch accent>, dte) AND ....

Or, out of deference to Optimality Theory we might say that the conjunction is actually "IS MORE IMPORTANT THAN AND". iMin then receives a consistent interpretation for the same reason that all instances of any variable in a mathematical formula are consistently interpreted.

VII) Conclusion

We have presented and analyzed a number of post-lexical phenomena involving alignment to a head prosodic position, modifying the framework of McCarthy and Prince (in press). Analysis of some of the phenomena requires stating alignment constraints of lexical items with respect to each other. This is a natural use of alignment functions, since lexical items can be viewed simply as extremely specific categories. However, it is not currently supported in Optimality Theory. The second significant conclusion is that head position functions as a location, like Left and Right. Third, we have argued that the use of relativized minimality resolves the problematic relationship between pitch accents and levels of the prosodic hierarchy. But it does so at the expense of using meta-rules to ensure uniform interpretation of a prosodic variable.

Footnotes

* This work has benefited from my discussions with Mary Beckman, Michael Broe, Morris Halle, John McCarthy, and Donca Steriade.

1) This account fails to cover forms such as /gr-um-adwet/, in which the affix follows an entire onset rather than a single consonant.Treating the onset as a prosodic constituent which can be invoked by prosodic circumscription or other processes has underdesirable ramifications for the moraic theory of the syllable. This flaw in the circumscriptional analysis is now rectified in the new treatment.

References

Beckman, M. (1986) Stress and Nonstress Accent. Foris, Dordrecht.

Beckman, M., M. Swora, M. Rauschberg and J. DeJong (1990) "Stress Shift, Stress Clash, and Polysyllabic Shortening in a Prosodically Annotated Discourse," Proceedings of the International Congress on Spoken Language Processing, Vol I, 5-8.

Beckman, M. and J. Pierrehumbert (1986) "Intonational Structure in Japanese and English," Phonology Yearbook III, 15-70.

Broe, M. (1993) Unification Based Phonology. PhD dissertation, Univ. of Edinburgh.

Burzio, L. (1991) "On the Metrical Unity of Latinate Suffixes," Escol 10.

Burzio, L (MS) "English Stress, Vowel Length, and Modularity," Johns Hopkins Univ.

Coleman, J. (1991) Phonological Representations -- Their Names, Forms and Powers. Ph.D dissertation, University of York, UK.

Crystal, D. (1969) Prosodic Systems and Intonation in English. Cambridge University Press, London.
Fromkin, V. (1973) Speech Errors as Linguistic Evidence. Mouton, the Hague.
Gazdar, M., E. Klein, G. Pullum and I. Sag (1985) Generalized Phrase Structure Grammar. Harvard Univ. Press, Cambridge MA.
Goldsmith, J. (1990) Autosegmental and Metrical Phonology. Basil Blackwell, Oxford.
Gussenhoven, C. (1983) On the grammar and semantics of sentence accents. Foris, Dordrecht.
Gussenhoven. C. (1991) "The English Rhythm Rule as an accent deletion rule," Phonology 8:1, 1-37.
Halliday, M. A. K. (1967) Intonation and Grammar in British English. Mouton, the Hague.
Hammond, M. (1990) "The Name Game and Onset Simplification," Phonology 7:1, 159-163.
Kubozono, H. (1990) "Phonological constraints on blending in English as a case for phonology-morphology interface," Yearbook of Morphology, 1-20.
Ladd, D. R. (1980) The Structure of Intonational Meaning. Indiana Univ. Press, Bloomington.
Langendon, T. L. (1976) "On the weak generative capacity of infinite grammars," CUNY Forum 1, City University of New York, 13-24.
Levelt , W. J. M. (1989) Speaking. A Bradford Book, MIT Press, Cambridge.
Liberman, M. Y. (1975) The Intonation System of English. Ph.D dissertation, MIT; reprinted by Garland Press.
Liberman, M. Y. and A. Prince (1977) "On stress and linguistic rhythm," Linguistic Inquiry 8:2, 249-336.
Liberman, M. Y. and I. Sag (1974) "Prosodic form and discourse function," Papers from the 10th Regional Meeting of the Chicago Linguistic Society, 416-426,
Lombardi, L. and J. J. McCarthy (1991) "Prosodic Circumscription in Choctaw Morphology" Phonology 8, 37-71.
McCarthy, J. J. (1992) "Template Form in Prosodic Morphology" Papers from the 3rd Meeting of the Formal Language Society of Mid-America, Northwestern Univ.
McCarthy, J. J. and A. Prince (1986 MS) Prosodic Morphology, Univ. of Massachusetts and Brandeis Univ.
McCarthy, J. J. and A. Prince (1990) "Foot and Word in Prosodic Morphology: The Arabic Broken Plurals," NLLT 8:209-282.
McCarthy, J. J. and A. Prince (in press) "Generalized alignment," Yearbook of Morphology.
Nespor, M. and I. Vogel (1986) Prosodic Morphology. Foris, Dordrecht.
Pierrehumbert , J. (1980) The Phonology and Phonetics of English Intonation. Ph.D dissertation, MIT. Reproduced by Indiana Univ Linguistics Club, Bloomington.
Pierrehumbert, J. and M. Beckman (1988) Japanese Tone Structure. Linguistic Inquiry Monograph 15, MIT Press, Cambridge.
Pierrehumbert, J. and R. Dhingra-Nair (in preparation) "Novel Word Games and Syllabic Structure".
Pierrehumbert, J. and J. Hirschberg (1990) "The Meaning of Intonation in the Interpretation of Discourse," in P. Cohen, J. Morgan, and M. Pollack, eds. Intentions in Communication, MIT Press, Cambridge MA. 271-311.
Pierrehumbert , J. and S. Steele (1990) "Categories of Tonal Alignment in English," Phonetica, 181-196.
Poser, W. (1985) The Phonetics and Phonology of Tone and Intonation in Tokyo Japanese. Ph.D dissertation, MIT.

Prince, A. (1983) "Relating to the Grid," LI 14, 19-100.
Prince, A and P. Smolensky (1993 MS) Optimality Theory: Constraint Interaction in Generative Grammar. Rutgers University and Univ. of Colorado at Boulder.
Ross, K. , M. Ostendorf, and S. Shattuck-Hufnagel (1992) "Factors Affecting Pitch Accent Placement in English," International Congress on Spoken Language Processing, 365-368.
Sag, I. and M. Liberman (1975) "The intonational disambiguation of indirect speech acts," Papers from the 11th Regional Meeting of the Chicago Linguistic Society, 487-497.
Scobbie, J. (1991) Attribute-Value Phonology. Ph.D dissertation, Univ. of Edinburgh.
Selkirk, E. (1984) Phonology and Syntax: the Relation between Sound and Structure. MIT Press, Cambridge.
Shattuck-Hufnagel, S. (1979) "Speech errors as evidence for a serial order mechanism in sentence production," in W. E. Cooper and E. Walker, eds. Sentence processing: psycholinguistic studies presented to Merrill Garrett. Lawrence Erlbaum, Hillsdale, NJ.
Steriade, D. (1988) "Reduplication and syllable transfer in Sanskrit and elsewhere," Phonology 5, 73-156.
Sternberg, S., E. Wright, R. Knoll and S. Monsell (1980) "Motor programs in speech; additional evidence," in R. A. Cole, ed. Perception and Production of Fluent Speech, Lawrence Erlbaum, Hillsdale, NJ.
Woodbury, A. (1985 MS) "Meaningful Phonological Processes: A Consideration of Central Yupik Eskimo Prosody," Univ. of Texas, Austin.