Diacritic Weight in the Extended Accent First Theory

Alexandre Vaxman
alexandre.vaxman@uconn.edu
Diacritic Weight in the Extended Accent First Theory

Abstract
In this article, I present the Extended Accent First theory, which is an offshoot of the Primary Accent First theory (van der Hulst 1996, 1997, 1999, 2010). While the latter is known to correctly account for accent location in a large variety of languages, it encounters difficulties accounting for lexical accent systems and systems sensitive to both phonological weight and lexical accent. The Extended Accent First theory makes such an account possible. In this theory, lexical accent is reanalyzed as “diacritic weight”, leading to the notions of “diacritic weight scale” and “hybrid weight scale”. The Extended Accent First theory is illustrated here with a case study from Central and Southern Selkup that shows how the theory works and, in particular, how it can account for dominance effects using a diacritic weight scale. A comparison of the Accent Deletion approach vs. the Extended Accent First theory with respect to accentual dominance suggests that the approach proposed here is more straightforward and economical. Interestingly, the existence of phonological and diacritic weight correctly predicts that there are accent systems which make reference to both weight types (ordered in a single language-specific weight scale).
Diacritic Weight in the Extended Accent First Theory

Alexandre Vaxman*

1 Introduction

This article presents the Extended Accent First theory (the “ExtAF” theory), a new theory of word accent. The ExtAF theory is an offshoot of the (Primary) Accent First theory (abbreviated as the “PAF”, or “AF”, theory) originally put forth by Harry van der Hulst and his colleagues in the 1990s (van der Hulst 1996, 1997, 1999).

The ExtAF theory follows the AF theory in that it separates word accent (primary stress) from rhythm (non-primary stress), assigning these separately and without using feet (as amply motivated in van der Hulst 2010, Goedemans and van der Hulst 2014). In this article, I will only consider word accent.

The AF theory correctly accounts for accent location in a large variety of languages, but encounters difficulties with lexical accent systems and systems that combine phonological weight with lexical accent (which I call “hybrid systems”). In this article, I will show how the innovative aspects of the ExtAF theory enable it to address these issues.

The article is organized as follows. Section 2 describes the main ingredients of the ExtAF theory. Section 3 is a case study in accentuation of Central and Southern Selkup varieties which shows that the ExtAF theory is able to straightforwardly account for accent assignment in a lexical accent system and, in particular, for dominance effects. Section 4 argues that the account of dominance in the ExtAF theory is superior to the Accent Deletion account proposed in some lexical accent theories. Finally, Conclusion sums up the results.

2 The Proposal

2.1 The Parameters of the ExtAF Theory

In order to adequately capture cross-linguistic variation in accentual patterns, the parameter system of the AF theory is revised as shown in (1). In particular, dependency relations among certain parameters are identified.

(1) *The accentual parameters of the ExtAF theory*

a. The Domain Size Parameter: the accent domain is {bounded/unbounded}.
b. The Domain Edge Parameter: a bounded accent domain is formed at the {left/right} word edge.
c. The Extrametricality Parameter: a peripheral element in the word is allowed to be accented. (yes/no)
d. The Extrametricality Edge: the extrametrical unit is on the {left/right} word edge.
e. The Extrametrical Unit: the extrametrical unit is a {syllable/segment}.
f. Project Weight: heaviest units are projected onto the grid. (yes/no)
g. Project Position: project {leftmost/rightmost} position in the word.
h. Select: choose the {leftmost/rightmost} gridmark in the accent domain.

All parameters in (1) are binary. Domain Size (1a) determines whether the system is bounded or unbounded, while Domain Edge (1b) places the bounded accent domain at or near the edge where word accent is located. In systems where the peripheral syllable is never accented, the posi-

---

*This paper owes much to seminal work on the PAF theory by Prof. Harry van der Hulst (University of Connecticut), and to his indispensable mentorship. I also wish to thank the audience at PLC39, in particular Prof. Mark Liberman (University of Pennsylvania), for their valuable feedback. Special thanks to Noémie Abitbol for her long-lasting support.
tive setting of the EM parameter (1c) makes it invisible to accent assignment, or extrametrical, shifting the bounded domain one syllable inside the word.

The last three parameters in (1), namely Project Weight (1f), Project Position (1g) and Select (1h), play a major role in accent assignment. The Project Weight parameter projects the heaviest syllables in the word, based on the Weight Grid (to be discussed), onto line 1 of the Accent Grid, while the other syllables are not projected, thus becoming invisible to the Select parameter. If there are no heavy syllables in a word, the Project Position parameter applies instead, projecting a peripheral position in the accent domain by placing a gridmark on line 1 over the leftmost or rightmost syllable in this domain. That is, the Project Weight parameter and the Project Position parameter are complementary: the former applies to words that contain heavies, while the latter applies to “all-light” words.

The Select parameter resolves the situation where a word contains more than one heavy syllable by choosing one of the gridmarks projected onto line 1 by the Project Weight/Project Position parameters and places a gridmark on line 2 on top of the chosen gridmark. The line 2 gridmark is then read off the Accent Grid as word accent. Note that the Project Position parameter, which applies to “all-light” words, places only one gridmark on line 1. In this case, the Select parameter does not choose between different gridmarks, but may be freely set to either value, which guarantees that the value chosen for all-light words will be the same as for words with heavies.

Another new feature of the ExtAF theory is that it reduces the dimensionality of the parameter system (1) by revealing dependencies among certain parameters. For example, since the gridmarks placed on line 1 by Project Weight and Project Position feed into Select, the latter is dependent on both Project parameters. At the same time, the latter are mutually independent because they operate on complementary sets of words (see above).

The dependency graph in Figure 1 displays the dependency relations on the parameter set (1).

```
<table>
<thead>
<tr>
<th>Domain Size</th>
<th>Extrametricality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Project Weight</td>
<td>Domain Edge</td>
</tr>
<tr>
<td>Project Position</td>
<td>EM Unit</td>
</tr>
<tr>
<td>Select</td>
<td>EM Edge</td>
</tr>
</tbody>
</table>
```

Figure 1: Dependencies among parameters in the ExtAF theory.

### 2.2 Representing Weight in the ExtAF Theory

#### 2.2.1 Diacritic Weight

Morphemes, like syllables, are capable of attracting or repelling accent: certain morphemes can be accented, while others cannot. Accordingly, van der Hulst (1999:19) calls this ability of morphemes “diacritic weight”. That is, accent attraction can be captured in terms of weight, rather than in terms of lexical accents. This implies a change in perspective: accent-attracting morphemes are diacritically heavy (rather than lexically accented), accent-repelling morphemes are diacritically light (rather than lexically unaccented).

Now, one might ask whether syllable weight and diacritic weight are different instances of the same notion “weight”. Indeed, they differ in that syllable weight is phonologically motivated (by syllable and/or segmental structure), while diacritic weight is not predictable and, as such, must be assigned in the lexicon. Nevertheless, diacritic weight and syllable weight form a class because they pattern together: (i) accent assignment may make reference to both types of weight within a single language, and (ii) in some languages, diacritic and syllable weights are ordered in a single weight scale.
Below, building on the notion of “diacritic weight”, I introduce several novel theoretical devices that augment the AF theory so as to extend its empirical coverage (which is why I call the new theory “Extended”).

2.2.2 Diacritic and Hybrid Weight Scales

It is well-known that, in some WS languages, accent is assigned with reference to a phonological weight scale. Examples of some such scales are given in Table 1.

<table>
<thead>
<tr>
<th>Language</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td>Klamath (isolate; Oregon, USA)</td>
<td>CVV(C) &gt; CVC &gt; CV</td>
</tr>
<tr>
<td>Moro (Niger-Kongo; Sudan)</td>
<td>CVC &gt; full V &gt; reduced V</td>
</tr>
<tr>
<td>Kobon (Trans-New Guinea; PNG)</td>
<td>low V &gt; mid V &gt; high V &gt; reduced V</td>
</tr>
<tr>
<td>Asheninca (Maipurean; Peru)</td>
<td>CVV &gt; Ca(C),Ce(C),Co(C), Ci &gt; Ci</td>
</tr>
</tbody>
</table>

Table 1: Examples of phonological weight scales (from Gordon 2006:27–28).

Similarly, in some accent systems with diacritic weight, diacritic weight distinctions are scalar rather than binary. In these systems, accent is assigned with reference to a diacritic weight scale, i.e. a language-specific scale in which (sets of) morphemes are ordered according to their relative diacritic weight. An example of a diacritic weight scale is given in (2).

(2) diacritically superheavy > diacritically heavy > diacritically light

Although diacritic weight scales resemble phonological weight scales in that both are ordinal, the two differ in that the former order morphemes, while the latter order syllables. This predicts that another type of scale is also possible: a language-specific scale that orders both syllables and morphemes. This prediction turns out to be correct: such scales, which I call “hybrid weight scales”, are attested cross-linguistically, e.g., in Eastern Literary Mari, Uzbek and certain Athabascan languages (Vaxman 2014, in prep).

2.2.3 The Weight Grid

A Weight Grid represents relative weight (according to the weight scale) in terms of columns of gridmarks: the taller the column, the heavier the relevant unit (syllable or morpheme); a light unit gets one gridmark.1 Diacritic, phonological and hybrid weight scales can all be translated into such Weight Grids. For example, the weight scale superheavy > heavy > light corresponds to the Weight Grid in (3).

(3) The Weight Grid

```
 sup  h  1
       *   *
       *   *
       *   *
```  

2.2.4 Projection

Project Weight and Project Position place gridmarks on the Accent Grid with reference to the Weight Grid. (I assume that these grids are located on separate planes.) Project Weight projects the heaviest units (syllables or morphemes), i.e., those which have the tallest column of gridmarks on the Weight Grid, by placing a (single) gridmark on line 1 of the Accent Grid for each heaviest

---

1The Weight Grid builds on proposals to grid syllable weight (Prince 1983, van der Hulst 1984:67–68) and sonority relations (Parker 1989:9–12). The crucial difference with the current proposal is that the Weight Grid proposed here can represent differences in diacritic weight that characterize morphemes.
unit. In “all-light” words, Project Weight is not applicable; instead, Project Position places a grid-mark over the left/right edge on line 1 of the Accent Grid.

Section 3 illustrates accent assignment in the ExtAF theory on the example of Central and Southern Selkup.

3 Accent Assignment in Selkup

3.1 Introduction

In Section 3, I will discuss accent assignment in Selkup, a Samoyedic language of the Uralic family (spoken in the Tomsk Region, the Krasnoyarsk District and the Yamal-Nenets Autonomous Area in Siberia), focusing exclusively on its Central and Southern dialects (henceforth, C. and S. Selkup). The goal of this section is to account for the accentual generalizations regarding these Selkup dialects in terms of the ExtAF theory. I must add that this is the first generative account of their accent system.

Dialectically, Selkup is not monolithic; in particular, N. Selkup dialects differ from C. and S. Selkup with respect to weight sensitivity. Thus, while the Taz dialect (N. Selkup) is a phonological WS system (McNaughton 1976:135, Idsardi 1992, Halle and Idsardi 1995), C. and S. Selkup dialects are not sensitive to syllable weight and assign accent with reference to diacritic weight alone (see Sections 3.3–3.5).

Section 3 is organized as follows. Section 3.2 provides general information about Selkup dialects and my data sources. Section 3.3 offers evidence that C. and S. Selkup has an unbounded lexical accent system. Sections 3.4–3.5 describe accent patterns in Napas and Parabel Selkup. Then, Section 3.6 identifies an important theoretical problem posed by some of these accent patterns. Finally, in Section 3.7, I present a solution to this problem in the framework of the ExtAF theory.

3.2 The Background

In this article, I mainly discuss Central Selkup, namely the Tym dialect (as spoken in the village of Napas) and the Narym dialect (as spoken in the villages of Parabel and Laskino). Although the Chaya dialect is the only dialect of S. Selkup mentioned here, the results of this article can be straightforwardly extended to Southern Selkup dialects as well.

All Selkup data used here are drawn from Normanskaya et al. 2011 and Normanskaya 2012a,b, which are based on fieldwork materials located in the extensive “Dulzon archive” (held at the library of the National Pedagogical University of Tomsk) and on the 2009 fieldwork with speakers of Parabel Selkup carried out by N. L. Fedotova and S. E. Šešenin.

3.3 Central and Southern Selkup as a Lexical Accent System

C. and S. Selkup dialects differ with respect to their vowel inventories. However, since accent is assigned in these dialects without reference to vowel properties, the vowel system is not relevant here.

The accent system of C. and S. Selkup is unbounded: while, in (4a–b), accent is inside the three-syllable window at the left word edge, it can also reach out of this window, witness (4c).

(4) a. ˈtʃɔndipuɡu cover
b. tʃonˈdiʃpuɡu girdle
c. kyʒɔmbuˈgu urinate

C. and S. Selkup has many minimal pairs, e.g., (4a–b). This implies that accent in C. and S. Selkup is contrastive and, therefore, not phonologically predictable. Indeed, these dialects of Selkup have been analyzed as a lexical accent system by Normanskaya et al. (2011) and Normanskaya (2012a,b), with lexical (un)accentedness of individual morphemes being determined based on their accentual patterning in complex words.
In Sections 3.4–3.5 below, the accent rule for two Selkup varieties, Napas and Parabel, is established.  

3.4 Accent Patterns in Napas Selkup

Let us begin with Napas Selkup. First, when a lexically accented suffix, e.g., /-e/ in (5), is attached to an unaccented root, accent falls on the suffix, which is the unique accented morpheme in the word.

(5) kiˈɡʲ-ɐ river

Further, in words that contain more than one accented morpheme, accent falls on the leftmost such morpheme (6).

(6) accented root-accented suffix
   a. ˈkomd-ɐ money
   b. ˈʃjɪb-ɐ fly

In words that contain an unaccented root followed by more than one accented suffix, accent falls on the leftmost accented suffix.

(7) unaccented root-accented /-eʃ/-unaccented /-pu/-accented /-gu/
   a. tʃˈond-ɐʃ-pu-gu girdle
   b. xel-ɐʃ-pu-gu sharpen

Note that lexically accented roots (here, /igʲ/) receive the accent in words with a lexically accented suffix (Selkup lacks prefixation):

(8) ˈigʲ-ɐʃ-pu-gu detach

I conclude that in words that contain lexically accented morphemes, accent falls on the leftmost lexically accented morpheme.

Finally, in lexically unaccented words, accent falls on (the initial syllable of) the leftmost morpheme; that is, default accent is initial in Napas. For example, when the unaccented suffix /-a/ is added to an unaccented root, as in (9), accent falls on the initial syllable.

(9) unaccented root-unaccented suffix
   a. ˈam-ɐ mother
   b. ˈloŋ-ɐ fox

I conclude that, in Napas Selkup, accent falls on the leftmost lexically accented morpheme; otherwise, accent is initial.

3.5 Accent Patterns in Parabel Selkup

Let us now turn to the accent patterns in Parabel, a variety of Narym Selkup.

In Parabel, when an accented suffix, e.g., /-a/ in (10a) or /-e/ in (10b), is attached to an unaccented root, accent falls on the suffix.

(10) a. kal-ˈa cup
     b. tʃˈond-ɐ waist

---

2The following description owes much to the analyses in Normanskaya et al. 2011 and Normanskaya 2012a,b.
In words that contain two lexically accented morphemes, accent falls on the leftmost one, as in (11):

(11) a. ˈarm-a  coolness
    b. ˈky3-e  urine

Then, the accent rule for Napas and Parabel can be approximated as in (12).

(12) The accent rule (preliminary)
Accent falls on the leftmost lexically accented morpheme in the word (if any); otherwise, accent is initial.³

In other words, Napas and Parabel Selkup have a First/First unbounded system.

3.6 The Problem: Accent-categorizing Suffixes

In Parabel and Laskino Selkup, certain suffixes, called “accent-categorizing”, are special in that they receive word accent, regardless of the presence or absence of a lexical accent on the other morphemes in the word (Normanskaya et al. 2011, Normanskaya 2012a,b). In particular, word accent always falls on the semelfactive suffix -ol/-al.⁴

(13) a. accented root-categorizing suffix-accented suffix
    tap-ˈol-gu  kick-SEMEL-INF
    kob-ˈal-gu  scour-SEMEL-INF
    b. unaccented root-categorizing suffix-unaccented suffix-accented suffix
    kad-ˈol-bi-gu  scratch
    yt-ˈal-ʒu-gu  make drunk

Crucially, when an accented root is followed (not necessarily immediately) by an accent-categorizing suffix, as in (13a), accent does not fall on the leftmost accented morpheme, thus violating the accent rule (12).

Note that the AF theory fails to capture the pattern in (13a): indeed, setting Select to “Left” would capture the general case described in (12), but does not derive the special pattern in (13a). In the next section, I will show how this problem is addressed in the framework of the ExtAF theory.

3.7 The Account

3.7.1 The Diacritic Weight Scale for Central and Southern Selkup

Viewing accent attraction as an effect of diacritic weight, whereby lexically accented morphemes are diacritically heavy and lexically unaccented morphemes are diacritically light (see Section 2.2.1), the pattern in (13a) can be analyzed as containing a diacritically superheavy suffix because this suffix always attracts accent, notably in words that also contain heavy morphemes. That is, we are dealing with a scalar weight distinction: diacritically superheavy vs. diacritically heavy vs. diacritically light.

Moreover, these three degrees of diacritic weight are ordered. Indeed, heavy morphemes are heavier than the light ones: for example, in (10), a heavy suffix receives the accent when combined with a light root. As noted above, superheavy morphemes are heavier than heavy mor-

³Note that, for Parabel, the default is not certain due to lack of relevant data.
⁴In the above-cited publications, Julia Normanskaya reports that there are several accent-categorizing morphemes in Selkup, but cites only the suffix -ol/-al discussed here.
phemes (13a). Finally, as can be seen in (13b), superheavy morphemes are heavier than light morphemes.

I conclude that the accent is assigned in C. and S. Selkup with reference to the diacritic weight scale in (14):

(14) diacritically superheavy > diacritically heavy > diacritically light

Given the scale (14), the accent rule can now be restated in its final version as (15).

(15) The accent rule (final)

In words that contain heavy and/or superheavy morpheme(s), accent falls on the leftmost diacritically heaviest morpheme; if all morphemes are diacritically light, accent is initial.

In this way, the rule in (15), stated with reference to (14), uniformly describes regular accentuation together with the systematically exceptional patterns involving accent-categorizing suffixes.

3.7.2 The Accentual Grammar

I will now present the ExtAF grammar needed to derive the accent patterns of Selkup described above.

Recall that, according to the ExtAF theory, an accentual grammar for a given language consists of a combination of parameter settings and of a language-specific Weight Grid. I submit that the accentual grammar of C. and S. Selkup consists of the Weight Grid (16) (a representation of the weight scale in (14) above) and of the combination of parameter settings in (17).5

(16) The Weight Grid

\[ \begin{array}{ccc} 
\text{sup} & \text{h} & \text{i} \\
* & * & * \\
* & * \\
* 
\end{array} \]

(17) Domain Size (Unbounded)
EM (No)
Project Weight (Yes)
Project Position (Left)
Select (Left)

3.7.3 Derivations

In this section, I will show how certain representative accent patterns described above are derived in the ExtAF theory.

(18) Words containing heavy morphemes
/tvel-gu/: heavy root /tvel/, heavy suffix /-gu/
* Select (Left)
* * Project Weight (Yes)

\[ \begin{array}{c} 
* * \\
* * \\
tvel-gu \\
[tvelgu] steal-INF 
\end{array} \]

5The subscript “\(\hat{d}\)” stands for “diacritically”.

Recall that, in the ExtAF theory, Project Weight only projects the heaviest morphemes from the Weight Grid onto the Accent Grid. Therefore, in (19), the superheavy morpheme is projected, while the heavy morphemes are not, which explains the accent on the superheavy.

(19) **Words containing a superheavy morpheme**

/tap-ol-gu/: heavy root /tap/, superheavy suffix /-ol/, heavy suffix /-gu/

* Select (Left)
* Project Weight (Yes)

* * * Weight Grid

* * *

* tap-ol-gu

[taˈpolgu] kick-SEMEL-INF

(20) **All-light words**

/lar-em-bu-gu/: a light root followed by three light suffixes

* Select (Left)
* Project Position (Left)

* * * * Weight Grid

lar-em-bu-gu

[ˈlarembuɡu] fear

4 **Diacritic Weight vs. Lexical Accent**

The ExtAF theory accounts for patterns with “accent-categorizing” morphemes using weight scales, which are possible because weight is ordinal. By contrast, given that lexical accent is categorical, lexical accent theories can have at most a binary distinction (accented vs. unaccented). Therefore, they cannot make reference to a weight scale (whether diacritic or hybrid). Instead, they typically have recourse to accent deletion (see Poser 1984), whereby word accent on the accent-categorizing morpheme results from deleting the lexical accents on all other accented morphemes.

However, the Accent Deletion approach has some significant drawbacks. Firstly, in order to delete lexical accents on both sides of an accent-categorizing morpheme, (at least) two accent deletion rules are required, whereas in the ExtAF theory, the superheavy morpheme “wins” simultaneously over all the heavy morphemes in the word because it is the heaviest morpheme, according to the diacritic weight scale.

Secondly, when a lexically accented morpheme is not adjacent to an accent-categorizing one, accent deletion rules must apply non-locally. Therefore, the formalism of the Accent Deletion approach needs to be augmented with non-local rules. By contrast, the ExtAF theory deals with this issue using the same device, viz. a weight scale.

I conclude that, with respect to “accent categorization”, the ExtAF theory is more straightforward and economical than the Accent Deletion approach of lexical accent theories.

---

6Due to the absence of words with more than one superheavy morpheme (at least in my corpus), I cannot verify that Select is set to “Left” for such words (if any).

7In the Chaya dialect (S. Selkup), from which this example comes, the infinitival marker /-gu/ is light.
5 Conclusion

In this article, I have presented the Extended Accent First (ExtAF) theory, a new theory of word accent resulting from an extension of the PAF theory.

The PAF theory is known to correctly account for accent location in a large variety of languages. However, it encounters difficulties with lexical accent systems and systems sensitive to both phonological weight and lexical accent. The ExtAF theory makes such an account possible.

In the ExtAF theory, lexical accent is reanalyzed as “diacritic weight” (Section 2.2.1), which leads to a generalized notion of “weight scale” as an ordering of morphemes and/or syllables in terms of their relative weight (Section 2.2.2).

In a case study (Section 3), I have shown how the accent system of Central and Southern Selkup (a First/First unbounded diacritic weight system) can be formalized within the ExtAF theory in terms of a diacritic weight scale (represented as a “Weight Grid”; see Section 2.2.3) and of a particular set of parameter settings (see Section 2.1).

Also, a brief comparison of the present ExtAF account of Selkup with the Accent Deletion approach suggests that the former is more straightforward and economical than the latter (see Section 4).

In conclusion, it is worth noting that abstracting away from the phonological motivation of weight (based on weight criteria), which is limited to syllable weight, allowed us to extend the notion of “weight” to morphemes by encompassing lexical accent (reanalyzed as diacritic weight). Interestingly, the existence of two types of weight, phonological and diacritic, predicts that, in some languages, accent assignment makes reference to both (ordered into a language-specific “hybrid weight scale”). We find that this prediction is effectively borne out, e.g., in Eastern Literary Mari, Uzbek and some Athabascan languages (Vaxman 2014, in prep).

References


Vaxman, Alexandre. 2014. The smell of morphemes in the PAF theory: the case of Eastern Mari. Poster pre-
sented at the 4th Workshop on Accent and Stress, Leiden University.

Department of Linguistics
University of Connecticut
Oak Hall (East SSHB)
365 Fairfield Way
Storrs, CT 06269
alexandre.vaxman@uconn.edu