TABLE OF CONTENTS

Preface ........................................... v

Part I. General and Experimental.

Alan Bell, ACCENT PLACEMENT AND PERCEPTION OF PROMINENCE
RHYTHMIC STRUCTURES ........................... 1

Victoria A. Fromkin, PUTTING THE EMPHASIS ON THE WRONG
SYLLABLE ........................................ 15

Eva Gården, THE IMPORTANCE OF TURNING POINTS FOR THE PITCH
PATTERNS OF SWEDISH ACCENTS .............. 27

Larry M. Hyman, ON THE NATURE OF LINGUISTIC STRESS ........... 37

Wayne A. Lea, ACOUSTIC CORRELATES OF STRESS AND JUNCTURE .. 83

Larry Nessly, ON THE VALUE OF PHONOLOGICAL EXPERIMENTS IN
THE STUDY OF ENGLISH STRESS ................. 121

John J. Ohala, THE PHYSIOLOGY OF STRESS .................. 145

Part II. Language Studies.

Wallace L. Chafe, ACCENT AND RELATED PHENOMENA IN THE FIVE
NATIONS IROQUOIS LANGUAGES .............. 169

Alice Wyland Grundt, SYNTACTIC ACCENT IN NORWEGIAN
MORPHOLOGY ..................................... 183

Mary R. Haas, TONAL ACCENT IN CREEK ............. 195

Paul Kiparsky and Morris Halle, TOWARDS A RECONSTRUCTION OF
THE INDO-EUROPEAN ACCENT .................. 209

Margaret Langdon, STRESS, LENGTH, AND PITCH IN YUMAN
LANGUAGES ...................................... 239
<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>James D. McCawley</td>
<td>ACCENT IN JAPANESE</td>
<td>261</td>
</tr>
<tr>
<td>Pamela Munro</td>
<td>TOWARDS A RECONSTRUCTION OF UTO-AZTECAN STRESS</td>
<td>303</td>
</tr>
<tr>
<td>Manjari Ohala</td>
<td>STRESS IN HINDI</td>
<td>327</td>
</tr>
</tbody>
</table>
The papers included in the volume were presented at a STRESSFEST, or Symposium on Stress and Accent, held at the University of Southern California on February 28-29, 1976. Together, they represent a broad spectrum of the approaches taken to the study of accentual phenomena and cover many of the important language areas of the world. The 15 studies are grouped into two sections: general and experimental (section 1) and language studies (section 2).

In the first section, the authors are concerned with the production, perception, and/or psychological implications of linguistic analyses of stress/accent. Bell's paper reports on experiments attempting to seek a correlation between one's accentual system (in particular, stress-placement) and the way sequences of like pitch-signals are perceived. Fromkin's paper discusses the nature of speech errors made in the realization of stress and relates these to the phonological analysis of English stress. Garding's paper demonstrates the importance of the timing of pitch changes in the realization of the two accents in Swedish. Hyman's paper investigates universals of stress-accent on the basis of a survey of more than 400 languages. Lea's paper treats the acoustics and perception of stress in English, especially with respect to the detection of syntactic boundaries. Nessly's paper tells of results of experiments where speakers assigned stress to nonsense words, and finally, J. Ohala's paper provides a thorough study of the physiological mechanisms underlying the production of stress.

In the second section, the authors are concerned with descriptive and/or theoretical aspects of accent in individual languages or language families. Chafe's paper surveys stress in the Five Nations Iroquois languages, while Grundt looks at the morphological side of accent in Norwegian. Particularly intriguing is the tonal accent system of Creek discussed in the paper by Haas, as well as the history of the Indo-European accent discussed in the paper by Kiparsky and Halle. In the paper by Langdon, stress (and some tone) is surveyed in the Yuman languages, and in McCawley's paper general statements are made concerning the pitch-accent systems of the various dialects of Japanese. In Munro's paper stress is reconstructed for Proto-Astecan with supporting evidence drawn from throughout that family. Finally, S. Ohala approaches the difficult and controversial subject of whether Hindi has, or does not have, stress.

As can be seen from the table of contents, this volume is both heavily experimental and heavily Amerindian. While most of the contributions deal exclusively with stress-accent, tonal or pitch-accent is covered in Scandinavian, Japanese, and some Amerindian.
In addition to the above, four other papers were presented at the symposium: "Alternating stresses: variations on a theme" by Jonathan Kaye, "The realization of accent in some nonstandard dialects of Serbo-Croatian" by Ilie Lehiste, "The nature and derivation of stress patterns in English" by Mark Liberman, and "Eskimo stress and consonant gradation" by Roy Wright.

The Symposium on Stress and Accent was partially supported by Humanities Division Development Funds at the University of Southern California, made available by Dean David H. Malone. The publication of these proceedings has been made possible by additional support from a National Science Foundation Grant No. SOC 75-16187, under which sponsorship this symposium was conceived. Finally, my personal thanks to all of those persons who made this symposium and this volume possible, especially (for hospitality shown to out-of-town guests) Vicki Fromkin, Peter Ladefoged, Steve Anderson, Jean-Marie Hombert, Pam Munro, Russ Schuh, Matt Shibatani, Tina Bennett, and Harriet Jies (the last of whom typed this volume).
ACCENT PLACEMENT AND PERCEPTION OF PROMINENCE
IN RHYTHMIC STRUCTURES *

Alan Bell
University of Colorado

The concern of this paper is the relation of the perception of auditory temporal patterns (henceforth rhythmic perception) to accentual systems of language. It deals mainly with one relatively straightforward aspect of the topic, the position of prominent (accented) elements and the sensory dimension on which they are marked. It is intended to mesh with the general enterprise of discovering what accentual systems are like, and how they came to be that way. For the proper direction of that enterprise it is well to look forward to the forces that may provide a deeper explanation of the regularities to be uncovered.

The existence of a relation between rhythmic perception and linguistic accent is most plausible under the assumption that the raison d'être of accent is at least partially functional. One conjecture, implicit in Trubetzkoy's characterization of stress as a culminative element of language, is that it provides an organizational framework for speech, which facilitates its use in some way.

This framework is a rhythmic structure, in the sense of a quasi-regular temporal succession of patterns. In the simplest cases, the pattern is based upon the contrast between a prominent (accented) syllable and nonprominent ones, and is composed of a succession of domains (usually words or phrases) in which the prominent element occurs once in a fixed position. The temporal pattern of one such accental system might be *---*---*---*---*---* where the number of nonprominent elements varies according to the length of the word.

Rhythmic perception. Psychological research on rhythmic perception parallels the linguistic work on stress and accent summarized by Hyman in this volume. One line of research emphasizes the role that the physical characteristics of the sounds of an auditory pattern play in its perception. This corresponds to research on the physical correlates of accent. Another focuses upon the organizational principles that people use in the perception of rhythmic patterns. This corresponds to the functional and structural studies of stress and accent.

A fundamental result concerning the physical qualities of succession of sounds is that their rate of succession must fall within a limited range for them to be
perceived in terms of a rhythmic pattern. Although there is substantial variation from
one person to another, if the rate is much greater than 5 repetitions per second, the
sounds tend to be heard as a continuous stream. As the rate falls below one repetition
every two seconds, the sounds tend to be perceived individually (Woodrow 1951, Fraisse
1956). This range of course includes the ordinary repetition rates of rhythmic elements
in speech, with syllable rate of conversational speech, about 5 syllables per second,
at the high end.

Within this range, even a perfectly regular succession of identical sounds will
often be perceived as composed of repeating groups of sounds. The length of the
group is influenced by the rate of repetition, longer groups being heard at higher rates.
This is the phenomenon called subjective rhythm (Woodrow 1951, Fraisse 1956).

The rhythmic structure that is perceived, i.e., the number of sounds in the
repeated groups and its internal structure (where the "accents" fall), is otherwise
influenced by the physical parameters of the sound sequence.

A rhythmic group of identical sounds may become salient by virtue of a separa-
tion between the end of one group and the beginning of the next that is greater than
the separation between the members of the group.

Alternatively, rhythmic structures may be marked by repeating combinations
of different sounds. The simplest cases consist of two different sounds, e.g., one
loud and one soft. Thus a regular alternation of soft (S) and loud (L) sounds will be
perceived as a succession of groups of two; a sequence such as ... LSLSSLSLSLSLSS ...
will be perceived as a succession of groups of five; etc.

It may have already struck the reader that sequences such as these are ambiguous.
That is, the first can be interpreted equally well as a succession of LS groups as of
SL groups. And there are five possible rhythmic groups for the second: LSLS,
SLSSI, LSSLS, SSLSL, and SLSLS. Psychologists have exploited a number of responses--
impressionistic reports by the listener of what he hears, forced choice among pattern
representations, and reproduction of patterns by tapping or otherwise--to infer which
of the possible groups a listener hears when presented with sequences of this kind.
One might expect that listeners would be more or less equally divided among the
possible rhythmic groups. This is not the case, in fact. There is almost always a
decided preference for certain groupings.

Influence of sensory dimension of prominence. The sensory dimension along which
the sounds differ influences which pattern will be perceived. Woodrow (1909) and
Fraisse (1956) report that there is a strong preference for groups beginning with
the louder sound. For example, a sequence ... SLSSLSLSSLSLSS ... will tend to be heard
as a succession of groups beginning with the loud sound—-... -LSS-LSS-LSS--, even
though all the sounds occur at equal intervals. The effect increases with increased difference in loudness between the loud and soft sounds according to an S-shaped curve. At repetition rates less than 3.5 per second the effect diminishes.

The tendency to perceive rhythmic structures with initial accent is found also for subjective rhythm.

On the other hand, if the sounds differ only in length, a different pattern of responses occurs. If the longer element is less than 1.5 to 2.0 times the shorter one, a tendency toward grouping with an initially prominent element is again found. For ratios above that, a tendency to perceive rhythmic groups with a long final element becomes evident.

The perception of rhythmic sequences whose sounds differ in frequency is more complex. Woodrow (1911) found that a higher pitch could either begin or end groups of two or three. In three similar experiments concerning the perception of 8- and 9-element groups, Royer and Garner (1966, 1970) and Handel (1974) different results were obtained. No tendency for either higher or lower pitched sounds to begin the perceived rhythmic groups was found in the first of these experiments, but two-thirds began with the higher pitched sound in the second, and about the same proportion began with the lower pitched sound in the third. Handel points out that sounds of different pitch are not quantitatively different in the manner of sounds of different loudness or duration, but rather have a more nearly qualitative difference. (This calls to mind the all-or-none effect of pitch difference in the experiments of Fry 1968 and others on the correlates of stress.) He conjectures that pitch itself may not determine element preference as much as secondary differences in timbre, etc.

**Structure influences.** The structure of the rhythmic group itself plays as large a role in perception as the sensory dimensions of prominence. Out of the four rhythmic groups possible for a sequence ...SLSSISSSLSSS..., LSSS and SSSL are preferred to the two groups with the loud sound in the interior of the group. If there are two loud sounds in the sequence, LLSS and SSLL are preferred to groups LSSL and SLLS (Fraisse 1966).

From the research by Royer and Garner (1966, 1970) and Handel (1974) who considered the perception of sequences made up of recurring groups of eight sounds, we know that these preferences extend to more complex sound combinations. They offer the explanation that the elements of temporal sequences are perceived either as figure elements or as ground elements. The sequences may be organized according to two principles. The gap principle states that a long string of ground elements ends a pattern; for simple patterns with one accent this corresponds to initial placement of accent, e.g. LSSS. The run principle states that a long string of figure elements begins a pattern. For the same example, this corresponds to final
accent placement, SSL, where now the soft sounds are figure elements. Note that for the sequences of Fraisse mentioned above, neither SLSS, SSLS nor LSSL, SLLS fit either principle.

Some kinds of sequences are harder to perceive and learn (i.e., reproduce or remember) than others. Sequences of the type ...SLSSSSSLSS... or ...SLLSSSLSSSS... containing unbroken strings of the two kinds of elements are easier than ones like ...SLLSLSSLSSLSSLSSL... Internal repetition of subgroups like SL appears to facilitate perception and production.

Summary. Perception of auditory temporal patterns possesses general characteristics which are consonant with the conjecture that the same mechanisms play a role in linguistic accentual systems.

1. Rhythmic perception occurs at rates comparable to the relevant repetition rates in language.
2. Where several organizations of a sequence are possible, perceived rhythmic groups tend to be based upon either initial or final position.
3. Sequences with essentially a culminating structure (one or few prominent elements or shifts from one element to another) are easier to perceive and learn.

The following characteristics of temporal perception pertain to the placement of accent and its sensory dimensions.

1. **Rhythmic groups** with an initial prominent element are preferred even if the physical stimulus is a regular succession of identical elements.
2. This tendency is increased if one of a group of elements is louder.
3. Next in preference are rhythmic groups with a final prominence; groups with medial prominence are relatively disfavored.
4. If the prominent element is marked by greater length, there appears a tendency for the longer element to end the perceived rhythmic group, as the ratio of short to long surpasses 1.5.
5. An element of a rhythmic group can be accented by receiving either higher or lower pitch than the other elements, in contrast to loudness. To the extent that Handel's results can be extended to rhythmic groups with a single prominence, there should be a tendency for such groups to be perceived with the different element (whether higher pitched or lower pitched) in initial position.²

Do accentual systems affect rhythmic perception? If accentual systems and rhythmic perception are significantly related, the relation could be from either one to the other, or both. The prior question is the effect of the accentual system of a person's language upon his perception of rhythmic structures in general. For if there is an effect,
explanations of accentual phenomena in terms of general rhythmic performance are jeopardized by circularity.

Jakobson, Fant, and Halle (1952) did indeed claim that there exists such an effect:

Interference by the language pattern affects even our responses to non-speech sounds. Knocks produced at even intervals, with every third louder, are perceived as groups of three separated by a pause. The pause is usually claimed by a Czech to fall before the louder knock, by a Frenchman to fall after the louder; while a Pole hears the pause one knock after the louder. The different perceptions correspond exactly to the position of the word stress in the languages involved: in Czech the stress is on the initial syllable, in French, on the final, and in Polish, on the penult. When the knocks are produced with equal loudness but with a longer interval after every third, the Czech attributes greater loudness to the first knock, the Pole, to the second, and the Frenchman, to the third.

**Experiment on perception of accent placement.**

An experiment was designed to investigate perception of simple auditory patterns of tone sequences. The sequences consisted of tones at equal intervals, in which every third tone differed in either frequency, intensity, or duration. Subjects were drawn from languages of different accentual patterns: non-fixed (English), and fixed (Bengali, with initial phrasal accent; Polish, with penultimate word accent; French and Persian, with final phrasal and word accent respectively).  

1. **Tone Sequences.** Four different sequences of tones were used. In all cases, the interval from tone onset to tone onset was 261 msec., or a repetition rate of 3.8 per second.

2. **Control sequence.** One sequence consisted of almost identical pulses of a 500 Hz tone, with an average length of 109 msec., separated by intervals of 152 msec.

3. **Frequency sequence.** In the frequency sequence every third tone was 540 Hz, the sequence otherwise being the same as the control sequence.

4. **Intensity sequence.** Every third tone was louder than the others by an amount between 2 and 3 dB, the sequence otherwise being the same as the control sequence.

5. **Length sequence.** Every third tone was 1.4 times the length of the others, with period between onset of the tones held constant.
To ensure that the tone sequences were ambiguous with respect to location of the prominent tone, it was important to eliminate any effect of a pattern beginning the sequence. Each sequence began with no prominent tone, and the prominence was then gradually increased on every third tone until the desired level was reached, after approximately 10 seconds. The sequences could be represented thus:

```

```

Each sequence of tones lasted 55 seconds.

**Procedures.** Each subject was presented a recorded set of 5 tone sequences at a peak sound level of 80 dB to the left ear and 74 dB to the right ear (an unintentional difference). The set began with the control sequence, followed by the three patterned sequences. The order of these sequences was counterbalanced across subjects. The control sequence ended the set. The sequences were 30 seconds apart.

The subjects were told that they would hear patterns of three tones repeated over and over. They were instructed to wait until they heard the pattern and to indicate which it was by writing either /- -, - / -, or --/ depending on the location of the prominent tone in the pattern.

**Subjects.** There were a total of 48 subjects. Eighteen were native speakers of English. Twelve were native speakers of Persian, and there were six each of Bengali, French and Polish. No subject was a native bilingual. Subjects all reported that they used their native language daily for several hours or more (one French speaker excepted). Instructions for the experiment were written in the subject's native language. Every effort was made to perform the experiment in a context favoring the use of their language (as in their home, in the presence of another person speaking their language, etc.). All subjects reported no history of hearing difficulty.

Some subjects were replaced for the following reasons: 1) post-experiment interview indicated that their choice of pattern was based on how the sequence ended (3 cases). 2) Post-experiment interview indicated that the pattern chosen was based strongly on the switching clicks in the background (3 cases). Two of these subjects were musicians. Some did not notice the clicks at all; others did, but reported not attending to them. 3) Either because of failure to understand the instructions or for other reasons, the patterns of two or more of the marked sequences were not clearly indicated.

**Results.** The responses of the listeners were scored according to the position of prominence in the pattern they reported. In Table I, I stands for initial prominence (/ - - ); M, for medial prominence (- / - ); and F, for final prominence ( - - / ). No
entry was made when a listener reported no or uncertain prominence, or when his response was not unequivocally interpretable. Listeners were told that they should write down another pattern if they changed their mind, and some did so. Only the first pattern recorded was scored.

Order effects. Inspection of the responses to the frequency, intensity, and length sequences by order of presentation did not show any marked effect or interaction. The differences between the initial control and final control are significant ($\chi^2 = 6.6$, d.f. = 2, p $\leq .05$).

Language. The tendencies that appear for each language do not follow the pattern of the language's accent placement. For all three sensory dimensions, French shows a tendency toward perception of groups with initial prominence, rather than final, as would accord with its accentual system. It is Persian, not Polish, that shows a tendency for perception of groups with medial prominence. Were the differences among languages significant, we would be hard put to explain them. In fact, the differences among languages for the three accented sequences are not statistically significant.

Sensory dimension of prominence. A preference for groups with initial accent is apparent both in the initial control sequence, with no prominent sounds and the frequency and intensity sequences. For the length sequence, on the other hand, there was a preference for groups with medial accent. The difference between the combined responses to the frequency and intensity sequences and the responses to the length sequences is significant ($\chi^2 = 10.6$, d.f. = 2, p $\leq .005$).

Conclusions. The results do not support the hypothesis that the accentual pattern of a language influences its speakers' rhythmic perception of non-speech sounds. They do not rule it out, of course. The influence may be small enough that the variation from speaker to speaker masked it, so that a larger sample would be required for its detection. It would also have been desirable to use monolingual listeners in their home environment. I doubt, however, that this invalidates the results. All the listeners, though bilingual in English to some extent, were strongly native-language dominant and participated in a subculture associated with their language. The effect of accentual structure upon rhythmic perception of non-speech sounds, if it exists, is unlikely to be of the same magnitude as the general psychophysical effects discussed earlier.

The results clearly show the effect of the sensory dimension of prominence. This appears even though the degree of prominence in each dimension is not as great as in some earlier studies, having been chosen as roughly comparable to values found for accented elements in speech. This may explain why no greater preference for initial position was found for the intensity sequence as for the initial control sequence.
Table 2. Number of responses reporting an accented pattern.

<table>
<thead>
<tr>
<th></th>
<th>48</th>
<th>48</th>
<th>48</th>
<th>48</th>
<th>48</th>
<th>48</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accented</td>
<td>24</td>
<td>44</td>
<td>44</td>
<td>43</td>
<td>46</td>
<td>31</td>
<td>39</td>
</tr>
<tr>
<td>Uninterpretable</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>No pattern</td>
<td>38</td>
<td>4</td>
<td>7</td>
<td>3</td>
<td>17</td>
<td>9</td>
<td>38</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Initial</th>
<th>Final</th>
<th>Length</th>
<th>Frequency</th>
<th>Intensity</th>
<th>Tones</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>7</td>
<td>18</td>
<td>1</td>
<td>1</td>
<td>12</td>
<td>I</td>
</tr>
<tr>
<td>15</td>
<td>15</td>
<td>30</td>
<td>1</td>
<td>2</td>
<td>15</td>
<td>II</td>
</tr>
<tr>
<td>22</td>
<td>10</td>
<td>32</td>
<td>2</td>
<td>1</td>
<td>22</td>
<td>III</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>10</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>I</td>
</tr>
</tbody>
</table>

Table 1. Number of responses according to reported position of

The table shows the frequency and intensity of responses for initial and final positions, as well as the length of the sequences. The sequences are matched for total, length, intensity, frequency, and control.
The similar preference for initial position in the perception of the frequency and intensity sequences supports Handel's analysis of frequency patterns.

The preference in length sequences for initially rather than finally accented groups accords with previous research, since the long/short ratio, 1.4, is less than the reported region where a shift to final accent preference appears. The relative preference is also about the same as the listeners exhibit for the frequency and intensity sequences.

What is surprising is that almost half of the responses indicated a rhythmic group with the longer sound in medial position ( / - / ). This is probably related to the fact that the longer sound was followed by a correspondingly short interval, since the repetition rate was held constant. The longer sound and its successor would then tend to be perceived within a rhythmic group.

In Handel's terms the sequence is composed simultaneously of a length pattern

\[
\ldots \ u \ u \ u \ \ldots
\]

and an interval pattern

\[
\ldots x \ x \ x \ x \ x \ x \ldots
\]

The results can then be explained thus: Some listeners perceived the length pattern, preferring the rhythmic group \( \ldots u \ u \) to \( \ldots u \ u \ u \). Others perceived the interval pattern, preferring the rhythmic group \( \ldots x \ x \ x \) to \( \ldots x \ x \ x \ x \).\footnote{This assumes that listeners were able to identify the interval organization \( x \ x \) with the notation \( - / - \).}

Does rhythmic perception affect accent placement? Recent contributions by Martin (1972) and Allen (1975) argue persuasively for the rhythmic basis of accent. From the perceptual point of view, Martin proposes general organizational principles underlying accentual systems and Allen points to the role of the physical characteristics of speech. I find it difficult to conceive that accent could be unrelated to rhythmic perception in the face of the evidence they marshal. But the mere existence of a relationship is not important for the enterprise of discovering the secrets of accent. What matters is the magnitude of its effect. Is it a major or minor actor in accentual events?

Non-speech stimuli in rhythmic research. Given that there are fundamental differences in the perception of speech sounds and non-speech sounds, the non-speech nature of the stimuli used for most research on rhythmic perception causes some concern. Clicks, hammer bangs, buzzes, and tones are the usual sound elements. The stimuli are usually presented at a slower than conversational syllable rate – 2 to 3 times per second instead of 4 to 6. The differences in intensity and length used to mark prominent elements are often greater than those usually found realizing accent in speech, e.g. 15-25 dB differences in intensity and length ratios of 3.0 to 5.0. Handel (1974) is the only study to consider the rhythmic perception of sequences marked in more than one sensory dimension, the usual case in speech. Finally, although pitch rise or fall is a common and highly effective mark of accent, its role in rhythmic perception has not been determined.\footnote{
Accent placement. How might the regularities of rhythmic perception influence accent placement? Seeking correlation with the synchronic distribution of accent placement appears relatively unpromising. It is unlikely that we will find that "languages with strong tonic accent...will have rhythmic groupings with the strong syllable leading; languages with accent based on duration...will have the strong syllable last." (Allen 1975:78). If perceptual tendencies were the dominant factor in accent placement, we should expect to find a much greater proportion of languages with initial stress than reported in Hyman's survey (1977). Nevertheless, the different nature of frequency, intensity, and duration suggests that typological and other studies of accent not entirely neglect their interaction with accent placement. It would be worth checking, for example, to see whether the distribution of accent position is different for languages with contrastive vowel length than for those without it; or whether there is a systematic difference between constraints on the occurrence of vowel length contrast in initial-stress languages, e.g., Czech and Finnish, and final-stress languages, e.g., Alabaman, Tübatulabal, and Turkish. Or, in languages with variable accent position, does position influence the sensory dimension of accent realization?

As with many complex linguistic phenomena, the trajectories of change in accentual systems should be more revealing. Here there should be more opportunities to distinguish possible perceptual influences from the influence of other competing explanatory factors such as the productive control of intensity, pitch, duration, etc., and interaction with intonation. Some places to start are suggested by papers in this volume. Hyman, analyzing Latin stress development, conjectures that there may exist a tendency to avoid too many unstressed syllables in a row, facilitating a shift $\text{SSSS} \rightarrow \text{SSSS}$. This is in conflict with the principles of rhythmic perception, suggesting the existence of an antagonistic mechanism. The pitch accentual patterns of Japanese, discussed by McCawley (1976) are quite similar in structure to those studied in detail by Royer and Garner (1966, 1970) and Handel (1974). A comparison of their hierarchies of difficulty and the paths of change in Japanese accent is obviously suggested. The same holds for studies of the development of accentual systems from tonal systems.

Do all languages have accent? If accentual systems have in substantial part a rhythmic basis with organizational function, then how could a language do without one? It is not surprising that the implicit answer of Martin (1972) and Allen (1975) is that they cannot. Both state that all languages have an accentual system, apparently on a prioristic grounds. Many tonal languages have no reported accentual system, and Hyman (1977) lists sixteen languages reported to have neither tonal contrast nor accent. Many of these may be cases of inadequate descriptions. Others are probably not. Examples cited at this conference are Greenlandic Eskimo (Roy Wright, personal communication) and Hausa (Will Leben, personal communication). It may be, however, that such languages possess characteristics that provide a rhythmic organization. Greenlandic Eskimo possesses strong and weak syllables that may function to provide an interval-marked rhythm. This is also true of some other languages like Amharic
reported to have no or very weak stress. Their careful examination in this light may extend our understanding of the rhythmic aspects of accent.

FOOTNOTES

*This research could not have been done without the generous assistance of Parviz Birjandi, Subhendu Datta, John Durrett, Zygmunt Frajzyngier, Ali M. Hedayati, Walter Kintsch, Ronald Leathem, Richard Olson, Carmen Plomin, Richard Sweetman, the anonymous listeners who participated in the experiment, and the facilities of the Computer Laboratory for Instruction in Psychological Research and the Phonetics Laboratory of the University of Colorado. I am also indebted to Dwight Bolinger, Larry Hyman, and Allan Taylor for their comments and suggestions. I express to all of them my grateful appreciation.

1 For an explanation of what an accentual system is, see Hyman (1977).

2 This does not accord with the conclusion of Woodrow (1911). However, his method of relying upon the "final" subjective report of his listeners would seem to be vulnerable to experimenter bias, so there may be room for disagreement here.

3 For descriptions of stress in Bengali, French, Persian, and Polish see Ferguson and Chowdhury (1960), Delattre (1960), Ferguson (1956), and Jassem (1959) respectively. A language such as Czech, Hungarian, or Finnish, with word-initial accent, would have filled out the typology. It was not possible to find suitable listeners for such languages, however.

4 The standard deviation of temporal intervals in the stimuli was 6 msec.

5 No statistical test is completely appropriate for the hypothesis of no overall difference among languages, because of the small samples in some conditions. It is possible to compare the most divergent pairs of languages, although this procedure is non-conservative, having the effect of making differences appear more significant than the actual case. French and Persian, though both possess final stress, are the most divergent pair for the intensity and frequency sequences. The difference in frequency of perception of initially prominent groups is nevertheless not significant ($\chi^2 = 2.0$, d.f. = 1, $p \approx .15$). Nor is the difference for medial position significant (Fisher exact probability $p = .09$). For the length sequences, Bengali and English are the most divergent pair, but not significantly (Fisher exact probability test for frequencies of perception of initially prominent groups gives $p = .10$; difference for medial position
not different by inspection.) The only significant pairwise differences are found in the responses to the initial control sequence between French and Persian, and between French and English (Fisher exact probability tests are significant at $p = .025$ and $.05$, respectively, for perception of initially prominent groups.). Again, this must be regarded skeptically, owing to the non-conservative statistical procedure.

Responses to the control sequences did, however, differ, from the marked sequences in another way. Fewer responses to the control patterns clearly reported an accented pattern. See Table 2. The difference among sequences is significant ($\chi^2 = 18.5$, d.f. = 4, $p < .001$), although the effect of replacing subjects should be taken into account.

Fraisse, however, explicitly specifies this type of sequence, not one with equal intervals between sounds, in summarizing the perception of length-accented sequences (1956:97–8).

Handel's results for the perception of interval patterns of eight elements do not lead to a prediction of which of these three-element patterns would be preferred.

Some evidence for similarity of rhythmic behavior for speech and non-speech does exist, e.g. the research on temporal control of rhythm reported by Allen. His suggestion that a general perceptual effect is found for speech in the perception of French phrases is doubtful, however. He attributes this to the realization of French accent by duration, postulating that the effect of duration dominates the accompanying pitch rise. Even ignoring the likely syntactic basis of rhythmic grouping, the usual length ratio in final position is 1.8 (Delattre 1966), probably not sufficient to cause a final-accent preference. And the assumption that duration dominates frequency is not supported by such perceptual evidence we possess (e.g. Handel's).

One possibility is that in longer words groups of two syllables can function as units in the rhythmic pattern. Royer and Garner (1970) found that repeated alternating groups of two facilitated pattern perception and learning. Then the rhythmic group still has initial prominence. This does not explain the shift in the initial dyad; it is however consonant with the absence of languages with third syllable accent.

REFERENCES


Hyman, Larry. 1977. On the nature of linguistic stress. (in this volume)


Woodrow, Herbert. 1911. The role of pitch in rhythm. Psychological Review 18, 54-77.

Putting the EmPHAsis on the Wrong SYLLAble*

Victoria A. Fromkin

University of California, Los Angeles

There was a time when people used to joke about putting the emphasis on the wrong syllable, by deliberately placing the stress on the second syllable of the words 'emphasis' and 'syllable'. That such a joke should have 'made the rounds' reveals that speakers of English are aware of stress placement and also aware that errors in the placement of stress are produced and recognized by hearers as deviant. This is not surprising in that knowledge of stress is part of the knowledge that speakers have about the languages which they speak. It is also not surprising that speakers produce speech errors in spontaneous speech which involve stress and/or accent. An analysis of speech errors provides evidence for testing linguistic hypotheses concerning the phonological representation of lexical items and the presence or absence of stress rules. It is the purpose of this paper to exemplify this claim and to hopefully provide some further insights into the phonology and the phonological description of English.

The examples used in the paper are taken from the UCLA corpus of speech errors, collected over the last number of years. They are all spontaneously produced 'slips of the tongue'. Boomer and Laver (1968) define a slip of the tongue as 'an involuntary deviation in performance from the speaker's current phonological, grammatical or lexical intention.' A deviation from a speaker's phonological intention can be illustrated by errors such as those given in (1). The intended or target utterance is labelled TU and the actual utterance, in which the error occurs, as AU.

(1)

a. TU: speech error
   AU: peach error
b. TU: fill the pool
   AU: fool the pill
c. TU: big and fat
   AU: pig and vat

A deviation from a speaker's grammatical intention (syntactic target) is illustrated by the examples in (2).

(2)

a. TU: He never forgets/
    He never does forget/
    He doesn't ever forget
   AU: he ever doesn't forget
b. TU: She made him do the
    assignment over.
   AU: she made him to do the
    assignment over.
c. TU: they seem to know where the problem is
   AU: they seem they know where...

d. TU: she swore me to secrecy
   AU: she promised me to secrecy

A deviation from a speaker's lexical intention is illustrated by a word blend or word substitution such as those in (3).

(3)

a. TU: terrible/horrible
   AU: horrible
b. TU: smart/clever
   AU: smeever
c. TU: she's marked with a big scarlet A
   AU: ...with a big scarlet R
d. TU: lady with the dachshund
   AU: lady with the Volkswagon

An examination of speech errors, then, shows errors made on the phonological, syntactic, and semantic/lexical levels. These may be errors in which phonemic segments, features, syllables, morphemes, words, phrases are anticipated, persevere or reverse, or they may be errors in which such units are deleted, added, or moved. They may also be errors in which 'rules' are incorrectly applied or where they are not applied when they should be. (Cf. Fromkin, 1973 for examples and discussion.)

Slips of the tongue therefore seem to provide clues as to "how the speaker's linguistic information (his grammar) interacts with other psychological mechanisms in the production of behavior" which, as is suggested by Fodor and Garrett (1966) is the aim of a linguistic performance model. The construction of performance models is however impossible without the assumption of an internalized grammar, which has been called, perhaps unfortunately, linguistic competence. The terms linguistic competence and performance, and the concepts, have been much debated recently. Derwing (1973), for example, asks: "Of what possible utility is a 'competence' model, after all, apart from an associated performance model?" He believes that "the first order of business for linguistic theory is the construction of tentative models of linguistic performance." This is also the view of Linell (1974) who states that "competence as typically conceived of by Chomsky is not a fruitful and workable notion in an empirical approach to language and linguistic behavior. It is much too abstract and abstruse to provide any real explanations of linguistic reality."

To criticise a theory as 'too abstract' is unworthy of a scientist. One is reminded of Einstein's statement that "the hypotheses with which a scientific theory starts become steadily more abstract and remote from experience" the greater the number of empirical facts the logical deduction for the basic postulates includes. (Einstein, 1934) Einstein concluded that "the predominantly inductive methods appropriate to the youth of science are giving place to tentative deduction...the theorist who undertakes such a labour should not be carped at as fanciful; on the contrary, he should be encouraged to give free reign to his fancy, for there is no other way to the goal."
This seemingly irrelevant (and overly simplistic) excursion into the philosophy of science is prompted by my belief that we can not understand the nature of speech errors if we fail to distinguish between the underlying grammar which determines the form of the target message and the actually produced utterances.

Derwing wishes to replace the notion of a grammar as 'a system of rules' which express the basic regularities of a language, with a notion of 'linguistic rule which can be directly elicited from surface structures'. But how can the linguistic rule concerning, for example, stress placement in noun phrases as opposed to noun compounds be directly elicited from surface structures' when one finds actual utterances like those in (4)?

(4)  
   a. TU: by averaging the six scores     AU: by averaging the six scores  
   b. TU: a three time loser         AU: a three time loser  
   c. TU: he lives in the big white house     AU: ...in the big white house  

Derwing suggests "that we replace Chomsky's abstract notion of rule with a reconceptualization specifically designed to represent part of a model of linguistic behavior (a performance model), that is, a model in which 'putting rules to use' means simply behaving according to the rules. This decision has the important immediate consequence" he continues "of implying that one kind of evidence is necessary if we are to justify the formation of any particular rule: we must demonstrate that the linguistic behavior of the speaker, at least is 'regular' in the manner stated by the rule... We must demonstrate, in short, that the speaker does behave according to the rule...We may then postulate that the rule in question is a general surface-structure constraint (or 'output condition') on the form of utterances, and the language user has learned he must conform to it."

Derwing concludes "Under a direct behavioral interpretation of the sort I advocate, it will no longer be possible to hide behind a competence/performance distinction...."

Suggesting that such a distinction exists, and that such a distinction makes it possible for us to account for otherwise unaccountable phenomena is not hiding behind anything. Speech errors cannot be understood, I believe, without this distinction, and, in fact, speech errors would invalidate any rule which Derwing would hope to propose.

Using Derwing's behavioral criteria we find that some posited rules do act as 'output conditions' some of the time. I shall discuss such cases in relation to stress rules. But even these rules are broken. Without the postulation of rules, however, we are unable to distinguish between error free utterances from those in which errors occur. Speakers and listeners are able to detect errors as errors. The perception
of an error requires knowledge that a rule exists and has been broken. Rules
therefore cannot be viewed solely as 'output conditions' because the output violates
the conditions.

The evidence from speech errors also argues against Derwing's statement that
"no one has ever demonstrated...that generative grammars are acquired by the
learner and put to use by the speaker or hearer or even that they provide the devices
that are employed....in the production and understanding of speech" or "underlie the
observed use of language in any way." The relationship between the grammar and
performance is revealed by the kinds of errors I shall discuss in this paper, and that
I and others have discussed previously. But this does not mean, as has been repeated
so often, that the rules posited in a grammar to account for regularities in the language
are rules necessarily used in the form or ordering which occurs in the grammar.

An analysis of speech errors provides us with 'a window into the mind'; errors
provide support for the kinds of constructs, structures, rules of grammar which are
not directly observable. Speech errors are observable 'facts'. They can be explained
by reference to abstract grammatical rules and representations. They also help us
to understand how we use these rules in speech performance -- when we actually
speak, thereby giving us some insight into the mechanisms of speech production.

It is of interest to examine speech errors in relation to the role of stress in
performance, and the role of stress in English grammar.

Nearly all the researchers who have been concerned with the analysis of speech
errors have noted that errors more frequently occur in stressed syllables than in un-
stressed. Garrett (1975) referring to the findings of Boomer and Laver (1968),
MacKay (1970), Fromkin (1971) and to the analysis of the M.I.T. corpus states that
"There is a clear relation between the involvement of a speech segment in an error
and both word and phrasal stress." MacKay (1970) was particularly concerned with
explaining the mechanisms which produce speech errors and concludes from the
frequency of the involvement of stressed syllables that there is a "higher level of
subthreshold activation (in stressed elements) than unstressed ones." Garrett
presents an alternative hypothesis, "namely, that (such errors) are a consequence
of the syntactic and morphological structure which underlies prosodic features." He
suggests that the MacKay hypothesis is weakened since errors of perseveration
occur as well as errors of anticipation and transposition. That is, if the phonological
errors were due primarily to the preactivation of stressed syllables because they are
in some way 'stronger' one would not expect this to show up in phonemic segments
or syllables which persevere. Yet, out of a random sampling of 427 simple consonantal
errors (errors in which only single consonants are involved) 26% were perseveration
errors, 38% anticipation errors, and 36% reversals.
The UCLA corpus of speech errors supports Garrett's position. First of all, the greatest number of errors occur in lexical formatives as opposed to grammatical elements. Since every 'content' word has at least one stressed syllable this raises the probability that stressed syllables will be more frequent. The large number of monosyllabic words involved in errors further increases the probability.

The computer program written by Georgette Silva (Fromkin et al. 1976) to analyze the UCLA corpus has to date analyzed only 612 errors. But the analysis does indeed show the greater frequency of stressed syllables involved.

At the top of the graph presented in Figure 1 we see that 82% of the target syllables in which segmental errors occur are stressed syllables. The earlier studies cited above, in addition to ignoring the role played by syntactic and morphological factors, also failed to consider the syllable position in the word. In our analysis, as is shown at the bottom of the graph in Figure 1, 73% of all syllables affected were word initial syllables. However, as shown by the middle bars, unstressed initial syllables count for only 2% of the errors.

It seems then that stress is indeed a factor, but so is syllable position. That both play a role is shown by the fact that 88% of all the stressed syllables involved are word initial syllables, whereas only 10% of the unstressed syllables are initial. We have yet to analyze these syllables in relation to grammatical category and syntactic function. When this analysis is complete we should have a better picture of the behavioral attributes of stressed vs unstressed syllables in speech performance.

This frequency analysis may tell us something about the mechanisms involved in both normal and deviant speech. It tells us little about the linguistic role of stress beyond suggesting the preponderance of word initial stress in English.

The examples in (5) are somewhat more interesting from a linguistic point of view.

(5)  
  a. TU: tremendously  AU: tremendous [trɪ'mendəs]  
  b. TU: specificity  AU: specific [spə'sifəti]  
  c. TU: polysyllabic  AU: polysyllabic [paʊlɪ'sɪlabɪk]  
  d. TU: hospitable  AU: hospitable [hɒsɪ'tabəl]  
  e. TU: extended care facility  AU: extended care facility [ɪn'tɜːdʒɪst ˈkeə fæˈlɪtɪ]  
  f. TU: San Vicente Blvd  AU: San Vicente Blvd [ˈvɪntə]  
  g. TU: infinitely many  AU: infinitely many  
  h. TU: philosophical  AU: philosophic [fəˈlɑːsəfɪk]  
  i. TU: obligatory  AU: obligatory [əˈblɪtərɪ]  
  j. TU: prosodic phenomena  AU: prosodic phenomena [prəˈsɒdɪk fəˈnəmənə]
Stressed (S) vs Unstressed (S) and Initial (Si) vs Non-initial (Sn) syllables in segmental errors

Figure 1
Notice that in these errors a syllable is deleted, added, or parts of words are transposed. As a result, the stress does not remain on the vowel which would have been stressed in the target utterance, but shifts in accordance with English stress rules. In addition, when the stress has moved to another vowel the emerging stressed vowel is given a full vowel articulation rather than the target schwa vowel quality. This is true even when that vowel is never produced in non-deviant utterances other than as a schwa, as in tremenly or San Vinte.

Note also that with the stress shift non-reduced vowels in the target utterance surface as reduced vowels. Such errors at least suggest that lexical items are represented in our mental dictionaries with unreduced vowels, i.e. that schwa is not phonemic in English (as was claimed earlier by Bloomfield). Since there are words in English with vowels which are phonetically always reduced it is conceivable that, if indeed speakers represent all vowels in unreduced form, these non-alternating schwas will be represented differently from speaker to speaker. There certainly appears to be a spelling effect on the vowels which emerge. That orthography influences the grammar should not be surprising, nor should it be regarded as negating the hypothesis put forth here. It is even possible that illiterate speakers of English would show a different phonological representation or a different phonemic system. Whatever factors influence the lexical representation, these errors provide some support for the hypothesis that speakers construct a representation of these surface schwas with full vowel qualities even when they never surface in other than reduced form (except in errors).

Further support is provided by Burton (1975). She reports on a case study of an emotionally disturbed 15 year old girl with severe dysprosody. The patient’s speech is completely monotonic. An analysis of her speech shows that syntactic stress is realized phonetically by increased duration and intensity with no change in pitch contour. More relevant to the discussion is the fact that the patient "does not reduce vowel quality of unstressed vowels; thus [open] rather than [opən]." Since she can read it is again possible that the orthography has influenced her phonological representation. It is, however, of interest that in the case of phonetically reduced vowels, she stores the morphemes with non-reduced vowels.

It has been suggested that an alternative hypothesis may account for the kinds of errors given in (5), namely that Vicente, for example, is pronounced [vɪntɪ] because of its orthographic similarity to a word like Vincent. For non-Los Angeles readers it may help to state that San Vicente in its normal pronunciation is [sæn vɪsəntɪ] with the first syllable of Vicente always pronounced with a schwa. It should be further stated that the error was made by someone who has lived in Los Angeles for many years and except for this error has always pronounced the name of the street as given above. I cannot disprove this alternative view. Nor can I prove that with the vowel deletion error in (5) e., the stress shift and vowel quality of facility [ˈfæsəlɪ] was not due to the influence of the orthographically similar word facile. I can merely
reiterate that such changes in vowel quality which accompany the stress shifts (due
to other errors in the string) at least are suggestive and provide some 'behavioral'
evidence in support of the hypothesis regarding vowel representation as non-schwas
in the lexicon.

However one wishes to view the question regarding the representation of reduced
vowels, the errors in (5) and those in (6) do point to the productivity of stress assignment
rules.

\[
\begin{align*}
(6) & \quad a. \text{ TU: the two political parties } & \text{ AU: } \ldots \text{polítichal} && [\text{pölítikjal}] \\
& \quad b. \text{ TU: professorial ranks } & \text{ AU: } \text{professoral ranks} && [\text{profésoral}] \\
& \quad c. \text{ TU: Improvisation (improvise\textsuperscript{ation}) } & \text{ AU: } \text{improvis} && [\text{impravížan}] \\
& \quad d. \text{ TU: the derivation (derive\textsuperscript{ation}) } & \text{ AU: } \text{deriva} && [\text{derájval}] \\
& \quad e. \text{ TU: communication (communicate\textsuperscript{ation}) } & \text{ AU: } \text{communic} && [\text{komjúnekatéjšon}] \\
& \quad f. \text{ TU: no necessary correlation } & \text{ AU: } \text{no necessitous} && [\text{nassésotes}] \\
& \quad g. \text{ TU: subtlety } & \text{ AU: } \text{subt} && [\text{subtili}] \\
& \quad h. \text{ TU: the moment of reckoning } & \text{ AU: } \text{moment of reckoning} && [\text{móménti}] \\
& \quad i. \text{ TU: description (describe\textsuperscript{tion}) } & \text{ AU: } \text{describ} && [\text{descripción}] \\
& \quad j. \text{ TU: specializing in } & \text{ AU: } \text{special} && [\text{spécialižin}] \\
& \quad k. \text{ TU: explanations } & \text{ AU: } \text{explan} && [\text{explánin}] 
\end{align*}
\]

The errors in some of the examples in (6) seem to be derivational errors, i.e. the addition of a wrong derivational affix resulting in possible but non-occurring words. The phonetic representations of such derived words, including the stress placement, must either have been determined or 'fixed up' after the word formation rules applied, i.e. after the error occurred. The stress placement rules and the vowel reduction rules do indeed operate in performance. Such errors seem to support Derwing's proposal that behavioral data are sufficient to show the existence of rules. Unfortunately it is not always the case that the rules apply when and how they should apply; that is, the 'output conditions' may fail to apply.

None of the errors discussed so far are errors of stress placement. They are errors involving segments or strings of segments or affixes in which the stress and vowel rules are applied after (or simultaneously with) the occurrence of the error and which prevent a compounding of errors, thus acting as 'output constraints'. But stress errors do occur, showing that stress or prosodic features can be disordered just as segmental features can be disordered. Gandour (1976) has shown that other prosodic features, such as tone, can also be disordered. The examples in (7) illustrate this stress disordering.
Such errors argue neither for nor against stress being marked in the lexical representation of morphemes or words. If stress is not specified, then these errors show an error in the application of the stress rules. Errors occur at different levels and stages in the production process. These errors may be due to a disordering of stress features which have already been assigned, either by rule or by lexical specification. The previous errors cited however show that stress rules exist whether or not stress is lexically specified. With either interpretation, once the stress is incorrectly placed, vowel changes may occur, supporting the suggestion that phonetic vowel quality is determined after stress assignment.

So far only errors involving word stress have been discussed. In Fromkin (1971), it was pointed out phrasal stress often remains as in the target utterance even when words are disordered. "Two aspects of stress must be accounted for: first, the word stress moves with the word itself (i.e. the syllable of the word which receives main stress in isolation also receives the primary stress when the word is moved); second, the stress contour of the phrase is fixed by the syntactic structure of the phrase and must be generated independently of the word order in the utterance." Garrett also supports this finding, pointing to the fact that errors such as the ones already cited, must have "occurred prior to the realization of the particular phonetic form of the elements involved in the error." He further states that "it appears to be the position which an element occupies within an independently specified structure which endows it with the detail of its physical form." My own analysis supports Garrett's suggestion that "we should look to syntactic descriptions for the relevant analyses (in that)... prosodic features are determined by surface syntactic descriptions." I am sure he would agree with the qualification that certain 'prosodic features' are so determined. The examples in (8) show that primary stress, i.e. the intonation contour, often remains as in the target position when words are transposed.

(8) (phrasal primary stress = /; word stress = \)

a. TU: apples of that origin  AU: "origin [ərˈdʒən]
b. TU: mobility  AU: "mobility [məˈbiləti]
c. TU: so that we can progress  AU: "can progress, uh, progress
d. TU: economists  AU: "economists, I mean, economists
e. TU: differences  AU: "differences [difəˈrɛnsiəs]
f. TU: phonetic  AU: "phonetic [fənˈetɪk]

The examples in (8) show that primary stress, i.e. the intonation contour, often remains as in the target position when words are transposed.
These examples are typical of what happens when words or morphemes are disordered. The phrasal stress does not move with the vowel but remains in the same position as in the intended structure. Word stress must be assigned or specified prior to phrasal stress, despite the fact that they are realized physically in the same part of the signal. That stress is determined by syntactic structure is further shown by the errors in (9).

(9)  

a. TU: I'm paying for it all together  
    AU: I'm pay foring it...  

b. TU: I have to go down town to get a transformer for my typewriter  
    AU: ...to get a transformer typewriter  

c. TU: for a good many years  
    AU: for a many years  

d. TU: I think it's reasonable to measure with care  
    AU: I think it's careful to measure with reason  

e. TU: he's undergone analysis  
    AU: he's gone under analysis  

These errors show that stress rules 'monitor' the utterance prior to production. When certain errors occur, the rules may reapply to correct an error in the placement of primary stress or word stress. Thus when the -ing suffix is disordered as in (9)a, the 'word' foring has the structure requiring stress placement. In b, the deletion of for my creates a possible noun compound and stress is changed appropriately. This correction of possible additional errors is also shown in d, where care is changed to careful making the sentence syntactically well formed if not semantically well formed.

I do not mean to imply that such monitoring always prevents new errors, nor that ill-formed strings do not occur. If our filters worked at all times no errors would occur. Errors in rule application also occur as shown by the operation of the regular past tense rule to irregular verbs and other such errors like those in (10).

(10)  

a. AU: *he swimmmed in the pool nude.  

b. AU: *he's more bigger than she is.  

c. AU: *it hasn't begun to rung yet.  

d. AU: *he scared the stiff out of all the students.  

e. AU: *she made him to do the assignment over.
Such incorrect rule application also occurs with nuclear and non-compound stress, as shown in (11).

(11)  

a. TU: Saul said 5000 marines were covered up by landing nets.  

b. TU: the town farm  

c. TU: pre season workouts  

d. TU: cancer causing qualities (or) cancer causing qualities  

e. TU: the Jerry West Night game  

f. TU: Larry Hyman's paper  

AU: ... by landing nets, uh I mean by landing nets.  

AU: the town farm -- the town farm  

AU: pre season work outs.  

AU: cancer causing qualities.  

AU: the Jerry West Night game  

AU: Larry's Hyman paper  

(then corrected)  

(corrected)  

In f. It may be that the stress error occurred because of the disordering of the possessive morpheme which thus changed the syntactic structure of the phrase.

All these examples have hopefully shown that stress rules are productive rules, that stress features are independent of segmental features, and that word and phrase stress are independent of each other. Deviant utterances therefore can help us understand something about the grammars which interact with speech mechanisms and something about normal speech production.

REFERENCES


Fromkin, V. A. 1971. 'The non-anomalous nature of anomalous utterances' Language, 47. 27-52.


Linell, P. (1974) 'Problems of psychological reality in generative phonology' _Uppsala University, Department of Linguistics._
THE IMPORTANCE OF TURNING POINTS
FOR THE PITCH PATTERNS OF SWEDISH ACCENTS

Eva Gårding
Lund University

A Swedish word has one of two accents, Accent 1 (acute) or Accent 2 (grave). For the larger part of the vocabulary the accent is predictable from the structure of the word and vice versa. For instance: if you recognize a noun in the definite form in the speech chain with Accent 1, say, the famous 'anden, you know that the indefinite form must be the monosyllabic 'and, which means wild duck. If you hear the same sequence with Accent 2, "anden, you can infer that the stem is a bisyllabic vande which means spirit. A speaker or listener who cannot distinguish the accents from each other is not in very great trouble. In connected speech the meaning is practically always clear from the context. The Swedish speaking population in Finland for instance communicates quite easily without making any accent distinction.

The accent assignment rules apply rather uniformly to all dialects but the pitch patterns differ a great deal as you will see in a moment. Actually the location of the turning points in these patterns has been used for two purposes which will serve as the outline of this paper, namely

1. Typology

The typology is based on Ernst A. Meyer's collection of accent patterns from different parts of Sweden and Scandinavia. (Fig. 1).

For each dialect the figure shows slightly schematized pitch curves from bisyllabic words in final position in statement intonation. The vertical line in each accent matrix denotes the syllable boundary. In the left hand corner there are five pairs of examples from Stockholm. Accent 1 is characterized by one turning point, a peak, at what is considered to be the boundary between the first and second syllable. Accent 2 is characterized by two peaks, one in each syllable.

At first glance, the dialectal variety of the accent manifestations looks rather impressive. But it turns out that these curves can be divided into two groups each
with two subcategories depending on the number and the location of the turning points in relation to the syllable.

Figure 2 shows my analysis of the accent manifestations and their geographical distribution.

Examples of the accent types are shown in the right hand corner of the figure:

Zero: No accent distinction.

Type 1: One peak for Al and one for A2.
   1A: Early in the first syllable for Al and late in the same syllable for A2.
   1B: Late in the first syllable for Al and early in the second for A2.

2: One peak for Al and two peaks for A2.
   2A: Al has a late peak in the first syllable.
   2B: Al has a late peak in the second syllable.

The map shows that the geographical distribution of the manifestation types co-occurs with well known dialect areas established mainly on lexical and morphological criteria. There are single peaked accents 2 in the south, that is the former Danish provinces, in Bergslagen and in Gotland, double peaked accents in the Göta dialects and the Svea dialects.

2. Generative schemes for Swedish intonation.

The second part of my paper shows how the turning points have been used to make generative schemes for Swedish intonation.

Our goal is to have a model that generates pitch curves for all kinds of sentences regardless of dialect and structure.

The method is to compare pitch tracings of accents in different types of dialects and different contexts and try to separate the contributions of the accents and the other variables to the total pitch curve.

I shall here limit myself to a single dialect, Central Swedish, and a limited material. In Figure 3 you have a sample of phrases to which our model in its present state can assign the correct pitch patterns. (The model is the result of joint work, notice in particular the references to Gösta Bruce.) The sample is taken from material which consists of noun phrases uttered as statements with all four combinations of accents and with a varying number of unaccented syllables between the accents. We have given preference to test words consisting of sonorants, and high vowels have been avoided in the accented syllables. The speaker used three different stress patterns for the phrases. One of them indicated that both the
Observed pitch patterns in phrases with two accented syllables

Stockholm
adjective and noun were equally important information, the other two stress patterns
gave prominence (focus) to either the adjective or the noun.

The curves showed that as far as pitch is concerned only two patterns were
used. When both words are important in this dialect, we say that prominence has
a pitch manifestation in combination with the last accent.

The model is presented in Figure 4. It can generate pitch for an input phrase
that is syllabified and has information about accent and prominence. This information
is also needed to give appropriate durations to the syllables. At the top of the page
there is a typical input phrase. The rules – or commands as they are called here –
represent the linguistic components of our model, accent, prominence and sentence
intonation. These rules are output oriented and prescribe pitch movements between
2 or 3 relative pitch levels, neutral, mid, high. High is not necessary for all speakers.

The Accent 1 command is a fall from mid at the beginning of the syllable to
neutral at the beginning of the vowel. Accent 2 has a similar fall but it is timed
differently. There are two adjustment rules that apply to accents before prominence.
It is assumed that the adjustment rules reflect the effect of tonal coarticulation. The
adjusted Accent 1 starts from a lower level. There is one prominence command, the
same for both accents. The intonation commands are also accent independent – a rise
up to the accent peaks before prominence and a fall down from the level reached in
prominence.

The conventions state what is happening when there is no command present.

Cl says that the beginning and end points of the curve are at a neutral level.
C2 is an interpolation rule.

As you can see at the bottom of the page the rules are linearly ordered but they
do not apply in a left to right order for an utterance. Rule 1 marks the pitch of all
the accents regardless of their position in the phrase. Rule 2 marks prominence
(phrase accent)

Rule 3 takes care of the phrase intonation which is here declarative. This
rule consists of two parts, a rise before the accents preceding prominence and
a fall on the accent after prominence.

Figure 5 shows the application of the rules.

The essential point of our model is that we have separated phrase prosody
from word prosody. Rule 1 operates at the word level, Rules 2 and 3 at the phrase
level. The pitch pattern of a phrase in this sample can be described as a rise-fall
GENERATIVE SCHEME FOR PITCH ($\nu$)

Central Swedish (Stockholm)

Phrases with two accented words of which one is prominent

Example of input phrase: en långre nunna 'a taller nun' [en-'lכn-re-'nun-na]

Pitch levels. Neutral, mid (high)

1. Accent commands

- **A1**. Full from mid at the beginning of the accented syllable, reach neutral in the beginning of V.
- **A2**. Fall from mid in the middle of the accented syllable, reach neutral at the end of the syllable.

2. Prominence command

- **P**. Rise to mid (high) after accent command.

3. Intonation commands (here statement)

- **S1**. Rise before accent followed by P.
- **S2**. Fall from P-level
  (P-level means mid or high)

Conventions

- **C1**. Beginning and end are neutral.
- **C2**. When all rules are applied the blank parts of the $\nu$-curve are filled in.

**Rules**

- **Rule 1**. A. Apply A1, A2 to accented syllables.
- **Rule 2**. P. Apply P after prominent accent.
- **Rule 3**. S1. Apply standard rise before accents preceding P.
- **S2**. Fall from P level
  a) above the fall of the last accent and delete this accent
  b) over the last syllable (timing of fall depending on preceding accent).
- **C**. Apply C.
Pitch patterns obtained by rules
in all cases. The accents are momentary setbacks on the way to the peak of the phrase which indicates that the preceding accent has carried the most important information.

The pitch curves give some interesting insight into the planning of speech. To get the correct pitch movement for the initial unaccented syllable in Figure 3 for instance, the speaker has to know what accent is going to follow. Our adjustment rules for the accents are most probably also examples of anticipatory coarticulatory effects. In sentence number 4 the speaker must know that a prominent word is going to follow in order to give the right height to the peak.

By comparing generating schemes for phrases derived from different dialects we expect to arrive at a better understanding of the dialectal variation that exists in Swedish intonation. Since one of the dialects – the Swedish spoken in Finland and in the far North – has no accent distinction – it will also be possible to contribute to the question of what is accent and what is stress.

REFERENCES

ON THE NATURE OF LINGUISTIC STRESS

Larry M. Hyman
University of Southern California

0. Introduction

Stress-accent (henceforth, stress) is a property found in many if not most languages of the world. It is often observed to play a key role in the phonological and/or morphological structure of languages, and has also figured highly in major developments in recent phonological theory. However, despite the recurrence of stress in descriptive studies, and despite its prominent position in phonological theory, relatively little is known about stress as a general (non-language-specific) phenomenon. The severity of the gaps in our current understanding of stress is made embarrassingly clear by our inability to provide satisfactory answers to such questions as the following: (i) what is stress? (ii) why do languages have stress? (iii) where does stress come from? (iv) what is a natural stress rule? (v) how does stress change through time? Of these five questions, only the first can be said to have been given the attention it rightfully deserves. The result is that most treatments attempt little more than a definition of stress, and that phonological and morphological rules involving stress have rarely been subject to the kind of naturalness conditions imposed on other rules. Needless to say, a general theory of stress change has yet to emerge.

The purpose of the present paper is to present the results of a cross-linguistic study of stress phenomena which has grown to encompass approximately 400 languages from all parts of the globe. My immediate goal has been to uncover cross-linguistic generalizations which may have a bearing on the answers to the above questions. In the following sections I shall address myself to general principles of stress placement, and then to the role of syllable weight in stress languages. But first, it will be necessary to summarize what is meant by "stress" and "stress language".

1. What is a Stress Language?

In an unpublished paper delivered at a Summer Linguistic Society of America meeting, McCawley (1964) raised the question "what is a tone language?". Since that time a number of papers and publications have addressed themselves to the linguistic nature of tone, as opposed to stress-accent and pitch-accent. As seen from the hyphenated nature of these terms, two types of accent are distinguished
which are represented, respectively, by languages such as English and Japanese.

While recognizing that tone and stress are not mutually exclusive, in this section I would like to raise what may appear to be the opposite side of McCawley's question; namely, "what is a stress language?". Attempts to characterize the difference between accent and tone, where stress is one kind of accent (see below), have been based on either 1) the linguistic function, or 2) the phonetic realization of the appropriate prosodic features. European phonologists such as Trubetzkoy (1939), Jakobson (1931), Martinet (1954, 1960) and Garde (1968) emphasize the culminative function of stress (accent) as a definitional requirement. While a tone language can allow a high tone to occur on more than one syllable of a word, the basic principle in a stress language is that only one syllable per word will receive primary stress. Thus, each stress signals a separate word unit (or other accentual unit) and, it is claimed, functions as an aid in processing utterances.

The culminative definition of stress therefore excludes cases where more than one (phonemic) primary stress occurs within a word. These are of two types. First, in a few languages such as Campa (K. Pike and Kindberg 1956), several or all syllables of a word can apparently be stressed. While data from such "multiple stress" languages are limited, it would appear that the notion of a stress-accent is not applicable in such cases. Rather, Campa may best be analyzed as involving a feature of intensity which is parallel with other features such as nasality and liprounding (cf. Pulgram 1969:375). Several other languages, however, including Nimboran (Anceaux 1965:39), Sarangani Manobo (Meiklejohn and Meiklejohn 1958:2), Tahitan (Cook 1972:231) and Yuma (Halpern 1946:150), have restricted or exceptional cases where a word has two phonemic primary stresses, sometimes even on adjacent syllables. Since this situation usually arises under well-defined circumstances, we needn't be too concerned. Although such sporadic exceptions do arise, there does not appear to be any tendency for multiple stress languages to arise via this pathway. The above languages should therefore still be seen as possessing a culminative stress-accent.

The culminative definition of stress is somewhat more vague in dealing with words occurring with no primary stress. It is usually stated that in a stress language each word (or other accentual unit) will have one and only one primary stress. Trubetzkoy's examples are mūka 'torture' and muká 'flour' from Russian. In a bisyllabic word, stress must occur either on the first syllable or on the second syllable. There is no word *mukā with stresses on both syllables, nor is there a word mūkā with no stress. Many stress languages however are described as having at least some words lacking a stress, e.g. Kitsai (Bucca and Lesser 1969:13), Saho (Welmers 1952:147) and Seneca (Chafe 1960:21). While it is often pointed out that certain "clitics", i.e. forms which attach themselves to other forms (which in turn take a stress), do not have an inherent stress, stresslessness seems to be of greater significance in the above languages. Thus, in Saho, Welmers points out
that unstressed words are "somewhat more common" than stressed words. Thus, it looks as though we will have to say that a stress-accent is culminative, but can contrast with an absence of stress, e.g. Saho /bəda/ 'son' vs. /bada/ 'daughter'.

Finally, it is not clear what can be said about languages where a stress is reported to be evenly spread over two (or possibly more) syllables. Such a case is reported as follows for Southeast Ambrym: "Bisyllabic major words are stressed on the first syllable if the second is short. Otherwise stress is variable or constant through both syllables" (Parker 1968:90). Languages which are reported to have stress on every long vowel or on every vowel preceding a consonant cluster, etc., are considered not to have a stress-accent (see Appendix).

With the above reservations, we have therefore established that stress-accent is culminative in nature. At this point stress languages are further differentiated: if the position of stress cannot be predicted from a word-boundary, one speaks of free stress. What this means is that stress will be at least in part lexical, as in Spanish examples término 'term', terminó 'I terminated' and terminó 'he terminated' (Martinet 1960:83-84). This lexical function of stress is also seen in Asmat éco 'kind of frog' vs. écé 'revenge' (C. Voorhoeve 1965:26) and Pashto gufa 'knot' vs. gufá 'pochard' (Shafeev 1964:5). On the other hand, stress can acquire a grammatical function. In the following Russian and Polish forms taken from Garde (1968:106),

<table>
<thead>
<tr>
<th>Russian</th>
<th>Polish</th>
</tr>
</thead>
<tbody>
<tr>
<td>raspráva</td>
<td>rozpráwa</td>
</tr>
<tr>
<td>raspráv</td>
<td>rózp raw</td>
</tr>
<tr>
<td>rasprávami</td>
<td>rozpráwami</td>
</tr>
<tr>
<td>'reprisal'</td>
<td>'discussion'</td>
</tr>
</tbody>
</table>

stress can be predicted on grammatical grounds. In the given nominal forms from Russian, stress falls on the same stem syllable throughout the declension. In Polish, on the other hand, the stress moves in order to remain in penultimate position in the three forms. Garde's examples thus illustrate the difference between the morphological function of stress (in Russian) and the demarcative function of stress (in Polish). In the latter case stress serves as a boundary signal: given a stress in Polish, the general case will be that the following syllable ends a word. We therefore distinguish between lexical and grammatical stress, with the latter being further differentiated into morphological vs. demarcative subfunctions. In any given language stress may function in any one or more of these, and in fact, the typical case is that a language with lexical stress may also have some grammatically predictable stress, and that a language with demarcative stress may be occasionally upset by morphological considerations. The final stress in terminó 'he terminated',
for instance, can be grammatically predicted on the basis of the -ó suffix meaning 'third person singular past tense', which always takes stress in Spanish. Other stress oppositions are lexically defined in Spanish (e.g. término 'term'), while the general (demarcative) pattern is penultimate stress (e.g. termino 'I terminate').

The other way stress has been dealt with is in its phonetic realization. A stressed syllable is frequently characterized by a pitch change, by greater duration, and by greater intensity, three factors which in that order Bolinger (1958) and others have shown to be important perceptual cues for stress. Bolinger has argued that the most important perceptual cue of primary stress is changing pitch. However, as he also notes, it would be a mistake to say that stress is changing pitch. Rather, as explicitly stated by Weinreich (1954:2) and Lehiste (1970:150), stress is an underlying mental phenomenon, which however must somehow be picked up by the hearer. I therefore choose to regard the above perceptual cues as "strategies" for the realization of stress. The experiments of Fry (1955, 1958), Bolinger (1958), Jasem (1959) and others have demonstrated that pitch may universally best fit our perceptual capabilities, followed by duration, and finally by intensity (see Lehiste 1970:125-132 for further discussion).

Thus, the two approaches (functional and phonetic) are not contradictory. In fact, as I will now demonstrate, certain cross-linguistic generalizations about stress cannot be explained unless both factors are taken into consideration. I will demonstrate this first in regard to general rules of stress placement, and second, in regard to the phenomenon of "syllable weight".

2. Stress Placement

I start with the assumption that lexical or unpredictable stress derives, historically, from grammatical stress. What this means is that a language which does not have stress will not directly develop lexical stress, but rather the latter will always involve an intermediate stage of grammatical (i.e. morphological or demarcative) stress. Many cases of the development of lexical from grammatical stress have been reported in the literature. The unpredictable stress contrasts found in most Romance languages, for instance, derives from an earlier distinction in vowel length. Thus, as seen below,

(a) Latin ánora \(\rightarrow\) ánora

Latin ad hanc hó:ram \(\rightarrow\) ancó:ra \(\rightarrow\) ancóra

the lexical contrast between Italian ánora 'anchor' and ancóra 'again, still' can be traced back to a vowel length contrast in Latin. In (a), since the penultimate syllable co contained a short vowel in open position, stress had to be assigned antepenultimately, according to the general Latin stress rule (see section 3 below).
In (b), stress was placed on the long vowel of the penultimate syllable co:. When
the vowel length distinction was lost, stress became unpredictable, i.e. lexical.
Other mechanisms by which predictable stress becomes unpredictable include the
loss of unstressed vowels, e.g. initially in certain dialects of Pomo (McClendon
1973:34ff) and finally in Biblical Hebrew (Gibson 1966:33), as well as the introduction
of openthetic vowels, e.g. in Modern Hebrew (Sadock 1970:543) and in certain
Iroquoian languages (Chafe, personal communication). Both of these processes
destroy a previously existing demarcative stress rule.14

Limiting our attention to demarcative stress, i.e. to stress which signals a
word boundary, it should be clear that the closer stress falls to that boundary, the
better it will fulfill its linguistic function. And not surprisingly, initial and final
stress are commonly attested in the world’s languages. It seems reasonable to
hypothesize that stress rules which require more calculation on the part of the
speaker and the hearer (for example, requiring that one count further from a boundary)
are less "natural" from a conceptual point of view than rules which require less calcu-
lation.15 Thus, while initial and final stress are common, no language has been found
which has regular stress on the third syllable of every word, and only six languages
(of which one is uncertain) have demarcative stress on the antepenultimate syllable
(see Appendix).16

Consider now the Table of Dominant Stress Placement given in the Appendix.
Looking at this chart there appears to be an asymmetry, as summarized below:

(a) languages with dominant initial stress ......................... 114
(b) languages with dominant second-syllable stress ................ 12
(c) languages with dominant penultimate stress ................... 77
(d) languages with dominant final stress ............................ 97

Although we have hypothesized that grammatical stress should tend to fall next to a
boundary, penultimate stress appears to be at least as natural as final stress, and
is observed to be almost as frequently attested in languages (77 vs. 97 respectively).
On the other hand, only 12 languages were found to have dominant second syllable
stress, as opposed to 114 initial-stress languages.17 The rare occurrence of second
syllable stress must therefore be explained.

The penultimate position, while it is equally distant from a boundary as is the
second syllable, is a much more natural place to put stress. Besides its great
frequency in the world’s languages, a number of language-specific facts reveal that
it is in some sense preferred over final position:

1) In Araucanian (Echeverría and Contreras 1965:134), stress is assigned to the
second syllable of a word; however, in bisyllabic words, where second syllable stress
would wind up in final position, stress can optionally be realized on the initial syllable.

2) In Czech (Jakobson 1926:615), primary stress falls on the initial syllable of a word; secondary or echo stress occurs on each successive odd-numbered syllable; however, in a five-syllable word, where stress is expected on the first, third and fifth (=final) syllables, in colloquial speech stress occurs on the initial and fourth (=penultimate) syllables.

3) In Tagalog (Bloomfield 1917: ), some verbs have stress on their penultimate, some on their final syllable; the latter, however, undergo a shift of stress to antepenultimate position in reduplication and other derivations involving prefixation; penultimate stress remains stable.\textsuperscript{18}

4) In Shiriana (Migliazza and Grimes 1961:35), utterances are divided into stress-groups or feet; stress will occur on the final syllable of a foot, unless that foot occurs before a pause, in which case stress will be realized on the penultimate syllable of the foot (cf. Onondaga (Chafe 1970:73ff)).

5) In Dyirbal (Dixon 1972:274) primary stress falls on the initial syllable of a word and is echoed on subsequent odd-numbered syllables; however, final syllables are never stressed (cf. secondary stresses in Mencmini (Bloomfield 1962:20), which do not occur word-finally).

We see from the above facts that languages often have mechanisms whose function is to remove stress from final position (often putting it in penultimate position). It is crucial that no language has general initial stress, except in bisyllables, where stress can optionally be realized in final position (contrast this with Araucanian above). That is, there is no corresponding tendency for stress to shift into final position.\textsuperscript{19}

There is, however, a tendency for stress to be shifted out of second position, as witnessed in Hawaiian, where a near mirror-image of the Czech situation is attested (Elbert and Keala 1961:6): stress is penultimate and on even-numbered syllables counting from the penultimate towards the beginning of the word. However, in a five-syllable word, where stress is expected on the penultimate as well as on the second syllable of the word (which occurs two syllables before the penultimate), the latter is frequently realized instead on the initial syllable, e.g. ʔélemakūle. It is important to recognize that no language will readjust such an "echo" stress so that it occurs in second position rather than initially.

The tendency to avoid second-syllable stress, as well as the somewhat lesser tendency to remove stress from final position, cannot be explained exclusively in terms of the underlying function of stress, but rather in terms of its phonetic realization. The purpose of stress is to highlight a given syllable in a word or other grammatical unit. And since pitch-change is the most effective cue of stress-prominence, languages will
tend to assign stress to those positions where a pitch-change will be most perceptible.

At this point I would like to suggest that a pitch fall is more basic as a strategy in the realization of stress than is a pitch rise. We know from the experimental work of Ohala and Ewan (1973), Sundberg (1973) and Hombert (1974, 1975), that a fall in pitch is articulatorily less complex than a rise in pitch (other things being equal), and this is borne out in contour simplifications in tone languages. 20 As seen in the following English words,

(a) perfect (verb) $\rightarrow$ perfect [\_\_\_] (H = high pitch, L = low pitch)
(b) perfect (adj.) $\rightarrow$ perfect [\_\_]
(c) perfection $\rightarrow$ perfection [\_\_\_\_]

the most common or neutral realization of primary stress involves a fall in pitch, but as seen in perfect in (b), not necessarily a rise in pitch. Similarly, a monosyllable such as boy or girl involves a fall in pitch and no rise. These facts have suggested to me that the "basic" realization of English stress may best be seen as the assignment of a falling HL pitch-accent, which as seen in (b) and (c), can extend over two syllables. 21 It is, of course, no accident that this realization of English stress corresponds with declarative intonation, which also involves a downward pitch curve (see Vanderslice and Ladefoged 1971 for an attempt to incorporate this fact into an accentual feature system). Linguists have often lamented the fact that the phonetic correlates of stress cannot be independently studied, because they are so intertwined with intonation. 22 I would like to propose, however, that it is in fact this inseparability which is a necessary requirement of a stress language. That is, stress-accent is an accent which incorporates intonational features in its realization.

It is of course well-known that an intonational contour can be superimposed on (and therefore affect) any kind of accentual system, and indeed any kind of tone system as well. However, in all but stress languages it is possible to isolate a phonetic feature (or features) which is distinct from intonation. In a stress language, the two are not distinct. Thus, one characteristic of stress languages which seems to be general is that a stress in utterance-final position will be realized as a HL fall in pitch, rather than as a H level pitch. This fall in pitch has often been attributed to intonation, and the last primary stress of a given grammatical unit can be said to be realized by incorporating an intonational falling pitch in just this fashion.

However, if the verb perfect in (a) were realized in English as a L pitch followed by a H pitch, i.e. as [\_\_\_] instead of [\_\_\_\_], we would not speak of English as a stress language. Such a possibility does in fact exist, as witnessed in Safwa (J. Voorhoeve 1973) and Kinga (Schadeberg 1973), two "restricted tone" languages of the Bantu family.
Other languages with H tone accents include Tolowa (J. Bright 1964:104), Cherokee (Bender and Harris 1946:14) and Mattole (Li 1939:50). It is significant, I believe, that in none of the above references is the accent referred to as a stress.

The interplay between stress and intonation is an intriguing aspect of the problem at hand, though perhaps less intriguing once we address ourselves to the origin of stress. It was said earlier that the function of a fixed stress is to demarcate grammatical units, normally words. Insofar as word boundaries are thereby made more distinct, the development of demarcative stress may serve as an aid in processing utterances (cf. Martinet 1960:83). Chiba (1935:89) sees stress as generally facilitating communication: "If the same word is uttered with a high pitch on the initial syllable by some, and on the final syllable by others, a great deal of inconvenience will be experienced, since the language is essentially intended for purposes of communication. Accordingly, within one community in which the same language is spoken, some sort of uniformity has been reached with regard to pitch and stress and they have gradually been fixed in certain forms". However, while the communicative advantages of having a stress-accent are obvious (and therefore, relatively few nontone languages lack a stress accent of some kind), nothing is said about where the seeds of such an accent are to be found.

The hypothesis which I would like to advance is that stress actually comes from intonation, and that this derivative relationship accounts for the interaction between the two. As an introduction to the position I am arguing for, consider the following description of Shilha: "...primary or heavy stress occurs on the last vowel of the stem" (Applegate 1958:9). The author later adds in a footnote, however: "It should be noted that the stress patterns referred to here apply only to utterances consisting of a single word. If the utterance contains more than one word, the stress is reduced slightly on all vowels except those in the final word. It can be said, therefore, that primary stress occurs only at the end of an utterance".

In a language such as Shilha, where primary stress is realized only in utterance-final position, it is hard to understand why one should speak of stress, rather than of intonation. The intonational pitch-fall which accompanies the last syllable of an utterance in Shilha is important, however, for our understanding of how word stress comes into being. In particular, we can hypothesize that a stress-accent comes into being when an intonational feature becomes associated with a grammatical unit smaller than a clause (where a pause is frequently expected). In other words, intonation becomes grammaticalized as word-stress when the suprasegmental features of pitch, duration, and intensity that would have characterized a word in isolation are encoded with the word, and thus come to function in words not in isolation.

This phenomenon of generalizing from a pause to a word boundary is quite general, I would like to claim. Thus, consider the parallel situation of a language which has the phonetically motivated process of utterance-final devoicing, i.e.
devoicing before pause. In this language, a word in its citation form will be realized, say, with a final voiceless consonant; the same word may, however, be realized with a final voiced consonant when in sentence-medial position (especially if the following word begins with a vowel). Final-devoicing may then become generalized to new environments, as when a speaker wishes to clearly identify the word boundaries for his listener. The result is that devoicing now characterizes voiced segments preceding any word boundary, whether that word boundary coincides with an utterance or pause boundary or not. Notice, however, that the resulting rule is not phonetically motivated, since there is no necessary phonetic pause between word boundaries. Nevertheless, the tendency for utterance-final processes to be generalized to word-final position cannot be underestimated (cf. Kim and Hyman 1973). The same may be said about processes which start out directly following a pause.

Coming back to stress-placement, it would now appear that initial and final stress correspond to the initial raising and final lowering of pitch as an intonational feature. The reason why stress is not frequently found on the second syllable of a word (or other grammatical unit) is that this would mean the assignment of a rising intonational contour, which, being articulatorily complex, is a rare phenomenon in itself.

Penultimate stress, on the other hand, feeds into the falling pitch which is characteristic of stress systems. It is generally accepted that final position is phonologically weak, i.e. whatever stands at the end of an utterance (or, by extension, at the end of a clause, phrase or word -- cf. footnote 24), is subject to various historical changes and/or loss. Both consonants and vowels have a tendency to devoice, for instance, or to drop in final position, and the same weakening processes can be observed with respect to pitch. In many African tone languages, a H tone is subject to lowering in utterance-final position, and a L tone typically undergoes what Stewart (1971:185) terms “downglide”. Thus, because of this lowering effect, a falling pitch realized on a final syllable will not be as prominent as a falling pitch realized over two syllables. A good example of where final stress is not as prominent as penultimate stress comes from Tagalog, where Bloomfield reports, first, that the pitch rise of a final stress is not as great as the pitch rise of a nonfinal stress (comparing like open syllables), and second, "... an accent on the last syllable of a sentence often entirely loses its pitch-rise" (1917:145).

Thus, to summarize, although final position is conceptually natural for demarcative stress, it is less natural phonetically for the realization of pitch prominence. There may, however, be a sense in which the functional nature of stress predicts penultimate over final stress, although this must be stated in speculative terms. If stress does have a culminating function as a definitional requirement, this should mean not only that one syllable per word received greatest prominence, but also that it receives its prominence at the expense of other syllables. I have emphasized
the importance of a pitch fall as a perceptual cue of stress. If stress is final, this pitch fall will have to be identified within the stressed syllable. If stress is nonfinal (e.g., penultimate), the general case will be that the stressed syllable has a high pitch, and the following syllable a low pitch. A number of language-specific facts suggest that a stressed syllable is more perceptible in terms of what follows rather than in terms of what precedes. If this is true, placing stress nonfinally will enhance the two functions of stress: 1) phonetically, the fall from H to L is enhanced; and 2) conceptually, the culmination of prominence is enhanced by virtue of the fact that a syllable lacking stress follows—cf. Bolinger (1962:35): "...if a syllable is to be accented it is convenient to have it flanked by unaccentable ones". While the above speculation remains to be verified experimentally, there can be no doubt but that what follows a stress is less prominent (and therefore contrasts more with the stress) than what precedes. This is stated explicitly for Portuguese as follows: "...in words that have antepenultimate or penultimate stress, the syllables that precede the stress are slightly stronger than those that follow it" (Camara 1972:76). In Palantla Chinantec, a tone language, "Pretonic syllables have tone but no stress; post-tonic syllables have neither tone nor stress" (Merrifield 1968:14). This would seem to suggest that syllables tend to weaken more after stress than before stress; but also, perhaps, that stress is better perceived by contrast with what follows, rather than with what precedes. 

We conclude, therefore, that a combination of functional and phonetic factors account for the high frequency of initial and penultimate stress. These positions are seen to be the most natural not only from frequency counts of dominant primary stress placement, but also, it is interesting to note, by investigating where secondary stress tends to fall, for example:

1) In Walmatjari, "In monomorphemic words of more than three syllables primary stress occurs on the first syllable, and a secondary stress occurs on the penultimate syllable" (Hudson and Richards 1969:183).

2) In Zoque, "Stress is usually on the penult of the stress group, with a secondary stress on the first syllable if there are three or more syllables" (Wonderly 1951:108).

3) In Auca (K. Pike 1964:425), there are two iterative stress rules ("wave trains"): on stems, the first syllable and every other syllable thereafter receive stress; on suffixes, the penultimate and every other syllable thereafter receive stress, e.g. CVCVCV#CVCVCV.

The cooccurrence of an initial primary and a penultimate secondary stress, or of a primary penultimate and an initial secondary stress is not fortuitous. Rather, it would appear that the intrinsic variations which give rise to stress-accent are always present, and can become grammaticalized at any time. These intrinsic variations
are, as I have argued, intonational in nature, and are ultimately derived from the
articulatory and perceptual universals which characterize human speech.

3. Syllable Weight

The second issue which arises in the study of stress is the phenomenon of
syllable weight. Many languages have stress placement rules which are sensitive
to the specific structure (or weight) of a stressable syllable. The general case
is that a short vowel in open syllable position shows a reluctance to accept a stress,
and may therefore pass the stress onto a neighboring syllable. The result is that
syllables of the shape V or CV (which are said to be light) sometimes function
differently from syllables of the shape CVC, CV:, VC, C:, etc. (which are said
to be heavy; see Newman 1972 for a general statement).

The best known example of syllable weight comes from Latin. As seen from the
following forms from Allen (1973:51),

(a) refé:cit
(b) refé:ctus
(c) ré:ficit

stress is assigned penultimately in (a) and (b), since the penultimate syllable is
heavy; stress is assigned antepenultimately in (c), since the penultimate syllable
is light. In many languages stress is normally assigned only to a heavy syllable
(see below for further discussion). It is tempting to seek an explanation in the
phonetic properties underlying the heavy-light distinction in syllables, and in fact,
Allen (1973) presents a sophisticated explanation involving Stetson’s (1951) model of
the syllable as a chest pulse. Heavy syllables are said to be "arrested", either
orally by a consonant or "thoracically" by vowel length (or a second vowel); light
syllables are not arrested. In a language such as Latin, penultimate stress can
only be assigned to a syllable which is physiologically arrested.

Even ignoring Ladefoged’s (1967) demonstration of the weakness of Stetson’s
contention, there are reasons for questioning in general a physiological approach to
the problem. Any explanation must first address itself to the generalization that
languages with heavy vs. light dichotomy always have a vowel length contrast, as
noted by Jakobson, Trubetzkoy and others. In the absence of a vowel length
contrast, we would simply have a contrast between open and closed syllables, as
Allen also notes. Since a CV syllable functions as lighter than a CVC syllable only
if it stands in contrast with a CV: syllable as well, we should not expect to find a
language which says: stress the penultimate syllable if it is closed, but the ante-
penultimate syllable if the penultimate syllable is open. In order to maintain a
physiological explanation of syllable weight, it would be necessary to show a difference in the production of a CV syllable in a language with vs. without a vowel length contrast.

I would like to suggest, however, that it is not any possible physiological difference which accounts for the heavy vs. light dichotomy. Rather, the syllable weight phenomenon owes its existence directly to the nature of stress itself. Besides changing pitch, a second frequent strategy in the phonetic realization of stress is duration. As stated by Fant (1957:43): "The tendency towards lengthening is the most obvious feature observed as a physiological correlate to stress". Thus, if a stress falls on a short vowel in open syllable position, e.g. on the first vowel of a CVCV sequence, one of two things can happen: 1) that vowel can lengthen; or 2) the intervocalic consonant can lengthen. While the first appears to be more common than the second, both are attested. Thus, in Terena, "if the stressed vowel is phonemically unitary (and phonetically short)... the following consonant is lengthened, and the syllable division appears to fall in the middle of it..." (Harden 1946:60). Thus, /apene/ 'it exists' is pronounced [ap.pe.ne], with the dots representing syllable divisions.

In Hyman (1975:207) I argued that stress may shy away from a light (CV) syllable in a language with a vowel length contrast, because the vowel of that syllable would tend to lengthen and therefore might merge with the corresponding long vowel. In a language without a vowel length contrast, the stressed short vowel would be free to lengthen, and therefore no syllable weight distinction between CV and CVC would be observed. I now believe this to be an oversimplification. One argument which makes this hypothesis somewhat unappealing, is that most languages which have a vowel length contrast exhibit vowel quality differences in the corresponding long and short vowels, e.g. [i, e, a; o, u] vs. [i, e, a, o, u]. Thus, if one of the latter vowels were to receive stress and become lengthened as a result, it would not necessarily merge with the corresponding long vowel.

A more satisfactory way of stating a principle would be to say that stress may be rebuffed by a vowel which has some reason for staying short, either articulatory or perceptual. Many vowel systems have built into them the "strategy" that certain vowels (e.g. lax vowels, reduced vowels) should be short or punctual. To stress one of these vowels in an open syllable would result in a strategy conflict. The vowel "wants" to stay short, but stress wants to make it long. In Chuvash (Krueger 1961:86), stress cannot be assigned to the "reduced vowel phonemes" /e/ and /a/. Thus, stress will be on the last nonreduced vowel of the word, or on the first vowel of the word, if the word consists entirely of reduced vowels. Part of the strategy of Chuvash speakers in the realization of /e/ and /a/ is that these vowels will be short. Since stress tends to lengthen (and otherwise highlight the vowel of a syllable), the resulting conflict is resolved in such a way that stress will only be assigned to one of these vowels when there is no other choice. That such strategies play a role in accentual phenomena will be seen more clearly in further examples below.
A seemingly unrelated phenomenon is found in Cayapa: "All heavy syllables (defined as CVC, CVV(C)--but not CV?) are stressed" (Lindskog and Brend 1962:39). While this situation does not represent a stress-accent, as defined earlier, the problem raised by unstressable CV? syllables is particularly relevant. Such syllables, which must be treated as "arrested" in Allen's framework, are considered to be light in Cayapa. The reason is that a syllable-final glottal stop has a considerable shortening effect on a preceding vowel. This fact is well-known to students of tone languages, though its relevance in stress languages is not fully appreciated. In fact, the development of a long falling tone in Mohawk (Postal 1969:296) may be directly due to this basic incompatibility of stress and glottal stop (as well as /h/, another glottal consonant). In Mohawk, Proto-Iroquoian *Vʔ and *Vh often turn up as V:, i.e. a falling tone over a long vowel. While tonogeneticists may propose an intermediate stage of creakiness and breathiness, respectively, I would like to propose that *ʔ and *h drop out so as to permit the preceding vowel to lengthen. In Cayapa, then, it would appear that in the conflict between stress and glottal stop, the glottal stop wins out (since stress cannot be assigned to a CVʔ syllable); in Mohawk, it is the stress (and its concomitant lengthening effect) which wins out.

Finally, let us note that if syllable weight depended on a strictly physiological distinction, we would expect a language to have the following stress-placement rule: 1) a vowel is "redundantly" lengthened before a voiced consonant; 2) stress is placed penultimately on a CVC syllable, or on a syllable whose vowel length results from the above lengthening rule (we assume no underlying long vowels); and 3), finally, if the penultimate syllable is light (i.e. CV), stress is assigned antepenultimately. Such a language would have the following hypothetical derivations:

(a) /papaba/ → papa:ba → papaːba
(b) /papapa/ → pápapa

First /a/ is lengthened before the voiced consonant /b/, and then stress is assigned: penultimately in (a) and antepenultimately in (b). Since there is still no vowel length contrast in this language, the approach I have outlined above would predict that such a language would not exist. The physiologically-based approach of Allen (1973), on the other hand, predicts that the phonetic difference in the penultimate syllables of the words in (a) and (b) could create a syllable weight distinction. Notice in another connection, that if this language were shown to exist, it would violate a generalization which I made in an earlier paper (Hyman 1973:154), which can be summarized as follows: While a tone (e.g. H tone) can be restricted in occurrence by an adjacent voiced vs. voiceless consonant, no stress language assigns stress on such a basis. The hypothetical stress phenomena described above could, however, be restated equivalently in the following way: 1) if the final syllable begins with a voiced consonant, assign stress penultimately; 2) if it begins with a voiceless consonant, assign stress antepenultimately. If a language were to have a lengthening rule operating before
voiced consonants, we would expect that lengthening due to stress could simply be superimposed on the segmentally-induced length difference. We can summarize by saying that stress will generally be assigned without reference to segmentals except in two potential situations: 1) there is a severe articulatory conflict (e.g. shortening before glottal stop); and 2) there is a perceptual conflict from within the phonological system (e.g. stress assignment to short/lax vowels which stand in contrast with long/tense vowels).

As a final example of where both articulatory and perceptual factors may be involved, consider the following situation in Southern Paiute: "...all words beginning with a syllable containing an organic short vowel, inherently unglottalized, are accented on the second syllable, unless the second syllable is final and therefore unvoiced, in which case the main stress is thrown back on the first syllable..." (Sapir 1930:39). Articulatorily, it would of course be difficult to manipulate the various parameters of stress on a voiceless vowel; perceptually, no matter what one did, the result would not be too audible. Thus, in Southern Paiute devoicing wins out and stress is sent back to the first syllable (which has a voiced vowel). In Northern Paiute, on the other hand, if stress falls on the last vowel, this vowel is exempt from the final devoicing rule and stress remains final (Nichols 1974:80).

In the foregoing we have seen that there may be reasons for segments to repel an imminent stress. This leads us to ask whether segments can attract stress. In Burushaski (Morgenstierne 1942:88), bisyllabic words which have a long vowel in initial position receive initial stress; bisyllabic words which have a long vowel in final position receive final stress. Since a bisyllabic word lacking a long vowel is characterized by a uniform stress throughout the two syllables, we can say that in Burushaski stress is attracted by length (which in turn detracts from the stress of neighboring short vowels). The same can be said about Southern Sierra Miwok (and several other American Indian languages), where stress falls on the first long vowel of a word. Since one of the first two syllables of a word will be long, stress will fall in either initial or second position (Broadbent 1964:16).

A more serious problem is raised by languages such as Ulithian: "In general, stress (which is non-phonemic) is not clearly recognisable, and it subphonemically accompanies a long vowel or a short vowel preceding a long consonant" (Sohn and Bender 1973:74). Or consider the following statement concerning stress in Amharic: "Une voyelle étymologiquement longue, et qui n'est pas abrégée par sa situation dans le mot, fait volontiers l'effet d'avoir une certaine force ou intensité; de même, l'effort articulatoire fait pour émettre une consonne geminée semble donner une intensité plus forte non seulement à cette consonne mais à une voyelle voisine..." (Cohen 1970:63). In many languages stress is reported to fall on all heavy syllables, or on all long vowels, etc. One wonders how much of such descriptions is physically present and whether such stress judgments are universal or language specific. It would seem that the intrinsic pitch, duration and intensity values for different segments and sequences of segments can become phonologized and be incorporated into stress-placement rules. 34
While there is undoubtedly some physical reality to statements such as those cited above for Ulithian and Amharic, it may be necessary to be skeptical in some cases. Sohn and Bender, for example, add the following remark to their above description of Ulithian stress: "However, a short vowel preceding a long consonant followed by a long vowel is not stressed". What this statement seems to indicate is that such judgments do not directly reflect what is present in the acoustic signal. Rather, a judgment is made as to the relative prominence of a segment or sequence in a given utterance. Thus, the much greater "intrinsic stress" of a long vowel following a geminate consonant may detract perceptually from the intrinsic stress of a short vowel preceding this geminate consonant.

However, the general tendency for heavy syllables or long and/or tense vowels to attract (culminative) stress cannot be denied. A number of languages have a stress-placement rule mimicking that of Maori: "Major stress will fall on the first syllable containing a geminate cluster [i.e. long vowel], if none, then on the first syllable with a non-identical vowel cluster, and if none, then on the first syllable" (Hohepa 1967:10). What seems to have occurred in Maori is that regular initial stress was disturbed by the competition of a subsequent long vowel or vowel sequence. Maori is thus best described in these terms: a stress-placement rule assigns initial stress, which is then attracted onto the first subsequent (i.e. the nearest) V: or VV (if present) by a second rule.

The mirror image of the Maori situation is found in Aguacatec Mayan (McArthur and McArthur 1956:76): "Within the word, stress occurs on the final long vowel or on the final short vowel if no long vowels are present". Since most Mayan languages have final stress (see Mayers 1966), we can assume that in addition to the historical rule of final stress-placement, a second rule has been added in Aguacatec which attracts the stress onto the last long vowel of a word, i.e. onto that long vowel which is closest in proximity to the original final stress.

What is interesting is that some stress languages go one step further and assign stress as in Huasteco (another Mayan language), where stress "...is located on the last long vowel of the word, or, if there are no long vowels in the word, on the first short vowel, regardless of the number of vowels in the word" (Larsen and E. Pike 1949:289). The important change which has taken place, conceptually, in the development of a Huasteco-type language from an Aguacatec-type language is that vowel-length, rather than being a secondary strategy captured by a patch-up rule, has come to be primary. Thus, the resulting stress-placement rule is accurately stated in such languages as one which scans the word for a long vowel (either from right to left, as in Huasteco, or from left to right), and if there is no long vowel, the stress falls where the scanning process has stopped.

What this suggests is that syllable weight may play an important role in the historical development of stress systems. In particular, it may offer a clue as to the
mechanism by which stress systems become diversified in related languages. The Altaic family, for instance, offers a great deal of diversity with regards to stress-placement. While the Turkic languages are all stress-final, the Mongolian languages tend to have initial stress with a secondary rule attracting stress onto the first long vowel. A representative of this situation is Khalka Mongolian (Poppe 1951:13). However, languages in the Tungus subgroup sometimes have initial, sometimes final stress. While I am not competent to comment in depth on the history of stress in Altaic (I will also leave Japanese and Korean out of this), I would like to present a possible scenario which could explain the passage from initial to final stress without external influence. I will illustrate the two subparts of the relevant stress rule of each stage by means of the hypothetical word structures CVCV:CV: and CVCV:CV.

(1) *CVCV:CV:
(2) CVCV:CV:
(3) CV.CV:CV:
(4) CVCV:CV:

In stage (1) it is observed that Proto-Altaic had word-initial stress in all cases (Menges 1968:74). The first change occurred when initial stress was attracted away onto the first long vowel of a word, i.e. onto that long vowel nearest to the original initial stress position, as in (2a). At this stage we hypothesize, as mentioned above, an initial stress rule, which is followed by a second rule moving stress onto the first long vowel of a word. In the event that there is no long vowel in a given word, stress will remain in initial position, as in (2b). In stage (3), on the other hand, although stress remains on the first long vowel in (3a), final stress is observed on words with no long vowel (3b). Here we hypothesize that stress is assigned by a rule which says to stress the first long vowel of a word (i.e. the initial stress rule has been lost), and as a natural consequence, if the whole word has been scanned and no long vowel found, the last vowel of the word will be stressed. Unfortunately I have not had access to Russian materials describing Tungus languages, but if such a language exists in Altaic (and it does exist elsewhere), it is in Tungus that we expect to find it. In the final stage, the stress of words with one or more long vowels shifts to final position (4a), generalizing the final stress of words with no long vowel (3b, 4b). In other words, the long-vowel strategy, which was primary in stage (3), has been dropped, and final stress has been generalized. Such a situation is found in the Turkic languages, all of which have final stress. It is quite clear that final stress is an areal feature encompassing the Turkic languages, the Iranian languages and Armenian. One might be tempted, therefore, to suppose that contact is responsible for the shift from initial to final stress in Turkic. However, one piece of evidence seems to support the view I have advanced above. This is that the Proto-Altaic vowel length contrast has been lost in all Turkic languages (Räsymänen 1949:64). It can thus be argued that as vowel length was being lost, Turkish
began treating all words as having only short vowels—and thus, *final stress* was generalized in all cases.37

To summarize, the development of stress in such languages clearly shows that stress can be attracted off of a short vowel onto a nearby long vowel. So great is this tendency that it is not clear which is more significant: the tendency for short vowels to rebuff stress or the tendency for long vowels to attract stress. This question is particularly important in view of the fact that Latin, a language on the basis of which many theories have been constructed, turns out to be exceptional when compared to other languages with syllable weight phenomena. What makes Latin exceptional is that a light antepenultimate syllable can attract a stress from a light penultimate syllable (e.g. réficit). This is a strange set of affairs both because one light syllable should not be favored over another, and also because, other things being equal, penultimate position is universally favored for stress over antepenultimate position. In most accounts of Latin, stress is assigned penultimately unless the penultimate syllable is light. Thus, Latin is seen to have dominant penultimate stress. However, we are faced not only with the problem of why stress is antepenultimate in réficit, but also why it is not pre-antepenultimate in conspicio, where there is a heavy syllable for stress to be attracted to. Unlike the other languages we have just looked at, the search for a heavy syllable cannot go back further than antepenultimate position. In fact, as we have just seen, it does not even seem appropriate to speak of such a search, because of words such as réficit.

The problem, I believe, is tied up with the general assumption that Latin is a dominant penultimate stress language. All of the difficulties encountered above can be eliminated if we simply assume that Latin has dominant antepenultimate stress, but that a heavy syllable in penultimate position will attract the stress onto it. This suggestion ties in nicely with the reconstruction of initial stress in Proto-Italic (Kent 1932:65), and allows us to say that stress was never penultimate in such words as réficit. The alternative, namely that stress was or is underlyingly penultimate in such words but is retracted to antepenultimate position, is totally unmotivated from either a conceptual or a phonetic point of view. Because such words as conspicio receive antepenultimate stress, it would appear that by the time of Classical Latin, stress had changed from initial to antepenultimate position; cf. Kent (1932:66): "The change from the one system to the other was almost complete in the time of Plautus, but not quite, for quadrasyllables with the first three syllables short, normally retain the initial accent in that poet's work: fácallus, séquímini, cécidero, múlherem; later facíllus, etc.". As to why a language would change from initial to antepenultimate stress, we can only conjecture that in words of four or more syllables, there might be a tendency to avoid too many unstressed syllables in a row (especially post-tonic). Thus, the stress in fácallus reduces the number of successive unstressed syllables from three to two.38
The association between stress and greater duration has thus been demonstrated in a number of cases. This association may also account for the process of "iambic shortening", which is discussed at length by Allen (1973:179-185). Allen cites the following statement by Abercrombie (1964:218): "There is felt to be something anomalous in a syllable which is stressed and yet short, followed by an unstressed one which is long...". When such sequences are found, one of two things can be observed to occur: 1) the stress can shift from the short vowel onto the long vowel, in which case stress and duration coincide, as in several languages cited earlier; or 2) the length of the vowels in question may be affected. In the history of Latin, as seen in the following forms (Allen 1973:179),

(a) \( \ddot{e}go: > \dddot{e}go \)
(b) \( \ddot{c}ito: > \dddot{c}ito \)
(c) \( \dddot{m}\ddot{o}do: > \dddot{m}\ddot{odo} \)

when in bisyllabic forms a short stressed vowel is followed by a long unstressed vowel, the latter becomes short. In Akkadian, as seen in the following forms (Moscati 1964:66),

(a) \( \dddot{imq}\dddot{\ddot{u}}tu: > \dddot{imq}\dddot{\ddot{u}}tu: \) 'they fell'
(b) \( \dddot{i}\ddot{\ddot{\ddot{s}}k}\dddot{\dddot{\ddot{k}}}\dddot{\dddot{u}}nu: > \dddot{i}\dddot{\ddot{s}}k\dddot{\dddot{\ddk}}\dddot{\dddot{\ddot{k}}}\dddot{\dddot{u}}nu: \) 'they put'

either the short vowel may be lengthened, as in (a), or the following consonant may be geminated, as in (b), so as to produce a stressed syllable which is heavy.

In Latin, notice that because stress is fixed, a conflict between stress and syllable weight will arise only under two circumstances: 1) as in the above examples, a bisyllabic word consists of a light penultimate and a heavy final syllable; and 2) a word consisting of four or more syllables has a light penultimate as well as a light antepenultimate syllable, the latter receiving stress, but there is an earlier heavy syllable, e.g. \( \dddot{m}ob\dddot{\ddot{\ddot{l}}}\dddot{\dddot{\ddk}}\dddot{\dddot{\ddot{i}}}\dddot{\dddot{\dddot{\dddot{a}}}t\dddot{\dddot{\dddot{\dddot{a}}}s}} \). The conflict created by the tendency of a long vowel to attract stress is resolved by shortening a long vowel which is adjacent to a stressed short vowel, as seen above. However, in a language where stress is free, i.e. where it can appear on different syllables in different words, the range of such conflicts is much greater. This, I assume is responsible for the observation that "languages where both length and stress appear as distinctive features are quite exceptional..." (Jakobson and Halle 1956:481). Such situations are unstable, I would claim, because the presence of a vowel length contrast obstructs the use of duration as a strategy in the realization of stress.
4. Conclusion

In the preceding sections we have seen the need to distinguish between conceptual and phonetic aspects of stress-accent, and have been able to explain a number of generalizations concerning stress placement and syllable weight by means of this distinction. It should be clear, however, that this represents only a first attempt to characterize some of the universal properties of stress, and that much of what has been said must be taken as tentative. In addition, several other important aspects of stress remain to be studied in further detail. Thus, it will be necessary to do cross-linguistic surveys of exceptions to stress rules and to look in particular at the ways in which the stem vs. affix distinction is relevant in determining stress placement. Also, the role of stress in derivational morphology, especially compounding, deserves a treatment of its own.40 In addition, while the model I have proposed for "accentogenesis" involving intonation predicts that utterance prosodies will be generalized to clause, phrase and word boundaries (and perhaps ultimately to stem boundaries, although the stem/affix distinction may involve an independent motivation for accentogenesis), the model will have to be extended to cover such languages as French, where the direction is the reverse (i.e. word stress in Latin has become phrase stress in French). The interaction of stress and syllabification seems appropriate for further investigation. If stress does in fact have a demarcative function in languages, i.e. if it actively serves as a boundary marker for the hearer, we should expect languages with regular (fixed) stress to be less likely to allow syllabification across a word boundary than languages with lexical (free) stress. In a language such as French, which has final phrase stress and which allows syllabification across a word boundary, e.g. fils unique 'only son' is syllabified [fl.sil.nik], one is tempted to ask whether one follows from the other. Did syllabification across a word boundary lead to phrase stress, or did syllabification across a word boundary result from the change from word to phrase stress? For the moment, I would like to simply point out that there may be generalizations to be discovered in the interaction between stress-placement and syllabification.

A rapid check in a few familiar languages suggests that syllabification across a word boundary may be less likely in a language with initial stress than in a language with final stress. Presumably this is because of the vowel-lengthening effect of a final stress which in turn favors syllabification of the type V(:)C#V, rather than V(:)C#V. Whether this holds up when checked against a larger sample of languages remains to be seen.
APPENDIX: Summary of Stress-Placement in the World's Languages

The present study has considered 444 languages, of which 300 show a dominant initial, second-syllable, penultimate or final stress placement. Six languages have been found with dominant antepenultimate stress, nine with syllable weight stress and 16 with no stress-accent and no tone. Finally, 138 languages have been found for which no statement can be made as to any dominant tendency in stress-placement. (Most of these are described as having "phonemic" stress in the literature.) These findings are summarized in the following tables. Languages which are known to me only through Merritt Ruhlen's Guide to the Languages of the World (1975) are marked by (R).

In summarizing stress-placement on a cross-linguistic basis, two immediate problems have had to be faced: 1) how do we decide what is "dominant"? and 2) how do we weight different languages from different parts of the world and from different language families? Neither of these can be provided a simple answer. In the first case, it has been necessary to consider a number of different criteria in determining whether a language has a dominant stress assignment tendency. One obvious one is sheer frequency. But then it is not clear where to draw the line: a language in which 90% of the words have penultimate stress is clearly a dominant penultimate stress language. But what if only 60% have penultimate stress? In most cases, however, we do not have such percentages to deal with, but rather a simple statement such as "most words in language X have penultimate stress". These statements are taken at face value, and language X is considered to have dominant penultimate stress.

This is, however, not the only problem associated with the determination of dominant stress patterns. Another concerns the grammatical unit within which a given phonological unit (normally the syllable) is stressed. In giving the distribution of the various stress-placements in different languages, I have not indicated whether it is the first syllable of a word, or of a phrase, or of a stem which receives stress. I have simply attempted to indicate whether a language counts from the beginning or from the end of a given grammatical unit, and how far it counts. I hope that this does not create any confusion. It may, of course, hide certain generalizations which cannot be studied from the following tables. For example, it seems to me that few languages have stem-penultimate stress, as opposed to word-penultimate stress, although stem-final stress is quite common. Whether this impression is correct will have to await further examination of the data. A potential problem which may arise should be pointed out, however. In some cases it may not be possible to determine whether a language has stem-final or word-penultimate stress, if there is exactly one suffix position in a word. Thus, there are two ways of describing the stress in a word such as gali+ma and the criteria for choosing between them may be unclear—at least potentially.

Just as I have collapsed the different grammatical units, so have I ignored the difference between syllable and mora. In most of the languages studied the relevant
unit is the syllable. The generality of the "mora" requires further attention, but it should be pointed out that I may have to revise my thinking so as to give explicit account of this concept. This problem is important because it affects the status of a language such as Delaware, which is said to have stress on the third mora from the end (Voegelin 1946:137).

The second problem mentioned above concerns the relative weight which should be assigned to different languages or language families. Certain areas are clearly underrepresented, while others appear to be over-represented (e.g. Turkic with its 21 final-stress languages). Even if we knew how to weight the different language groups it would of course be impossible to find the right kinds of languages to do so. One question which is difficult to resolve, for instance, concerns the problem of determining how many languages from a related genetic family should be counted, especially if they all have the same kind of stress-placement, as in the case of Turkic. If an area or group is too heavily represented, this may disturb the relative values of the different stress-placement possibilities. It is of course not so easy to say that Turkic should get only one vote as opposed to 21 (or some reduced number). A comparable group in which each language has a different stress-placement would presumably get one vote per language, but this would discount the possibility that conservativeness may be as much a clue in distinguishing the relative naturalness values of different stress placements as is change. Thus, I have simply given every language I have found, and have indicated what is judged to be its dominant stress-placement pattern.

A final word should be said about so-called "syllable-weight" stress languages. These are those languages which say, for example: stress the last long vowel of a word; in the absence of a long vowel in a word, the first syllable is stressed. In this language, as was argued in the text, syllable-weight constitutes the primary strategy for stress-placement. If the same language had assigned final stress in words lacking a long vowel, it would be considered to have a primary final-stress strategy with a secondary patch-up rule detracting the stress from the ultima onto the nearest long vowel of the word. Languages of this type are classified as initial, final, etc. and are marked "syl.wt." in the following tables.

**ABBREVIATIONS**

(R) : Ruhlen (1975)

(L) : Limited stress placement; e.g. on only the first or second syllable, or on only the final or penultima, etc.

(syl.wt.) : Syllable weight plays a role in stress-placement

(V-quality) : Vowel quality plays a role in stress-placement
<table>
<thead>
<tr>
<th>AREA or family</th>
<th>+1</th>
<th>+2</th>
<th>-3</th>
<th>-2</th>
<th>-1</th>
<th>Other</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFRICA</td>
<td>7</td>
<td>-</td>
<td>1</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>19</td>
</tr>
<tr>
<td>Afro-Asiatic</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>5</td>
<td>1</td>
<td>10</td>
<td>17</td>
</tr>
<tr>
<td>Altaic</td>
<td>6</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>26</td>
<td>2</td>
<td>35</td>
</tr>
<tr>
<td>AUSTRALIA</td>
<td>22</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>1</td>
<td>-</td>
<td>27</td>
</tr>
<tr>
<td>Austro-Asiatic</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>3</td>
<td>7</td>
</tr>
<tr>
<td>Austronesian</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>18</td>
<td>8</td>
<td>14</td>
<td>48</td>
</tr>
<tr>
<td>Caucasian</td>
<td>5</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>2</td>
<td>14</td>
<td>23</td>
</tr>
<tr>
<td>Dravidian</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>Indo-European</td>
<td>15</td>
<td>-</td>
<td>1</td>
<td>8</td>
<td>12</td>
<td>8</td>
<td>44</td>
</tr>
<tr>
<td>Mayan</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>13</td>
<td>5</td>
<td>18</td>
</tr>
<tr>
<td>NEW GUINEA</td>
<td>4</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>20</td>
</tr>
<tr>
<td>NORTH AMERICA</td>
<td>19</td>
<td>5</td>
<td>2</td>
<td>17</td>
<td>14</td>
<td>45</td>
<td>102</td>
</tr>
<tr>
<td>Paleo-Siberian</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>-</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>SOUTH AMERICA</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>16</td>
<td>7</td>
<td>14</td>
<td>46</td>
</tr>
<tr>
<td>Uralic</td>
<td>17</td>
<td>1</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>OTHER</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td><strong>TOTAL:</strong></td>
<td>114</td>
<td>12</td>
<td>6</td>
<td>77</td>
<td>97</td>
<td>138</td>
<td>444</td>
</tr>
<tr>
<td>Languages with Dominant Initial Stress</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AFRICA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Balanta (Ganja) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diola-Fogny</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fulani</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mba (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sango</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolof</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zande (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Altaic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Altaic (proto)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buriat (syl.wt.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dagur (syl.wt.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Khaikha Mongolian (syl.wt.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olcha (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Udihe (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>AUSTRALIA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arabana Wangaguru</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bidyara &amp; Gungabula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Burera</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dieri (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Djadjala (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dyirbal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Juat (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kalkutung (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kardutjara (Gugada) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kardutjara (Pitjantjatjara) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitabal (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kitja (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kokoj-Jelandji (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mantjiltjara</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maranungku</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nunggubuyu (syl.wt.) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nyulnyul (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pintupi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pitta-Pitta</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walmatjari</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Western Desert Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wik-Munkan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Austronesian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dehu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iai</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lifu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Austronesian (cont.)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maori (syl.wt.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nenema (Koumac) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tigak (central) (R)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Caucasian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Archi</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bats</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chechen</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Georgian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xinalug</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Dravidian</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Badaga</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brahui</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kolami</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kota</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Koya</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Telugu</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Toda</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Indo-European</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Irish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Danish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dutch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>German</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Icelandic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Swedish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yiddish</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bengali</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gujarati (V-quality)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kashmir</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ossetic (syl.wt.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Italic (proto)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Czech</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Latvian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serbo-Croatian</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>NEW GUINEA (non-Austronesian)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guhu-Samane</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kewa (tone)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
NEW GUINEA (cont.)
Orokaiva (R)
Selepet

NORTH AMERICA
Cahuilla
Chinook (Wishram)
Comanche
Chitimacha?
Cuicateco (tone)
Haida
Kiliwa? (tone)
Konkow
Luiseno (syl.wt.)
Miwok (Lake) (syl.wt.)
Miwok (Western) (syl.wt.)
Mixe (Totontepec)
Papago
Pomo (Southeastern)
Shoshone
Timucua
Tunica
Yana (syl.wt.)
Zuni

URALIC
Estonian
Finnish
Hungarian
Ingrian (R)
Karelian (R)
Komi? (R)
Lapp (R)
Livonian (R)
Mordvin (R)
Ostyak
Selkup (R)
Tavgy (R)
Uralic (proto)
Vepsian (R)
Vogul (general)
Votic
Yenets (R)

SEMITIC
Akkadian (Assyrian dial.)

SOUTH AMERICA
Arabela?
Aykaumeli
Cayapa
Parintintin
Siona

TOTAL: 114
LANGUAGES WITH DOMINANT SECOND SYLLABLE STRESS

Altaic
Mongolian? (Modern Standard)

Austronesian
Tolai

Caucasian
Tsaxur (R)

SOUTH AMERICA
Araucanian
Baure
Ignaciano

Uralic
Vogul? (Tavda dialect)

NORTH AMERICA
Dakota
No. Paiute
Pomo (Central)
Seneca (V-quality)
So. Paiute

TOTAL: 12
# Languages with Dominant Penultimate Stress

## Africa

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chichewa (tone)</td>
</tr>
<tr>
<td>Mangbetu</td>
</tr>
<tr>
<td>Shona (tone)</td>
</tr>
<tr>
<td>Swahili</td>
</tr>
</tbody>
</table>

## Australia

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alawa</td>
</tr>
<tr>
<td>Anyula</td>
</tr>
<tr>
<td>Ngarndji</td>
</tr>
<tr>
<td>Wailbri</td>
</tr>
</tbody>
</table>

## Austronesian

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Southeast) Ambrym</td>
</tr>
<tr>
<td>Chamorro (R)</td>
</tr>
<tr>
<td>Futuna-Aniwa</td>
</tr>
<tr>
<td>Hawaiian</td>
</tr>
<tr>
<td>Iraya Mangyan</td>
</tr>
<tr>
<td>Kaliai (R)</td>
</tr>
<tr>
<td>Luanguia (R)</td>
</tr>
<tr>
<td>Malay</td>
</tr>
<tr>
<td>Nengone</td>
</tr>
<tr>
<td>Nggela</td>
</tr>
<tr>
<td>Niue (R)</td>
</tr>
<tr>
<td>Sabah Murut</td>
</tr>
<tr>
<td>Samoan</td>
</tr>
<tr>
<td>Sarangani Manobo</td>
</tr>
<tr>
<td>Sedik</td>
</tr>
<tr>
<td>Tsou</td>
</tr>
<tr>
<td>Tuamotuan</td>
</tr>
<tr>
<td>Wolio</td>
</tr>
</tbody>
</table>

## Caucasian

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rutul</td>
</tr>
</tbody>
</table>

## Indo-European

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albanian?</td>
</tr>
<tr>
<td>Bhojpuri</td>
</tr>
<tr>
<td>Breton</td>
</tr>
<tr>
<td>Polish</td>
</tr>
<tr>
<td>Portuguese</td>
</tr>
<tr>
<td>Roumanian</td>
</tr>
<tr>
<td>Spanish</td>
</tr>
<tr>
<td>Welsh</td>
</tr>
</tbody>
</table>

## New Guinea (non-Austronesian)

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentani?</td>
</tr>
<tr>
<td>Suena (tone)</td>
</tr>
</tbody>
</table>

## North America

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aleut</td>
</tr>
<tr>
<td>Chumash</td>
</tr>
<tr>
<td>Huichol</td>
</tr>
<tr>
<td>Hupa</td>
</tr>
<tr>
<td>Klamath?</td>
</tr>
<tr>
<td>Kutenai</td>
</tr>
<tr>
<td>Menomini?</td>
</tr>
<tr>
<td>Mohawk</td>
</tr>
<tr>
<td>Nahuat (Isthmus)</td>
</tr>
<tr>
<td>Nahuatl1 (North Pueblo)</td>
</tr>
<tr>
<td>Nahuat1 (Sierra)</td>
</tr>
<tr>
<td>Nahuatl1 (Telelcingo)</td>
</tr>
<tr>
<td>Nisenan</td>
</tr>
<tr>
<td>Pomo (Southern)</td>
</tr>
<tr>
<td>Yokuts</td>
</tr>
<tr>
<td>Zapotec (Isthmus) (tone)</td>
</tr>
<tr>
<td>Zoque</td>
</tr>
</tbody>
</table>

## Paleo-Siberian

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kamchadal (Napan)</td>
</tr>
</tbody>
</table>

## Semitic

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Maltese) Arabic? (syl.wt.)</td>
</tr>
<tr>
<td>Arabic (Tunisian)? (syl.wt.)</td>
</tr>
<tr>
<td>Arabic (Egyptian/Syrian)? (syl.wt.)</td>
</tr>
<tr>
<td>Aramaic (East)</td>
</tr>
<tr>
<td>Syriac (Modern Spoken)</td>
</tr>
</tbody>
</table>

## South America

<table>
<thead>
<tr>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amuesha</td>
</tr>
<tr>
<td>Aymara.</td>
</tr>
<tr>
<td>Ayoweos-Moros</td>
</tr>
<tr>
<td>Bacairi (R)</td>
</tr>
<tr>
<td>Campa</td>
</tr>
<tr>
<td>Camsa (R)</td>
</tr>
<tr>
<td>Cavinena (R)</td>
</tr>
<tr>
<td>Cofan</td>
</tr>
<tr>
<td>Itonama</td>
</tr>
<tr>
<td>Motilone? (syl.wt.)</td>
</tr>
<tr>
<td>Piro</td>
</tr>
</tbody>
</table>
SOUTH AMERICA (cont.)

Quechua
Siriono
Tacana (R)
Icua Tupi
Warao

TOTAL: 77
### Languages with Dominant Final Stress

#### Africa
- Malinka (R)
- Manding

#### Altaic
- Even (syl.wt.) (R)
- Monguor
- Nanaj (R)
- Oroch (R)
- Udihe (R)

#### Turkic:
- Altai (R)
- Azerbaijani (R)
- Bashkir
- Chulym (R)
- Chuvash (V-quality)
- Gagauz (R)
- Karachay (R)
- Karaim (Trakaj) (R)
- Karakalpak (R)
- Kazakh (R)
- Khakas (R)
- Kirghiz (R)
- Nogai (R)
- Orkhon Turkic
- Shor (R)
- Tatar
- Turkish
- Tuvan (R)
- Uighur (R)
- Uzbek
- Yakut

#### Austro-Asiatic
- Cambodian (R)
- Gutob-Remo
- Khasi (R)
- Mundari (R)

#### Austronesian
- Agta
- Atayal (R)
- Cham (R)

---

#### Indo-European
- Non-Iranian:
  - Armenian
  - French

#### Caucasian
- Udi
- Zan (Mingrelian)

#### Indo-European
- Non-Iranian:
  - Armenian
  - French

#### Austronesian (cont.)
- Ivatan
- Kalagan
- Sa?ban
- Saisiyat (R)
- Yogad (R)

#### Indo-European
- Non-Iranian:
  - Armenian
  - French

#### Iranian:
- Baluchi (Southwest) (R)
- Kurdish (Salaimania) (R)
- Persian
- Sanglechi (R)
- Shugni (R)
- Tajik
- Talysh (R)
- Wakhi (R)
- Yazghulami (R)

#### Mayan
- Achi (R)
- Aguacatec Mayan (syl.wt.)
- Cakchiquel (R)
- Chorti (R)
- Chuj (R)
- Kekchi (R)
- Mam (R)
- Maya (Mopan) (R)
- Pocomchi
- Quiche
- Tojolabal
- Tzeltal
- Tzotzil (Huixtec) (R)
NEW GUINEA (non-Austronesian)

Asmat
Dani (Lower Grand Valley)
Morwap (R)
Wahgi (Middle)
Weri (Sim) (R)

NORTH AMERICA

Alabaman
Chatino (R)
Chinantec (Tepetotutla) (tone)
Diegueño
Kiliwa (R)
Maku
Mohave (R)
Onondaga?
Popoluca (Western) (tone)
Shawnee?
Tillamook
Trique (tone)
Tóbatulabal
Yuma

AUSTRALIA

Kunjen

SOUTH AMERICA

Apinaye
Chavante (R)
Guananq (R)
Guarani
Shiriana?
Tenetehara (Guajajara) (R)
Yaruro (R)

Uralic

Votyak

ISOLATE

Ainu

Semitic

Hebrew (Modern)

TOTAL: 97
### LANGUAGES WITH DOMINANT ANTENPELULTIMATE STRESS

<table>
<thead>
<tr>
<th>Language</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cayuvava (S.America)</td>
<td>Latin? (see text) (Indo-European)</td>
</tr>
<tr>
<td>Cora (N.America)</td>
<td>Mae (Austronesian)</td>
</tr>
<tr>
<td>Delaware (moras) (N.America)</td>
<td>Temein (Nido-Saharan) (R)</td>
</tr>
</tbody>
</table>

**TOTAL:** 6

### LANGUAGES WITH DOMINANT SYLLABLE WEIGHT STRESS

<table>
<thead>
<tr>
<th>Language</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Akkadian (Semitic)</td>
<td>Duwamish (N.America)</td>
</tr>
<tr>
<td>Arabic (Classical and other)</td>
<td>Hebrew (Proto) (Semitic)</td>
</tr>
<tr>
<td>Arabic (Iraqian) (Semitic)</td>
<td>Huasteco (Mayan)</td>
</tr>
<tr>
<td>Carib (S. America)</td>
<td>Miwok (Southern Sierra) (N.America)</td>
</tr>
<tr>
<td>Cheremis (Eastern) (Uralic)</td>
<td></td>
</tr>
</tbody>
</table>

* (see also languages marked "syl.wt.")

**TOTAL:** 9

### LANGUAGES WITH NO STRESS-ACCENT OR TONE

In the following languages there is either no stress as distinct from intonation, or there is stress which is predictable from segmental information (e.g. on each long vowel or on each vowel followed by a geminate or consonant cluster, etc.).

<table>
<thead>
<tr>
<th>Language</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bella Coola (N.America)</td>
<td></td>
</tr>
<tr>
<td>Berber (Tamazight) (Afro-Asiatic)</td>
<td></td>
</tr>
<tr>
<td>Chitimacha (N.America)</td>
<td></td>
</tr>
<tr>
<td>Cuna (Panama)</td>
<td></td>
</tr>
<tr>
<td>Gafat (Afro-Asiatic)</td>
<td></td>
</tr>
<tr>
<td>Goajiro (Arawak) (S.America)</td>
<td></td>
</tr>
<tr>
<td>Hindi (Indo-European)</td>
<td></td>
</tr>
<tr>
<td>Javanese? (Austronesian)</td>
<td></td>
</tr>
<tr>
<td>Kato (N.America)</td>
<td></td>
</tr>
<tr>
<td>Korean (Altaic)</td>
<td></td>
</tr>
<tr>
<td>Mundari (Austro-Asiatic)</td>
<td>(Some of these are found</td>
</tr>
<tr>
<td>Puluwat (Austronesian)</td>
<td>in other columns as well,</td>
</tr>
<tr>
<td>Shilha (Afro-Asiatic)</td>
<td>because of uncertainties.)</td>
</tr>
<tr>
<td>Tsotsil (Mayan)?</td>
<td></td>
</tr>
<tr>
<td>Ulithian? (Austronesian)</td>
<td></td>
</tr>
<tr>
<td>Yurok? (N.America)</td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:** 16
LANGUAGES WITH NO DOMINANT STRESS PLACEMENT

In the following cases either I or Ruhlen (R) were not able to ascertain any dominant stress placement. Most of these languages were reported as having "phonemic" stress. Any information leading to a clarification of the stress status of these languages would be greatly appreciated. (L=limited stress placement)

Afro-Asiatic

Beja (R)
Hebrew (Biblical)
Saho

Altaic

Evenki (R)

Austronesian

Batak (Toba) (R)
Buhid (Mangyan)?
Cebuano (L)
Hanunóo (L)
Ibanag (R)
Kalinga
Malagasy (Merina) (R)
Palauan (R)
Pampangan (Bacolar) (R)
Tagalog (L)
Tiruray (L)

Caucasian

Abkhazian (R)
Avar (R)
Bagvali (Tlondoda) (R)
Botlix (R)
Botlix (Godoberi) (R)
Chamalal (R)
Dido (R)
Ginux (R)
Lak (Kumux) (R)
Svan (Upper Bal) (R)
Tabasaran (R)
Tindi (R)
Ubyx (R)
Xvarshi (Inxokari) (R)

Chinese

Mandarin (tone)

Indo-European

Bielorussian (stem?)
English
Friulian
Greek (Modern)
Pashto
Russian
Ukrainian

Mon-Khmer

Sedang (stem?)
Sre (stem)

NEW GUINEA

Iwam (R)
Kamano
Kaugel (R)
Keigana (Yagaria)
Manambu (Yambon) (R)
Nimboran
Rawa (R)
Waffa (Kusing) (R)
Yareba (Bibira) (R)

Nilo-Saharan

Barya (R)
Berta (R)
Ingassana (R)
Kunama (R)
Nyimang (R)
### NORTH AMERICA

<table>
<thead>
<tr>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atsugewi</td>
<td></td>
</tr>
<tr>
<td>Blackfoot</td>
<td></td>
</tr>
<tr>
<td>Chinook Jargon?</td>
<td></td>
</tr>
<tr>
<td>Chol? (Mayan)</td>
<td></td>
</tr>
<tr>
<td>Chontal (Highland)</td>
<td></td>
</tr>
<tr>
<td>Chontal (Lowland)</td>
<td></td>
</tr>
<tr>
<td>Chontal (Mayan)</td>
<td></td>
</tr>
<tr>
<td>Clallam?</td>
<td></td>
</tr>
<tr>
<td>Cocopa</td>
<td></td>
</tr>
<tr>
<td>Cupeno?</td>
<td></td>
</tr>
<tr>
<td>Dhegiha (Osage)</td>
<td></td>
</tr>
<tr>
<td>Hopi</td>
<td></td>
</tr>
<tr>
<td>Ioway-Oto (L)</td>
<td></td>
</tr>
<tr>
<td>Jacaltec (Jacaltenango) (R)</td>
<td></td>
</tr>
<tr>
<td>Keresan (Mayan)</td>
<td></td>
</tr>
<tr>
<td>Kiowa</td>
<td></td>
</tr>
<tr>
<td>Kitsai</td>
<td></td>
</tr>
<tr>
<td>Maidu</td>
<td></td>
</tr>
<tr>
<td>Mandan</td>
<td></td>
</tr>
<tr>
<td>Miwok (Northern Sierra)</td>
<td></td>
</tr>
<tr>
<td>Mixe (Tlahuitoltepec) (R)</td>
<td></td>
</tr>
<tr>
<td>Molale (R)</td>
<td></td>
</tr>
<tr>
<td>Nez Perce (Upper) (R)</td>
<td></td>
</tr>
<tr>
<td>Okanagan (R)</td>
<td></td>
</tr>
<tr>
<td>Popoluca (Sayula)</td>
<td></td>
</tr>
<tr>
<td>Popoluca (Sierra)</td>
<td></td>
</tr>
<tr>
<td>Sahaptin (R)</td>
<td></td>
</tr>
<tr>
<td>Salinan</td>
<td></td>
</tr>
<tr>
<td>Seri</td>
<td></td>
</tr>
<tr>
<td>Squamish</td>
<td></td>
</tr>
<tr>
<td>Tahlita (stem)</td>
<td></td>
</tr>
<tr>
<td>Tarascan (L)</td>
<td></td>
</tr>
<tr>
<td>Twana (R)</td>
<td></td>
</tr>
<tr>
<td>Upper Chehalis</td>
<td></td>
</tr>
<tr>
<td>Walapai</td>
<td></td>
</tr>
<tr>
<td>Wappo (stem)</td>
<td></td>
</tr>
<tr>
<td>Washo</td>
<td></td>
</tr>
<tr>
<td>Winnebago</td>
<td></td>
</tr>
<tr>
<td>Wiyot</td>
<td></td>
</tr>
<tr>
<td>Yaqui</td>
<td></td>
</tr>
<tr>
<td>Yupik (Kuskokwim) (R)</td>
<td></td>
</tr>
<tr>
<td>Yurok?</td>
<td></td>
</tr>
</tbody>
</table>

### SOUTH AMERICA

<table>
<thead>
<tr>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cashibo (R)</td>
<td></td>
</tr>
<tr>
<td>Chacobo (R)</td>
<td></td>
</tr>
<tr>
<td>Chama</td>
<td></td>
</tr>
<tr>
<td>Guahibo (R)</td>
<td></td>
</tr>
<tr>
<td>Island Carib</td>
<td></td>
</tr>
<tr>
<td>Jivaró (Aguaruna) (R)</td>
<td></td>
</tr>
<tr>
<td>Jivaró (Huambisa) (R)</td>
<td></td>
</tr>
<tr>
<td>Paez</td>
<td></td>
</tr>
<tr>
<td>Waiwai (R)</td>
<td></td>
</tr>
<tr>
<td>Waunana (R)</td>
<td></td>
</tr>
<tr>
<td>Yucuna (R)</td>
<td></td>
</tr>
</tbody>
</table>

### Uralic

<table>
<thead>
<tr>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ostyak (Eastern) (R)</td>
<td></td>
</tr>
<tr>
<td>Yurak (Tundra)</td>
<td></td>
</tr>
</tbody>
</table>

### ISOLATES

<table>
<thead>
<tr>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basque</td>
<td></td>
</tr>
<tr>
<td>Burushaski (syl.wt.)</td>
<td></td>
</tr>
</tbody>
</table>

### Paleosiberian

<table>
<thead>
<tr>
<th>Language</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilyak (Amur) (R)</td>
<td></td>
</tr>
<tr>
<td>Ket (Imbat) (R)</td>
<td></td>
</tr>
<tr>
<td>Koryak (R)</td>
<td></td>
</tr>
</tbody>
</table>

**TOTAL:** 113
FOOTNOTES

1 The research upon which this paper is based was supported in part by a Postdoctoral Fellowship from the Miller Institute for Basic Research in Science, University of California, Berkeley. I am particularly grateful to John Ohala, Director of the Phonology Laboratory at Berkeley, who was instrumental in my obtaining this support, and to Joseph Greenberg, who made available to me his extensive notes on prosodic systems. In addition, I would like to thank Wallace Chafe and Karl Zimmer for discussions on stress in Iroquoian and Turkic languages, respectively. Finally, my thanks to those people who commented on earlier versions of this paper, especially Josh Ard, Sandra Thompson, Paul Newman, and Bernard Comrie.

2 Ironic as it may seem, when the word "tone" is substituted for the word "stress", questions (i) - (v) become much more answerable. With all of the recent interest, it can reasonably be argued that tone is now better understood than stress.

3 Because of languages such as Japanese (McCawley 1968) and Otomi (Leon and Swadesh 1949), where an underlying single accentual mark is translated into phonetic H and L tone syllables or moras, it is important to emphasize that the notion of a culminative feature may yield different taxonomies depending on whether one looks at the underlying or the phonetic level. The term pitch-accent often refers to languages which have phonologically culminative accent, but phonetically nonculminative tone.

4 Except where stated otherwise, the term "word" will stand for the phrase "word or other accentual unit".

5 One would like to have more information about such languages as Campa, both on the phonetic parameters of such multiple stresses and on their history and phonological nature. K. Pike and Kindberg point out that there is considerable variation in how these multiple stresses will be realized phonetically, which variation is singled out by E. Pike (1974:174) as typical of multiple stress languages. One wonders from the available descriptions whether Campa might be amenable, for instance, to a re-analysis involving echo stresses and perhaps different internal boundaries within a word. Thus, it is possible that the five possible realizations of a quadrisyllabic word, i.e. oxoo, xoxx, oxoo, xooo, and xoxo, be phonemicized as /ooox/, /ooxo/, /oxoo/ and /xooo/, with the fifth phonetic sequence perhaps being a realization of /oooo/. A rule would be introduced to provide a secondary stress two syllables before a primary stress.

6 Welmers mentions (p.147) that "...forms which are not clearly stressed on a vowel before the last one are pronounced in isolation with a final high pitch." While I shall argue below that an utterance-final high pitch should not be identified as stress,
we might ask whether the "stress" in /bája/ cannot be treated as a H pitch-accent, in which case 'son' and 'daughter' would contrast in having, respectively, initial and final H tone.

7 This statement does not apply to languages which, in addition to having stress on every long vowel or heavy syllable, have a fixed stress on a given syllable of a word, e.g., in addition to having dominant penultimate stress, every long vowel is stressed in Hawaiian (Elbert and Keala 1961:6).

8 The term "free" is actually confusing. It originally meant that the stress, rather than being fixed on a particular syllable of each accentual unit (and therefore predictable by rule), was not constrained in terms of word structure. However, it is sometimes used in an opposite sense, as witnessed in the following statement: "Tibatulabal employs a free word-stress which expresses no systematic meaning; words are never distinguished by a difference in accentuation. The stress given to an element is not fixed, but may be changed according to a rhythmical pattern, the alternation of stress" (Voegelin 1935:75).

9 There is some question as to whether penultimate stress should be considered as demarcative, since stress is not realized next to a boundary: "Il y a démarcation imparfaite lorsque l'accent fixe n'est pas automatiquement sur une des syllabes extrêmes du mot" (Martinet 1954:20). Demarcation is even less perfect in languages which place stress on the first (or last) long vowel of a word. While one could probably predict the word boundaries in such a language, it does not seem likely that stress in such languages actually serves such a function. Rather, as I shall argue in section 3, an original demarcative stress has been affected by the tendency for stress to coincide with greater duration.

10 Where stress is determined in some way by a grammatical affix, it should of course be regarded as having a morphological status. Affixes have been observed to interact with stress in the following ways: 1) They can be ignored in stress-placement; e.g., in Syriac, "Final enclitics do not normally affect the stress pattern and primary stress remains on the penultimate syllable of the root form" (Solomon and Headley 1973:145). 2) They can require that the stress be placed on them; e.g., the -ei suffix in German obligatorily takes stress, as in Konditoré (Garde 1968:119). 3) They can require that stress be placed on a particular syllable of a word; e.g., -ity in English requires a directly preceding stress, as in serendipity, alacrity, etc. 4) The addition of an affix can cause a stress shift (or other change). In this last case it would appear that stress functions to highlight the difference between a stem and an affix. Thus, in Luiseno (Koreber and Grace 1960:8), stress is usually on the first syllable. However, when a suffix is added to a bisyllabic stem (which would be expected to have initial stress), stress jumps one syllable to the right to coincide with the stem boundary. The same phenomenon is reported for Western Miwok: "When
followed...by any syllabic increment (which may be a suffix, a stem in composition or some logically quite unrelated element) such stems tend to undergo a lengthening of the second syllable together with a shift of stress" (Freeland 1947:33). Although this question requires further investigation, it would appear to be a frequent occurrence that (2) and (3) result from borrowing affixes along with the normal stress placement from the loaning language. Thus, stress is placed finally on -oi (as in French) and antepenultimately with -ly (as in Latin).

11 For a detailed discussion of stress functions and typologies, see Garde (1968). I have not considered here all of the logical types one might distinguish, e.g. I have not mentioned what Garde refers to as "limited" stress languages. These are languages which do not have completely predictable stress, but in which stress is limited to particular syllables in a word. In Ioway-Oto (Whitman 1947:238), Ossetic (Abaev 1964:11) and Tarascan (Pester 1969:23), stress is found on either the initial or the second syllable of a given word. In Hanunoo (Conklin 1953:5), Sarangani Monobo (Meiklejohn and Meiklejohn 1955:2) and Tagalog (Llamzon 1966:35), three Philippino languages, stress is found on either the penultimate or the final syllable of a given word. Such languages can be viewed as partially fixed, partially free in their stress placement. In distinguishing between lexical and grammatical stress, a further qualification should be made. There are languages where stress cannot be predicted on the basis of external (i.e. word) boundaries, but only on the basis of where a stem boundary occurs. In this sense, lexical stress does not mean phonemic stress, as can be observed in the following statement concerning Burera: "Stress is phonemic and contrasts as in áma 'mother' and ama 'will get him'... Stress appears to be grammatically predictable, usually occurring on the first syllable of the word stem" (Glasgow and Glasgow 1967:7).

12 There have been several attempts to provide a production-based definition of stress, as opposed to a perception-based one: "Definitions of what is meant by stress are often vague. But the data reported in Ladefoged (1967) clearly show that an accented syllable in English is accompanied by an increase in subglottal pressure, which in normal conversational utterances is produced by the action of the internal intercostal muscles" (Ladefoged 1974:299).

13 In many languages secondary stress lacks one or more of the above perceptual cues. Thus, in Walmatjari: "The predominant contrastive features of primary stress are high pitch and greater loudness, whereas the predominant contrastive feature of secondary stress... is length" (Hudson and Richards 1969:183). In some languages having iterative secondary (or, to use Garde's term, "echo") stresses on alternate syllables, the same perceptual cue(s) may be involved (normally pitch), although even here there may be differences of degree, as in Cahuilla: "...pitch drops about a fifth from main stress to unstressed and then raises and falls about a third with the alternation of secondary stress and unstressed" (Sailer 1965:52).
This is not to say that the reverse, namely the development of grammatical stress from the levelling of lexically unpredictable stress, does not occur. The latter occurs as a stress change, while the development of lexical stress occurs as the result of segmental changes. (I ignore possible cases of stress changes by contact.) The only point I am interested in here is in explaining how stress becomes unpredictable. Having done so, it is now possible to address the discussion to where predictable or regular stress comes from.

Of course it will be necessary to determine with greater precision what constitutes a greater or lesser complexity from a production or a processing point of view. While I hope to address myself to this question in future studies, let it simply be stated that the less information necessary for the placement of stress, the less complex the stress placement rule.

One problem with assigning stress too far from a boundary is that short words may require a separate treatment. In a language with third syllable stress, a bisyllabic word would presumably get initial stress, while a monosyllabic word would receive stress on its only syllable. It is only initial and final stress which allow a general statement without complications (e.g. without parenthesis notation); cf. Kurylowicz (1958:375n): "Dans les langues à accentuation initiale ou finale les monosyllabes équivalent, respectivement, à la syllabe initiale ou finale d'un polysyllabe". It is interesting that no language has a rule which 1) places stress ante-penultimately, but 2) fails to place a stress on words consisting of one or two syllables (i.e. words which do not have an antepenultimate syllable).

Of the second syllable stresses languages covered in this survey, all but four are spoken in the Americas, and some of these languages (e.g. Dakota; Wallace Chafe, personal communication) may turn out to be best analyzed as having had one prefix preceding the stem, and hence can be reconstructed as having stem-initial stress at some earlier time.

The way this shift of stress is effected is summarized in the following hypothetical derivation:

\[ *CV-CVC\textsuperscript{\tilde{\nu}} \rightarrow C\textsuperscript{\tilde{\nu}}-CV\textsuperscript{\tilde{\nu}} \rightarrow CV-CVC\textsuperscript{\tilde{\nu}} \]  
(cf. \( CV-CVC\textsuperscript{\tilde{\nu}} \))

First, a CV- prefix (representing a syllable) is attached to a CVCV stem with final stress. In the second stage a secondary stress is assigned two syllables before the primary stress, in this case on the prefix (which is in antepenultimate position). In the third stage this secondary stress becomes primary, while the final stress is reduced accordingly (cf. the statement concerning final stress quoted from Bloomfield below).
There are, however, numerous reports in the literature of a penultimate stress language developing final stresses by the loss of a final vowel, e.g. Latin partí:re 'to leave' becomes partír in Portuguese, since *e is deleted after *r and -l at the end of a word (Camara 1972:24). In French, deletion of final vowels is more general, and the result is general final stress being assigned to "groupes rhymiques". However, this should not be seen as a stress change, per se, but as a segmental change which has brought about a restructuring of the stress system (cf. footnote 14).

In Hombert's work there is also some evidence that a pitch rise is more complex perceptually than a pitch fall, given the same interval.

A similar proposal is made by Goldsmith (1974), who has developed a framework for mapping "melodies" onto segmental matrices.

"The special difficulties in determining and describing the stress phonemes of a language like English lie in the relative nature of their phonetic characteristics and in the wide variability of the stresses in response to expressive prosody, i.e. intonation" (Newman 1946:171). Compare also Lehiste (1970:127): "Additional experiments with more complex patterns of fundamental frequency change suggested that sentence intonation is an overriding factor in determining the perception of stress, and that in this sense the fundamental frequency cue may outweigh the duration cue".

This interaction between stress and intonation is in fact much more significant than I have so far made it out to be. See, for instance, the numerous studies of Bolinger (especially Bolinger 1955), where intonation in English is seen to affect the actual placement of stress.

It is important to recall in this connection that the terms "word" and "word-final" actually refer to the word or other grammatical unit. Thus, it is likely that utterance-final processes will either first or alternatively be generalized to clause-final and/or phrase-final position. The word is a particularly enticing unit for such generalization, since the word boundary represents the smallest boundary constituting a "potential pause".

It is possible that features other than pitch are involved in the development of stress from intonation. Thus, consider the following situation in Cherokee: "Phonemes at the beginning and at the end of each contour are pronounced with phonetically loose contact, and those in the middle part of the contour stretch are pronounced with relatively close contact. As a result the impression obtained is that the beginning and end of a contour stretch is spoken slowly and the middle portion spoken fast:
dehigwenugi 'he pinched me' (-higwe- is spoken fast and de- and nugi slow)..." (Bender and Harris 1946:16). I may therefore be underestimating the role of rhythmic factors in stress, as pointed out to me by David Stampe. It is important to point out that in the above account I am assuming that it is the "neutral" or "unmarked" declarative intonation which will serve as input to the development of stress. This seems consistent with the function underlying stress, namely the demarcation of word boundaries (see footnote 9, however). When a speaker isolates an utterance-medial word and gives it the pronunciation it would have had in isolation, this word is generally "new information". If it were old information, the need for isolating it among the other elements of the utterance would not be felt. A logical direction for further study would be an investigation of how the functions of intonation naturally feed into the genesis of stress-accent.

26 There are a few accounts of languages having neutral rising intonation. It is interesting to note that only one, Chamorro, has a lowered pitch as a primary cue of word stress: "In normal continuous speech... primary stress is normally accompanied by a lower pitch level, and is followed immediately by a higher pitch level on the following syllable which is sustained until the next primary stress or juncture. Most terminal contours (except discourse final) are marked by rising pitch"(Topping 1969:71). The author later takes note of the aberrant nature of this stress system: "This is a very unusual stress-pitch relationship, and is one of the major sources of the 'foreign accent' of Chamorros learning English and of Americans learning Chamorro. I have not heard this particular kind of intonation reported for any of the other languages of the Pacific". One is tempted to propose a recent intonational "flip-flop" rule for Chamorro, changing high pitches to low and low pitches to high.

27 This generalization is confirmed in still another way. If it were important for there to be an unstressed syllable preceding a stress, we would expect to find some tendency to move stress out of initial position. Since initial stress does not exhibit such a tendency, the facts are consistent with the above view. However, the problem of final stress should not be overlooked. Reference has already been made to cases where penultimate stress has become final through the loss of final segments. The very fact that this happens would seem to cast some doubt on the hypothesis that a stressed syllable is optimally followed by an unstressed syllable (i.e., even more so than being preceded by one). On the other hand, it was also said that such developments are not motivated directly by stress considerations, but by segmental ones. While it is not entirely clear how to interpret these facts, the picture I have tried to portray is, I hope, a consistent one which will be confirmed by further investigation.

28 In fact, it is tempting to attribute certain stress changes to a second grammaticalization, as it were. Most Australian languages have initial stress, sometimes
with secondary stresses on subsequent alternating syllables, sometimes with a secondary penultimate stress, as in Walmatjari (Hudson and Richards 1969:183). General penultimate stress is found, however, in Anyula (Kirton 1967:25) and Wallbri (Reece 1970:9). As seen in the following derivation,

\[
*CVCVCVCVCV \rightarrow CVCVCVCVCV \rightarrow CVCVCVCVCV \rightarrow CVCVCVCVCV
\]

primary and secondary stress exchange relative values, and then the secondary stress falls out (or is transferred to occur on the second syllable before the newly created penultimate stress). This same derivation may also account for the change from Proto-Celtic initial stress to penultimate stress in Welsh (Fowkes 1966:45) and Breton (Smith 1946:53). The forces of intonation may, therefore, introduce a new (secondary) stress which may at some point become more prominent: "An optional secondary stress may occur [in Tojolabal, a Mayan language with primary final stress] on the first syllable of a long word, but this appears to coincide with stresses superimposed by sentence rhythm" (Supple and Douglass 1949:169) [my emphasis].

29 This statement needs to be somewhat clarified in view of the fact that languages such as Eastern Cheremis (Sebeok and Ingemann 1961:9) recognize full vs. reduced vowels in stress placement, rather than long vs. short vowels. Since reduced vowels will generally be shorter than full vowels in languages with such a distinction, this situation need not be interpreted as a counterexample to the above generalization, cf. below.

30 One reservation which is necessary concerns the possibility of a CVC syllable attracting stress in final position. Thus, in Bhojpuri, "Words with only short vowels and no vowel clusters or diphthongs are stressed on the penultimate syllable unless the word ends in a checked syllable, in which case the final syllable receives primary stress..." (Trammel 1971:139). Other languages, e.g. Araucanian (Echeverría and Contreras 1965:134), place secondary stress on final closed syllables. Note, first, that many languages have stress-placement rules which refer to the distinction between long and short vowels, but which do not refer to closed syllables to attract stress in positions other than word-final, i.e. as opposed to syllables with long vowels. It therefore seems justified to hypothesize that the Bhojpuri situation represents a mechanism for reinforcing a consonant in a vulnerable position, namely word-finally. In some languages, however, where a word final CVC syllable takes precedence over a penultimate syllable in attracting stress, it may be the case that a final vowel was lost historically (e.g. Spanish decir 'to say'). In this case, the final CVC reconstructs as an earlier CVCV and regular penultimate stress can be seen to have once applied. It would therefore be wrong to claim that such examples necessarily reveal a tendency for final CVC syllables to attract stress away from the penult. Similarly, in Arabic dialects, where a final VCC sequence preempts the regular stress placement rule, it is probable that historically two syllables were involved.
A similar argument might be made with respect to consonant lengthening (cf. Allen 1973:81). Namely, the consonant following a short stressed vowel could not lengthen if the language already had an opposition between simple and geminate consonants. While I have no data on this, it is interesting to note that most of the languages with syllable weight phenomena have not only a vowel length contrast, but also a consonant length contrast. It would seem that depriving a short vowel in open syllable position of the possibility of stress was the only way to keep long and short distinct, whether of vowels or of consonants.

In tone languages it frequently is the case that contour tones, which require or at least "prefer" greater duration in their realization, are not found in CV? syllables. Thus, Swadesh (1947:228) reports that Ixtlan has low, high, rising and falling tones, but of these four tones, rising and falling are not found before glottal stop. In Fe?fo -Damlileke a bisyllabic sequence rising-low (i.e. with rising tone on the first and low tone on the second syllable) is realized as a low followed by a falling tone, if the first syllable ends in a glottal stop. Because a rising tone requires greater duration than a low tone, the high part of the rise is shifted to the following syllable.

There are, however, languages where a vowel before a glottal stop is in fact reported as stressed (even exceptionally so); e.g. in Mundari: "Syllables ending in a phonetic glottal stop or a glottalized allophone of a stop are considerably more stressed than others" (Gumperz 1957:9). The case of glottal stop represents an interesting phenomenon since besides having a shortening effect on a preceding vowel, it has also a pitch-raising effect. Thus, depending upon whether duration or pitch is the primary cue reported for stress, conflicting results are obtained; cf. in Diegueño: "A stressed syllable containing a short vowel followed by a word-final /?/ is always higher in pitch than preceding parts of the word" (Langdon 1970:19).

Thus, to summarize, a glottal stop will deprive a preceding vowel of length-prominence, though it may provide a preceding vowel with pitch-prominence.

These intrinsic stress properties have been studied experimentally in several tone languages, e.g. Klowa (Sivertsen 1956), Navaho (Hung 1959). These variations most likely play some role in "accentogenesis", which can be said to have occurred as soon as stress becomes culminative in nature, i.e. as soon as one stress is limited to each word.

A similar stress assignment rule is found in Classical Arabic and probably other older Semitic languages. The main difference is that stress is normally not assigned to a final CV: or CVC syllable, but rather will fall on the last pre-final heavy syllable of a word, and on the initial syllable should there be no heavy syllable (Socin 1942:16).
This type of stress-placement rule has received considerable attention from phonologists studying the formal properties of phonological rules, e.g. Kiparsky (1972:190), Howard (1973:51), Leben (1973:119), Anderson (1974:101) and Halle (1974). While the languages I have mentioned above assign stress on the basis of vowel length, other languages scan the word for a full (vs. reduced) vowel, or simply for a heavy syllable. The proposals found in the above works vary from one to another, although the general conclusion is that the two subparts of such stress rules (e.g. stress the last long vowel; if none stress the first syllable) collapse into a single rule, given the proper formal apparatus.

This development is of course different from that which occurred in Spanish, Portuguese and Italian, where the loss of Latin vowel length resulted in phonemic stress.

In Bhojpuri, where we have already seen a tendency for penultimate stress, stress will be realized as antepenultimate if there is at least one syllable preceding it (i.e. if there are at least four syllables in the word), e.g. अगुआँकाराब 'to lead', where the final secondary stress is conditioned by the closed syllable (Trammell 1971:139). Compare Içü TÜpi: "In the predominant word patterns accent occurs on the penult in two, three, and four syllable words... and on the antepenult in five or more syllable words..." (Abrahamson 1968:17). Thus, there may be some tendency for stress to be centralized in longer words, although closer scrutiny of the attested cases is required.

Cf. Trubetzkoy (1939:196-197): "If a language has a free accent in addition to the nonculminative differentiation of prosodemes, it cannot employ the means utilized for this differentiation for the realization of accent as well". What this means is that in a tone language, pitch will tend not to be the primary perceptual cue of stress; and in a language with a vowel length contrast, duration will tend not to be the primary perceptual cue of stress.

One question worth raising, for instance, is whether one can predict what will happen to stress in compounding on the basis of stress placement in simplex forms. In familiar languages with initial stress (e.g. German, Hungarian, etc.), the second noun of a compound is reduced in stress. In a language such as Tajik (Rastorgueva 1963:10), which has final stress, it is the first noun of a compound which undergoes a reduction in stress. A systematic study of this relationship is badly needed.
REFERENCES

Fry, D. 1955. Duration and intensity as physical correlates of linguistic stress. J. Acoust. Soc. of Amer. 27. 765-768.


ACOUSTIC CORRELATES OF STRESS AND JUNCTURE *

Wayne A. Lea
Speech Communications Group
Sperry Univac DSD, P. O. Box 3525
St. Paul, Minnesota

1. INTRODUCTION

Stress is usually considered to be associated either with increased vocal effort on the part of the talker, or with a listener's perception of increased prominence on a syllable. Its vital role in linguistic analyses is well established, but worth reviewing. We shall briefly review the importance of stress analysis (section 1.1) and the companion need for studies of acoustic cues to syntactic junctures (section 1.2), and suggest (section 1.3) that these prosodic analyses are vital enough to demand extensive work on listeners' perceptions of stress patterns (section 2) and careful studies of procedures for determining stress and syntactic junctures from acoustic parameters (section 3 to 6). References are listed in Section 7.

1.1 The Importance of Stressed Syllables

Some words such as noun/verb pairs (permit/permit) are distinguished from each other almost entirely on the basis of stress. Phonological rules require information about whether a syllable (vowel) is stressed, unstressed, or reduced. Stress is acknowledged to be a crucial factor in the accomplishment of correct meter and rhyme in poetry, and the time interval between stresses is known to be the basic unit of rhythm even in spoken prose and conversation.

Stress patterns (and other prosodies) are important in the early stages of language acquisition, and must be properly mastered in second-language learning if one is to speak without a "foreign accent." Recent experiments (Lehiste, 1975; Pisoni, 1976) suggest that stress plays an active role in speech perception. Listeners perceive vowels with inflections in pitch as being longer than monotone vowels. Presumably this is because they decide that the inflected vowel is stressed, and then assume that since it is stressed it is also longer in duration. The fact that inadvertent phonetic errors in speech production (spoonerisms or "slips of the tongue"; Fromkin, 1971; 1973) usually exhibit the correct stress patterns suggests that stress plays a significant role in the organization of speech production and perception.

Stressed syllables occur in the important ("content") words of speech. When we want to emphasize a word, we give it particularly strong stress, by making it particularly loud, long or prominent in pitch. Contrastive stress (such as in "I said productive, not predictive" or in coordinate constructions like "I saw an old house and
a new house") particularly exemplifies how important words or syllables get stressed.

We thus see that stress is important to: the distinction between words (permit/permit); the conditions for application of phonological rules; correct rhythm; correct pronunciation without foreign accent; listeners' perceptions of phonological structure; the organization of articulatory units; and the location of important words, emphasis, and special contrasts.

Also, it has often been assumed that consonants and vowels should be more clearly articulated and easier to distinguish in stressed syllables than in unstressed or reduced syllables. Stressed syllables provide islands of phonetic reliability in vowel, fricative, and stop recognition, and reduced syllables appear to offer islands of reliability in sonorant consonant categorization (Lea, 1973a).

1.2 The Value of Syntactic Junctures

For decades, linguists have claimed that prosodic features indicate the immediate constituent structure of English sentences. Gleason, for example, considered intonation and stress as "the dominant elements in the syntax-signaling system," which not only mark boundaries of constituents, but "also mark the kind of constituents and the type of construction in which they are found" (1961: 169; cf. also Chomsky and Halle, 1968).

Phrase boundaries ("syntactic junctures") are closely intertwined with stress patterns, since the acoustic correlates of stress also are acoustic correlates of junctures. Such syntactic boundary markers are also important in acoustic analysis of speech, synthesis of speech, and models of speech production and perception. Linguists argue that "In general, the perceiver of speech should utilize syntactic cues in determining the phonemic representation of an utterance" (Chomsky and Miller 1963: 314, emphasis added). Perception theorists (e.g., Miller 1962: 81) have also argued that large units (on the order of phrases and clauses) are used in early stages of human perception of speech, and that detailed phonemic decisions are made later, and then only where they are needed to fill in information about the large units.

Experiments have shown that acoustic clicks superimposed on speech were perceived as occurring near certain major syntactic boundaries, regardless of the actual timing of the clicks within the speech continuum (cf. review by Gleitman and Gleitman, 1970). The perceiver appears to wait until the ends of such syntactic units before making decisions about detailed sound structure. Speech perception thus appears to involve making use of certain expectations and received cues to determine the syntactic structure of a sentence, and then using such syntactic information in guiding phonemic decisions.
1.3 Detecting Stress and Juncture from Acoustic Data

Despite the growing importance of prosodies in language descriptions (cf. Crystal 1969), speech perception models (Lea 1973; Lehiste 1975; Pisoni 1976), machine synthesis of speech (Mattingly 1966; Allen 1973), and speech understanding machines (cf. Newel, et al. 1971; Reddy 1975), it is only recently that the acoustic correlates of stress and juncture have been understood enough to permit the successful implementation of rigorous (in fact, automatic) procedures for detecting stress and junctures from acoustic data.

In this paper, we shall review the various studies of acoustic correlates of stress, by first considering how listeners' perceptions provide a 'standard' specifying which syllables are actually stressed (section 2), then proceeding from a review of earlier work on acoustic correlates (section 3), through a specification of complicating factors that interfere with the detection of stress from acoustic correlates (section 4), and culminating in descriptions of successful automatic methods for extracting stress and junctures from acoustic data (section 5). Conclusions and references are then provided (in sections 6 and 7, respectively).

2. Listeners' Perceptions of Stress

Acoustic correlates of stress can only be evaluated in comparison with some "standard" specifying which syllables are actually stressed. For studies of isolated words, such as minimal pairs of noun versus verb, a desk dictionary or a researcher's own intuitions may be sufficient. However, for studies of the stress patterns throughout sentences and discourses, that "standard" for stress assignment is not as readily established. The standard should be consistent from time to time, and largely independent of talker and listener idiosyncrasies. In this section, experiments will be described which show that listeners' perceptions of stress levels, obtained by any of several procedures, provide a consistent standard for stress assignment.

2.1 Selecting Procedures for Obtaining Stress Perceptions

A variety of techniques have been tried for obtaining listeners' perceptions of stress patterns. Three general issues are involved in contrasting such techniques: (1) What is the form of speech being studied? (2) How is the speech presented to the listener for his perceptions? and (3) What is the listener asked to do? We will consider these questions in turn, indicating the advantages and disadvantages of each alternative, and leading toward at least one procedure which gives a consistent standard for evaluating acoustic correlates.

The speech studied for stress judgments has generally been one of the following types: (a) synthesized speech; (b) undistorted natural speech in the listener's native language; (c) undistorted natural speech in a foreign language; (d) speech that has been distorted to convey only prosodic features; or (e) no speech at all.
(the individual is asked to provide stress judgments when given only the written text of the speech).

It is advantageous to be able to separately control acoustic features, as is done with synthesized speech, and to correlate those controlled acoustic changes with listeners' perceptions of stress (cf. Fry 1955, 1958; Bolinger 1958; Morton and Jassem 1965; Mattingly 1966). Yet, one must be cautious about simply extrapolating from results with unnatural synthetic speech to corresponding claims about natural speech. Also, when whole sentences or paragraphs are synthesized, it is difficult to control a local correlate of stress at one point in the utterance without giving careful consideration as to how that correlate fits into the total prosodic contours of the utterance. Given these cautions, synthesized speech can be an extremely effective method of providing well controlled tests of acoustic correlates.

When the listener hears undistorted speech in his native language, he is able to use all his knowledge of the language and all the incoming acoustic data to decide on stress patterns. It is possible that he may use the wording and syntax of the sentence to decide where stresses "should" be and not decide where they are based on the actual acoustic data. For that reason, attempts have been made to remove the "intelligibility" or word content, syntax, and meaning of the utterance, so the stress perceptions will be based on acoustic prosodic features only. (We will describe these acoustic prosodic features in detail in sections 3 and 4.) One method is to have the listener hear a foreign language, for which he presumably cannot establish the wording, syntax, or meaning. His perceptions are obviously then a complex composite of what he knows to be the correlates of stress in his own language, and what he hears as prominent syllables in the incoming foreign speech. (If stress correlates were different from language to language, he then would be providing judgments which would differ from those provided by a native speaker of that language.)

Another way to insure that the wording, syntax, and meaning of the utterance are not used by the listener to predict stresses is to spectrally distort the signal so that the vowels and consonants (the phonetic string) cannot be identified. Specific distortions that have been used include (1) adding enough noise so that the formant structure and spectral features of the speech are masked, but the overall energy envelope and fundamental frequency can still be determined; (2) low-pass filtering the waveform (below the first formant frequencies) so that only slowly-changing energy contours and the fundamental frequency are preserved; and (3) inverting the speech spectrum (high frequencies converted to low ones, and vice versa) so the speech is garbled, yet energy and fundamental frequency are preserved.

Individuals can also be asked to mark stress when given only the written text. They would presumably indicate stresses based on an internal model of how that sentence should be pronounced, or how they would pronounce it. Such an approach obviously
disregards speaker differences, and provides stress expectations, not judgments based on acoustic features.

On the second issue, dealing with how the speech is presented to the listener, a few obvious points can be made. Use of headphones rather than loudspeakers avoids effects due to room noise. A listener probably should not hear a syllable, word, or short phrase removed from its context, but rather hear something on the order of a clause or sentence, to provide a consistent context and total prosodic contour within which local variations can be given their proper attention and significance. The major issue is whether the listener should hear the sentence or clause once, a controlled number of times (such as once for each syllable within it), or an unlimited number of times. It is extremely difficult to provide many decisions based on hearing the utterance only once, although that may be argued to be the only way the listener normally and naturally hears utterances. Allowing the listener to hear an utterance an unlimited number of times (by rewinding a tape, using a tape loop, or having a computer repeat an utterance on demand) permits him to firmly decide on the stress of each syllable.

The third issue concerns what the listener is asked to do once he hears the utterances. He may be asked to tell which syllable in an utterance seems loudest, which syllable is the most stressed of the syllables in the utterance, which of two specific syllables is more stressed than the other, or what is the stress level for every syllable in the utterance. In the latter case, he might be asked to mark stress levels according to any of several theories about stress assignment, such as Pike's (1945) system, Trager and Smith's (1951) four level system, or the conventional system of primary, secondary, and tertiary stress. Alternatively, he may be asked to specify which syllables are stressed, which are unstressed, and which are reduced (Lea 1973a, 1974a). Another interesting alternative is to have him mark the stress level for each syllable on a continuum of possible values, so a "graph" of relative stress throughout an utterance could be obtained (Cheung 1975; Cheung and Minifie 1975).

Selecting among these various options in procedures for perception is not simply a question of right versus wrong. Presumably, for most purposes, one will be interested in total stress patterns throughout an utterance, so the methods which assign stress to each syllable are preferable to those which inquire about single syllables or pairs of syllables. Studies (Lieberman 1960, 1965, 1967) have shown that listeners have difficulty consistently assigning four or five levels to prosodic features, so that an approach such as marking syllables as stressed, unstressed, or reduced seems preferable to ones using more levels. A recent study (Minifie and Cheung 1975) showed that perceptions correlated slightly better from one trial to the next for the procedure using stressed, unstressed, or reduced labels than for procedures using a continuum.
To study acoustic correlates of stress in natural spoken English sentences, it thus seems that one of the best procedures is to ask native English listeners to mark for each syllable in the undistorted utterance whether they hear that syllable as stressed, unstressed, or reduced. This is the procedure we will be discussing further in sections 2.2 to 2.5.

2.2 Consistency of Perceptions

Experiments have been conducted that demonstrate the effectiveness and stability of listener's perceptions of stressed, unstressed, and reduced syllables in continuous speech, (Lea 1973a). Individual listeners were allowed to repeatedly listen to portions of tape recordings until they could mark each syllable as either stressed, unstressed, or reduced. These perception tests were repeated three times (separated by at least a few days), by each of three phonetically trained listeners. For two listeners, an individual listener's perceptions of which syllables were stressed agreed from one trial to the other trials, for 95% of all syllables. Thus, a listener confused about 5% of the syllables between levels of stressed and unstressed from one trial to another. His confusions between unstressed and reduced levels were much more frequent. Also, two of the listeners agreed with each other on 95% of their judgments as to which were the stressed syllables.

Recently, Lea's experiment (1976) was repeated for other speech texts with ten phonetics students at the University of Minnesota, with similar results. With three repetitions spaced apart one week or more, these less-highly-trained listeners showed 2 to 11% stressed/unstressed confusions from time to time. Listener-to-listener confusions were also generally around 7 to 15% for stressed/unstressed distinctions, though a few listeners gave stress perceptions which were quite different (as much as 37% different) from those of other listeners.

In a similar study at the University of Washington, Cheung and Minifie found that perceptions of stress on a continuum of possible values also were quite consistent (that is, with values closely correlated) from time to time or listener to listener.

Lea concluded that it was possible to determine, within about a 5% tolerance, which syllables are actually stressed in connected speech, by obtaining listeners' perceptions in the manner that was used in his experiments. Then, if a procedure for locating stressed syllables from acoustic correlates could locate 95% of all syllables perceived as stressed by majority votes of two or more listeners, it would be doing as well as one repetition of the perception tests would do for predicting the perceptions from another repetition of the experiment. It would also be doing as well as one listener would do in comparison to another listener. Thus, one can demand no better than 5% precision in stressed syllable location from acoustic data.
2.3 Is There a Normal, Unmarked Stress Pattern?

Bolinger (1972) has taken exception to the idea that the stress pattern in a sentence is systematically determined by the wording and syntactic structure of the sentence. He suggests that semantics and speaker intentions play an indispensable and critical role in stress assignment, so that you can predict stress "only if you are a mind reader". It is crucial to the utility of generative rules for stress assignment, and to the use of acoustic correlates in making claims about the stress patterns in analyzed sentences, to know whether there is, in any consistent sense, an "unmarked", normal stress pattern that will be exhibited in any utterance token of a given sentence structure, provided that no special emphasis or nuances are added to the intended token.

There is considerable evidence that there is a normal stress pattern which is predictable from the wording and syntactic structure of a sentence. Some of the linguistic evidence is to be found in the agreement among most aspects of various published generative rules for stress assignment, and the ability of those rules to predict stress patterns that agree with the intuitions of most linguists and speech scientists. Most published objections to available rules (e.g., Chomsky and Halle 1968; Halle and Keyser 1971; Bresnan 1971; Halle 1973) have been concerned either with details and subtleties in special cases of wording and syntax, or with the idea that one can devise some circumstance (no matter how bizarre) in which the predicted pattern would not occur. These minor objections do not seem to warrant the rejection of a syntactically-dictated normal stress pattern. Further discussions of such linguistic evidence is beyond the scope of this paper.

However, there is experimental evidence that strongly supports the existence of a normal stress pattern which is based on the wording and syntax of a sentence. For one thing, Lea found (1973a, Appendix B) that perceived stress patterns were very much the same for six different talkers reading the same text. They agree on how the sentences ought to be said. In addition, when subjects were asked to mark syllables in a written text as being stressed, unstressed, or reduced, without speech being provided, those individuals agreed substantially with what they marked when listening to the speech. Stressed/unstressed confusions between judgments with speech versus without speech involved about 10% of the syllables. Thus, what the individual judges the stress pattern to be, given only the wording and syntax that are intrinsic to the written text, is substantially the same as what he hears when presented with the speech. Finally, when individuals mark stress for syllables in the written text, without any speech, they are consistent from time to time and person to person (with stressed/unstressed confusions about 5 to 10%). They are consistent, and agree about which are the stresses, given only the wording and syntax as provided by the written text.

It might be argued that the predictions (or judgments) of normal stress patterns given only the wording and syntax (or written text) may be nice and consistent
from talker to talker, time to time, and 'listener' to 'listener', but they might not agree with stress as determined from the acoustic data. That is, the judgments may be based on an internal model of what stress patterns 'should' be there, and not depend upon or agree with the acoustic data at all. Experiments actually indicate that that hypothesis is wrong. Listeners hearing undistorted speech (or those who mark stresses given the written text alone) mark most of the same syllables as stressed as do listeners who hear only the prosodies in spectrally distorted speech. The reader can quickly verify this by taking any published results about stresses in distorted speech and comparing them to perceptions provided by a few available listeners (or a few individuals who will mark stress judgments given the written text alone). Also, published results show that listeners to a foreign language can successfully identify the stressed syllables (Fonagy 1966). They presumably cannot use the wording, syntax, or semantics to guess at stresses, but instead successfully use the acoustic data to determine stresses.

We may conclude that there is a normal, unmarked stress pattern, which most talkers will exhibit in reading a sentence, which listeners will predict given only the written text, which good generative rules will presumably predict, and which will agree with the actual stress patterns determined from acoustic data. Listeners' perceptions of stress patterns will provide a good standard for establishing the actual stress patterns which should be determined from acoustic correlates.

If one does not wish to go to the trouble of setting up an experiment where listeners provide stress perceptions from hearing recorded speech, a close substitute is to obtain judgments given only the written text. One caution is due in that case, however. Spontaneous speech (taken from conversations, etc.) may exhibit non-normal pronunciations such as special emphasis on contrastive words, etc., so that judgments given only the written text do not always agree with perceptions with the speech when the speech is spontaneous conversation or when special semantic or idiosyncratic nuances are conveyed. Listeners' perceptions provide the most reliable standard for evaluating acoustic correlates of stress.

3. PRELIMINARY STUDIES OF ACOUSTIC CORRELATES OF STRESS

Stress is not simply associated with a single acoustic correlate. It was traditionally associated with loudness (Sweet 1892; Bloomfield 1933), whose primary acoustic correlate is intensity, but we shall see in this section that it is even more closely associated with fundamental frequency contours and durations of vowels or syllabic nuclei. A brief review of earlier studies of the simplest acoustic correlates of stress is given in section 3.1, with some of the author's recent results superimposed over those earlier results, to reinforce earlier conclusions. In section 3.2, we will consider how the primary cues to stress may be combined to produce even better correlates. A summary of the primary and secondary cues to stress is given in section 3.3.
3.1 Simple Prosodic Correlates of Stress

Stressed vowels generally have higher intensity, higher fundamental frequency, and longer durations than their neighboring unstressed vowels. In early work, this was clearly shown with synthesized speech. In words (e.g., permit/permit) which could be either a noun or a verb depending on which syllable was stressed, it was determined that, other things being equal, the more intense vowel was heard as more stressed, thus changing which word function (noun or verb) was heard (Fry 1955, 1958). Figure 1 shows Fry's average results for the word pair object/object (large circles in the figure), superimposed on results that I obtained for 54 such word pairs (dots, for the second syllable stressed, crosses for the first syllable stressed; cf. Lea 1973c). Although there is some overlap, so that some intensity pairs appear ambiguous as to which stress pattern is intended, it is apparent that the more intense vowel is usually intended to be the stressed vowel.

However, Fry (1955) showed that duration dominates intensity as a stress cue, in that percentage changes in the relative durations of the two vowels caused more variation than variations in intensity ratios did in the percentage of listener's judgments that declared the word to be a noun versus verb. That is, more variation in judgments of stressedness occurred due to duration variations than due to intensity variations. Figure 2 shows Fry's results for twelve talkers saying the word pairs object/object (Figure 2a) and contract/contract (Figure 2b). In Figure 2a, there is a clear clustering of vowel durations depending on which word (x=object, o=object) was intended. When both vowels are similar low vowels, as in contract/contract, there is more overlap, illustrating that the stress effects are compounded with phonetic influences on durations.

Superimposed on Fry's results are my results for one talker's readings of similar stress pairs project/project (open and closed squares in Figure 2a) and compact/compact (squares in 2b).

Later (1958), Fry showed that variations in fundamental frequency (Fo, or "pitch") are even better cues to stress than durations or intensities (cf. also Bolinger 1958; Morton and Jassem 1965). The syllable with the higher fundamental frequency is heard as more stressed. This was confirmed by Lea's results (1972, ch. 5) for peak Fo values in vowels, for 54 word pairs, as shown in Figure 3. Interestingly enough Fry found that the magnitude of frequency difference between the two synthesized syllables had no marked effect; a frequency difference has an "all or none" effect, resulting in perceived stressedness on the higher-frequency syllable regardless of how much higher it was.

3.2 Combinations of Cues

If amplitude, duration, and Fo values are individually good correlates of stress, we might expect that some combination of them would be even better in
Figure 1. Intensities of the First and Second Vowels of Disyllabic Words with Contrasting Stress. Shown are Fry's average results for object (o) and object (●), superimposed over Lea's results for 54 wordpairs with the first vowel stressed (+) or second vowel stressed (●).
(a) Fry's results for 12 talkers saying object (o) and object (x). Lea's results for 1 talker saying project (o) and project (■).
(b) Fry's results for 12 talkers saying contract (o) and contract (x). Lea's results for 1 talker saying compact (o) and compact (■).

Figure 2. Durations of First and Second Vowels of Disyllabic Words with Contrasting Stress
Figure 3. Peak $F_0$ Values in the First and Second Vowels of Disyllabic Words, Showing Clustering Due to Stress Contrasts
determining stress levels. In 1960, Phillip Lieberman suggested a procedure for combining such cues to determine stress in dissyllabic words (noun/verb pairs spoken in carrier sentences). One major point which Lieberman observed (1960: 454) was that "the integral of the amplitude with respect to time over the entire syllable" provided some improvement in stress determination, compared to using amplitude alone or duration alone. This was confirmed in another study of stress in isolated words, by Medress, et al. (1971), in which the intensity peak, duration of a vowel, and peak fundamental frequency succeeded in obtaining 72%, 70% and 68% of the stressed syllables, respectively, when used separately. When duration and intensity were combined into the energy integral, 85% of the stresses were correctly found. The energy integral was thus established as an excellent cue to stress.

Medress and his colleagues also found that the shape of an Fo contour within a syllable was an important cue to stress. Rising Fo contours were very closely correlated with stress, while level or falling contours were not. They were then able to devise a modified procedure for peak picking on fundamental frequency contours, whereby they ruled out certain falling-Fo contours from being in stressed syllables, then asserted that syllables with prominent Fo rises were definitely stressed, and of the remaining vowels, they picked the one with the highest Fo peak value. This procedure correctly found 90% of the stressed syllables, which was a marked improvement from the 68% found from the peak Fo alone. Thus, Fo increases were found to be excellent cues to stress. Bolinger (1958, 1965) also claimed that rising Fo (and, in some cases, abrupt decreases in Fo) were among the best cues to stress. In fact, Bolinger avoided the use of the word "stress", preferring "accent" instead, to emphasize the fact that what are called stress distinctions in English are really pitch accents (1965: 17f). More recently, Lea (1973c; 1974a) found that a computer program which associated stress with every substantial rise in the Fo contour of a sentence successfully found 77% of all the syllables which listeners perceived as stressed, while obtaining 24% false alarms (that is, 24% of all Fo rises were not associated with stressed syllables).

The study of Medress and colleagues, as mentioned above, dealt with the most-stressed syllables in isolated words. However, they also recorded some of their words in the context of simple sentences. They found that the success in stressed syllable location went down substantially for the peak intensity and Fo peak, when used alone as simple correlates, while the duration worked about as well in context as it had in isolated words. The energy integral, on the other hand, improved in performance as a stress cue, when the words were placed in sentences. Lea (1973c) found that a computer program which located stressed syllables as high-energy chunks of speech surrounded by substantial dips in energy (that is, a program which used a form of energy integral) located 84% of all stresses in connected speech, with 28% false alarms. These studies clearly demonstrate the value of the energy integral as a robust cue to stress.
3.3 Primary and Secondary Cues to Stress

There are a few other acoustic features that correlate with stress. One that has frequently been mentioned is vowel quality (or, equivalently, "spectral content"). A vowel is expected to be more neutral and schwa-like the more unstressed (or reduced) it is, while stressed vowels should be more likely to exhibit the extremal articulatory positions (such as /ʌ, ʊ, æ, /) than their unstressed or reduced counterparts. The spectral content of such stressed vowels should thus be more like that expected with distinguishable articulatory target positions, since they are easier to reach in the long-duration stressed vowels.

While spectral content may be somewhat useful in distinguishing reduced from unstressed vowels, it seems to be of little value in distinguishing stressed versus unstressed vowels. Fry (1965) found that vowel quality is not a very dependable cue to stress. We shall see in section 4 that the intrinsic identity of the vowel (/ɪ/ vs /u/ vs /æ/, etc.) is more significant in how it interferes with the prosodic correlates of intensity, syllable duration, and Fo contours.

In addition to vowel quality, there are some other segmental phonetic correlates of stress. Plosives are more likely to be aspirated when they precede stressed vowels (Lehiste 1970: 141). Syllabic sonorant consonants (/l, n, m/) are cues to the presence of reduced syllables. Also, although it is rarely noted in the literature (exception: Potter, Kopp and Green 1947: 80), the occurrence of glottal stops is an indication of stress. Glottal stops occur in English just before word-initial vowels. (Though it is a complication we cannot discuss here, they are also occasionally used as a substitute for oral stops, as in [ b a ʔ I ] for [ b a D 1 ], "bottle"). Lea (1975b: 26-29) found that glottal stops are twelve times more likely to occur before stressed word-initial vowels than unstressed ones. If a glottal stop occurs, it very probably precedes a stressed vowel, and is often likely to be just after a major constituent boundary. The glottal stop is thus another potential cue to stress and constituent structure.

In a study of contrasting word sequences like "light housekeeper" versus "lighthouse keeper", Lieberman (1967: 147-159) showed that the time interval between onsets of vowels (called a "disjuncture") was a cue to which meaning was intended. Since some linguists consider the contrast in these structures to be one of stress contrasts (light housekeeper versus lighthouse keeper, in Lieberman's notation), it appeared that disjunctures were cues to stress patterns. A long disjuncture occurs between two vowels when the first vowel (the one spanned by the interval) is stressed. In general, since stressed vowels are longer, we would expect that the time intervals spanning them (that is, the disjunctures) would be longer.

In general, these "stress contrasts" which Lieberman studied can also be considered to be contrasts in syntactic bracketing. Then, if a disjuncture spans a
syntactic boundary, it will be longer. A brief 'pause' (filled or silent) seems to be inserted as a marker of the structural boundary. Since such structural contrasts often go hand in hand with stress contrasts, the assumed stress contrast is indeed evidenced by a longer disjuncture.

Perhaps more convincing evidence about the correlation of long disjunctures with stressed syllables would be found in the cases where syntactic junctures are not also involved. Even within a word, we expect that the time interval between onsets of two vowels will generally be greater if the vowel spanned is a stressed vowel.

In summary, we have found that the primary cues to stress are Fo inflections (particularly, rises in Fo) and high values of energy integrals. Simpler, but less reliable, cues are vowel (or syllable) durations and peak (or average) values of Fo and intensity. Disjunctures, which are closely associated with long durations of syllables, are also good stress cues. Secondary phonetic cues are vowel quality and occurrences of aspirated plosives, syllabic sonorant consonants (which mark reduced syllables), and glottal stops.

4. INTERACTIONS AND INTERFERING FACTORS

The acoustic correlates of stress are not completely invariable because they are also affected by other aspects of speech structure. In this section we consider some of the complicating factors that interfere with the extraction of stress from intensities (section 4.1), durations (4.2), and Fo contours (section 4.3). Because of the importance of Fo contours in stress detection, we shall discuss their interaction with sentence intonation in section 4.4. A model of how phonetic structures, stress, constituent structures, and overall sentence intonation combine to produce Fo contours is sketched in Section 4.5.

4.1 Factors Affecting Intensities

Stressed vowels have higher intensities. However, each vowel category has its own "intrinsic intensity", so that a stressed /i/ may be less intense than an unstressed /a/. As shown in Figure 4, the vowels which are articulated with the tongue high in the mouth, such as /i/ or /u/, have much lower intrinsic intensity than the low vowels like /a/ or /ə/. The vowels are arranged in Figure 4 with high front vowels on the left, high back vowels on the right, and low vowels in the center. The narrow solid line shows average relative intensities of the vowels as found by Lehiste and Peterson in their 1961 article. The bold solid line shows a similar trend for the averages of stressed vowels in Lea's 54 pairs of disyllabic words. The dashed line is for unstressed syllables in Lea's word pairs. (Lea's results have been shifted in relative dB values so they could be compared to Lehiste and Peterson's results, with the unstressed /a/'s made equal to Lehiste and Peterson's average results for /a/).
Figure 4. Peak Intensities in Various Stressed and Unstressed Vowels of 108 Disyllabic Words (Lea, 1972). Lehiste and Peterson’s (1961) Average Results are also shown.
Obviously, an absolute level of intensity (such as 85dB) cannot be declared as stressed, since some unstressed low vowels may have that energy, as well as some stressed high vowels. One might hypothesize that if intensities are normalized or adjusted on the basis of F1 (since F1 is low for high vowels and high for low vowels), intensities might be more directly relatable to stress.

Intensities are also influenced by the position of the syllable within the sentence and phrase (early syllables tend to be more intense, while phrase-final or clause-final syllables are often less intense). Vowels are less intense when surrounded by unvoiced consonants than when in the context of voiced consonants (Lehiste 1970: 123). Intensity and fundamental frequency are both affected by subglottal pressures and laryngeal tensions. Also, intensities are higher when formant frequencies coincide with harmonics of the fundamental frequency, so that interactions between Fo and formant frequencies affect relative intensities in such a way as to make stress determination more difficult.

4.2 Factors Affecting Durations

Stressed syllables tend to have longer vowels (or longer syllabic nuclei), but the intrinsic identity of the vowel has a significant effect on durations, just as was true for intensities (Lehiste 1970: 18-20). Low vowels have longer durations than high vowels, in a trend that is very similar to that shown in Figure 4 for intensities. (See House and Fairbanks 1953: 111; or Lehiste 1970: 120 for such plots for durations.) Thus, an unstressed /a/ can sometimes be longer than a stressed /i/. One might hypothesize that if durations were adjusted on the basis of F1 (so low vowels, with longer durations and higher F1 are reduced in duration relative to the shorter high vowels with lower F1), then durations might be improved as a stress cue.

Durations of vowels are also substantially affected by their position within an utterance. Phrase-final and pre-pausal syllables have vowels whose durations are lengthened by 20 to 50% over their values in other positions (Lehiste 1970; Klatt 1973, 1975; Kloker 1975; Lea and Kloker 1975). Thus a phrase-final unstressed vowel may be longer than the same vowel in a stressed phrase-medial position.

Other factors that are known to affect vowel durations include: the voicing (and other phonetic features) of following consonants (vowels are longer before voiced consonants than before unvoiced ones (Lehiste 1970: 10-23); the position in a word (Oller 1971; Lindblom 1968); and the rate of speech (including how the complex articulatory units such as syllables and phrases are timed). When the speech rate is increased, unstressed syllables are shortened more than are the stressed syllables, so relative ratios of stressed and unstressed durations are increased (Lehiste 1970: 38).

4.3 Phonetic Influences on Fo Contours

Figure 5 illustrates the fact that vowels that are articulated higher in the mouth have intrinsically higher Fo. Results are shown for various consonantal contexts.
Figure 5. Effects of Vowel Identity on Peak F₀ in the Stressed Vowel V₂ of /hCV₂C/ Utterances, for Talker GWH (Lea 1972, p. 110). Consonants are Pooled by Manner of Articulation. Superimposed on such results are the results of Lehiste and Peterson (1961), pooled for all their consonantal contexts.
(specifically, for various manners of articulation of the preceding consonant), as obtained for one talker (Lea 1973b; cf. also Lea 1972: 111). Superimposed on Lea's results are Lehiste and Peterson's (1961: 422) average results for several talkers. In contrast to the higher intensity and longer durations that accompany low vowels, the low vowels have lower Fo. Thus an unstressed /i/ may have a higher Fo than a stressed /a/.

In addition to the effects of vowel identity on relative Fo values, there are influences of consonants on Fo contours, which complicate the determination of stress from Fo contours. Another look at Figure 5 (for example, focusing on the difference in Fo for /i/ after an unvoiced fricative versus after a voiced fricative) should clearly illustrate that the consonant can affect the peak Fo in a vowel even more than the vowel identity (i.e., /i/ vs /a/) does.

During unvoiced consonants, Fo contours are blanked out. Also, sudden increases or "jumps" in Fo occur following unvoiced consonants, and are followed by rapid fall of Fo within a few centiseconds (hundredths of a second, equivalent to tens of milliseconds). Fundamental frequency will also decrease suddenly, or "dip", within voiced obstruents, then increase suddenly again at opening of the vocal tract, then increase gradually in the first few centiseconds of vowels (or sonorants) following such voiced obstruents. Such Fo changes due to consonant-vowel transitions interact with stress patterns in the utterances, so that an Fo fall in a vowel following an unvoiced consonant is more likely in unstressed than stressed syllables, while Fo increases following voiced obstruents are more likely in stressed than in unstressed syllables (Lea 1972, ch. 5).

These influences of phonetic sequences on Fo contours have been studied in isolated words and connected speech (Lea 1972; 1973b), and have been asserted to be a phonetic origin of the fact that high and falling tones usually follow unvoiced consonants in tone languages, while rising or lower tones are more frequently observed after voiced consonants (cf. Hyman 1973). However, the most important implication they have for the current discussion is their interference with stress correlates. For example, in a study of disyllabic words with contrasting stress patterns, Lea (1973b: 58-60) found that only 72% of all rising Fo contours in vowels accompanied stressed vowels. A rising Fo contour either indicated that a vowel was stressed, or the vowel was word-initial, or the previous consonant was voiced. Stressed vowels preceded by unvoiced consonants frequently showed falling Fo contours. Similarly, a procedure for automatically locating stressed syllables in continuous speech from acoustic correlates (such as rising Fo contours) sometimes failed to locate a stressed syllable when Fo fell in the stressed vowel after an unvoiced consonant. Also, some unstressed vowels would have rising Fo after voiced obstruents, which occasionally led to their falsely being called stressed.
4.4 Stress and Intonation

Other factors that are asserted to influence Fo contours include emotion (Huttar 1968), special emphasis (Lieberman 1967, ch. 7), co-reference (Cantrall 1969), and the rate of speech (Clean and Tiffany 1973; Lehiste 1970). Detailed discussion of these factors is beyond the scope of this paper. On the other hand, several factors that affect Fo contours are encompassed in the "intonation contour." These include the initial and terminal Fo contours in sentences and clauses, the position of a syllable within phrases, and effects of such syntactic constructions as subordination and coordination.

The overall intonation of English sentences has been characterized (Armstrong and Ward 1929) in terms of two alternative "tunes" or contours. Tune I has a characteristic rising of pitch (or its acoustic correlate, Fo) until the first stressed syllable in the sentence is reached, and a falling of the pitch from the first stressed syllable to the last. Sentence-final intonation was said to fall dramatically (for declarative sentences, exclamations, and questions with interrogative words). Tune II is like Tune I, but is terminated by a brief rise (or leveling) in pitch (within the last stressed syllable, if it is sentence-final; otherwise in the final unstressed syllables). Tune II was said to mark yes-no questions, uncertainty or indifference in expression, and various forms of incompleteness (cf. Lehiste 1970: 99).

I tested these claims with a set of 31 test sentences (involving eight talkers) and found that 88% of all 'initial' unstressed syllables (those before the first perceived stress in a sentence) had a lower Fo than the first stressed syllable. Over 93% of the 'final' unstressed syllables (those after the last stress in a sentence) had a lower Fo than the last stressed syllable. In 75% of all cases the Fo on a non-initial stressed syllable was lower than the Fo on the preceding stress. An "Fo staircase rule" does seem to apply, so that Fo progressively decreases from one stress to the next. Exceptions were found, in that 13% of the time Fo increased from one stress to the next, but we shall soon see that these increases are associated with the marking of major syntactic breaks in the sentence.

Joseph Olive (1974) of Bell Telephone Laboratories presented average Fo contours for some 50 declarative sentence structures, for which these same trends (rise in Fo up to the first stress, decrease in Fo from stress to stress, and a terminal fall in Fo after the last stress) were quite evident.

Another question regarding the interactions between intonation contours and stress patterns concerns how the Fo values in unstressed syllables compare with Fo values in neighboring stressed syllables. In the 31 test sentences, I found that 55% of all unstressed syllables had peak Fo values which were lower than both the preceding and following stresses, while another 30% of the unstressed syllables had Fo values which were lower than the previous stress only. The other 15%, for which Fo was higher in the unstressed syllable, were due to Fo variations associated with voiced
and unvoiced obstruents. These results agree with previous observations (Armstrong and Ward 1929: 4) that unstresses usually will either ride along on the falling intonation contour, or exhibit local dips in Fo below that contour.

Armstrong and Ward observed that complex sentences may be broken up into more than one "sense group", each sense group being marked by a Tune I or Tune II contour. They argued that different people divide their speech into different sense groups (cf. also Scholes 1971: 38), and that the division depends upon the type of message and whether the speech is conversational or deliberate (such as the reading of written texts). Despite this disclaimer, the intonation contours they gave actually frequently showed a close correlation between pitch rises and the beginnings of syntactic units.

Kenneth L. Pike (1945) divided intonation contours of sentences into primary contours and precontours. A sentence may be broken into several primary contours and associated precontours. Pike characterized these contour portions as follows:

A stressed syllable constitutes the BEGINNING POINT for every primary contour; there is no primary contour without a stressed syllable, and every heavily stressed syllable begins a new contour. (p. 27)

Immediately preceding the stressed syllables of primary contour there will oftentimes be one or more syllables which...are unstressed. These syllables may be called PRECONTOURS... (p. 29)

Trager and Smith (1951) asserted a one-to-one relation between boundaries between constituents and prosodic cues. They said (cf. Lieberman 1967: 189) essentially that each linguistic unit always is represented by a prosodic pattern clearly present in the acoustic speech signal. Lieberman summarized their position as follows:

The suprasegmentals always provide 'acoustic' cues that tell the listener how to divide the sentence for syntactic analysis. (p. 190)

In summary, there has been a spectrum of expressed viewpoints about intonational cues for constituent structure. One of the weakest hypotheses (Armstrong and Ward 1929) is that sentences may (but need not necessarily) be divisible into parts by intonation contours associated with any arbitrary (but fairly long) sequences of syllables or words. The units need not be syntactic constituents, and indeed the individual talker may divide (or not divide) a sentence differently from time to time, and different talkers may divide utterances differently. At the other extreme, all sentences are assumed to be divisible into syntactic units by intonational (or other prosodic) cues that always occur at unit boundaries (Trager and Smith 1951; Wells 1945). The phonological principle of invariance (Chomsky and Miller 1963) applied to such prosodic aspects of language would imply that a syntactic boundary always has an associated acoustic (or phonetic) boundary.
Marker manifested, and only when the syntactic boundary occurs will that acoustic marker appear (Trager and Smith 1951: 51). Linearity (Chomsky and Miller 1963) would imply that a boundary between two syntactic units would be manifested by acoustic features at the time stretch ("pause" or such) after the time stretch associated with the last phoneme of the earlier constituent and before the time stretch associated with the phonemes of the later constituent.

Malmberg (1963: 69) implicitly rejected the linearity condition for structural boundaries. He broke up utterances into "measures" on the basis of perceived intonation, yielding such divisions as the following:

The boys are ----- playing in the ----- street.

(ɪə bolz ar) ----- (pλeIŋ iə) ----- (strɪt)

Each measure or group has an "accented" (stressed) syllable and zero or more unstressed syllables. The breaks he shows occur just before the stressed syllables, and not necessarily at the points in the phonemic string where structural boundaries occur. We shall see in section 5.1 that this breakdown into groups is similar to what is obtained from automatic analysis of fundamental frequency contours. That is, strict linearity must be rejected if one is to succeed in finding acoustic cues to the syntactic breaks.

Recent experiments on Fo in several forms of connected texts have confirmed the ability to detect syntactic structures from Fo contours. Lea's thesis research at Purdue University (Lea 1972a chapters 2 and 3) involved four male and two female talkers reading a five-sentence paragraph composed of only monosyllabic words, portions of three weather reports, and a portion of a news report. Also, Lea processed some excerpts from conversational interviews, bringing the total amount of speech processed to 500 seconds of speech, from nine talkers. More recently, Lea (1973a; 1975) extended such studies to include six talkers reading a paragraph of the "Rainbow Passage" (Fairbanks 1940), ten talkers spontaneously speaking instructions (mostly questions or commands) for simulated or actual interaction with computers, and three talkers reading certain subsets from a database of 1100 sentences.

Lea performed intuitive constituent structure analyses of the readjusted surface structures (cf. Chomsky and Halle 1968) of these utterances, giving some thought to the transformational history of each sentence, and thereby predicting where major syntactic breaks might be expected to be manifested in phonological structure. The boundaries thus predicted by this independent syntactic analysis (which resemble the types of structural boundaries Malmberg displayed) were then compared with Fo contours of all the utterances. His results showed that a decrease (of about 7% or more) in Fo usually occurred at the end of each major syntactic constituent, and an increase (of about 7% or more) in Fo occurred near (but not necessarily coinciding with) the beginning of the following constituent.
Figure 6 illustrates the Fo contour (that is, Fo values, in Hertz, versus time in seconds throughout the utterance) of a typical sentence taken from a weather report. Fall-rise “valleys” (marked by vertical dotted lines) accompany the syntactically predicted boundaries (marked by arrows labelled with the categories of surrounding constituents).

4.5 A Model of Fo Contours

In summary, sentence intonation, constituent structure, stress, and phonetic content of utterances have all been shown to influence the shapes of Fo contours in English sentences. A simple model of the various factors interacting with stress in Fo contours is illustrated in Figure 7. (cf. Lea 1973b: 14-16). Each clause has an overall Tune I contour,1 if incompletion was to be marked on the first clause, as when it is a subordinate clause, a Tune II might occur in that clause; if a yes-no question were involved, the second clause could have ended in a Tune II terminal rise in Fo. The first clause has the highest Fo, while the large rise associated with the beginning of the second clause will normally not attain the full height of the earlier clause. There is a general trend toward decreased Fo values as the utterance proceeds. (This is even true for paragraphs, in that the Fo peaks in later sentences rarely are as high as that of the first sentence of a paragraph (Lea 1972)).

Next suppose that we know the first clause consists of five constituents, and the second clause consists of four constituents. We expect each such constituent to exhibit a Tune I-like contour (or, under some conditions, a Tune II-like contour), that "rides on", or centers around, the overall clause contour. Then, we might expect the general Fo contour shapes of Figure 7b.

Stress effects may then be superimposed on the archetype contours, by increasing Fo in the region of stressed syllables, and reducing the relative Fo below the archetype for reduced syllables. If the stressed syllables are at those time stretches marked by the dark horizontal bars and "S's" shown in Figure 7(c), then the contour of Figure 7(c) results.

Finally, the effects due to phonetic sequences are introduced in Figure 7(d). During unvoiced consonants, Fo contours are blanked out. Immediately following the unvoiced consonants, Fo is assumed to start high (higher than the contour from 7(c)) and rapidly fall. During voiced obstruents, Fo dips slightly below the contour of Figure 7(c).

Now compare Figure 6 with Figure 7(d). Hopefully, it is clear that the superimposing of influences due to sentence intonation, constituent structure, stress placements, and phonetic influences has yielded an Fo contour quite similar to the actual one in Figure 6, which the author was modelling as he sketched the contours of Figure 7.
Figure 6. An $F_0$ Contour (Vertical Axis Frequency; Horizontal Axis, Time), of a Sentence, with Predicted Constituent Boundaries Shown by Arrows (Labelled with Category Symbols for Surrounding Constituents), and Detected Boundaries Shown by Vertical Lines.
(a) Overall Sentence Intonation Contour for Two-Clause Sentence.

(b) Archetype F₀ Contours of Constituents Riding on the Overall Sentence Contour.

(c) Local Increases in F₀ Assigned Where Stressed Syllables (S) Occur, All Riding on Contours of the Constituents.

(d) F₀ Variations Due to Phonetic Contents Superimposed on Contours that Incorporate Stress, Constituent Structure, and Sentence Intonation Influences on F₀ Contours.

Figure 7. Progressive Composition of a Complex Intonation Contour from the Superposition of Effects of Sentence Type, Constituent Structure, Stress Patterns, and Phonetic Sequences. Each figure shows the contour of the prior step as a dotted curve upon which the effect of the new factor is superimposed. The sentence is that shown in Figure 2.
Finer-grain adjustments in the contours could be made based on such additional aspects as relative levels of stress, syntactic subordination (Crystal 1969), vowel heights, and consonantal place of articulation, etc. (Lea 1972b). In general, however, it is apparent that the superimposed effects of sentence intonation, constituent structure, stress, and phonetic content give an encouragingly accurate account of the structure of Fo contours. Such a model of Fo contours may be very useful in developing procedures for locating stresses and junctures from acoustic data.

5. ALGORITHMS FOR PROSODIC ANALYSIS

Since the acoustic correlates of stress and juncture include Fo contours and intensity and durations in vowels (or syllabic nuclei), procedures are needed for obtaining those cues from the acoustic waveform. Lea, Medress, and Skinner (1973b: 6) found that the total broadband energy in the speech waveform does not delimit syllabic nuclei very well, since it remains relatively high even during obstruents like unvoiced sibilants. A frequency-band-limited function which measured the energy in the range of 60 Hz to 3000 Hz, called a "sonorant energy function", performed best in isolating the syllabic nuclei. This band-limited energy function may be obtained by summing the power values in all frequencies from 60 Hz to 3000 Hz of a software linear predictive analysis (Lea, Medress, and Skinner 1973b) or by a simple analog filter. Values are converted to decibels for energy analysis.

Durations of syllabic nuclei can be determined from the sonorant energy function. For example, Lea (1973a; Lea and Kloker 1975) used the points at which sonorant energy dropped 5dB below the maximum value in the syllabic nucleus as the endpoints of the nucleus. Vowel durations would require more sophisticated phonetic segmentation, or the use of an additional energy function limited to 650 Hz to 3000 Hz, which seems to detect many vowel–nonvowel sonorant boundaries, and thus delimits the vowel portion of a syllabic nucleus.

Mermelstein (1975) uses a similar band-limited energy function, which is weighted more highly at some frequencies than others, to achieve a "loudness function" which is good for delimiting syllabic nuclei.

There are a variety of methods for determining the fundamental frequency in speech. These include cepstral analysis (Noll 1968), the residual error function of a linear predictive analysis (Makhoul 1973), harmonic identification in the frequency spectrum (Snow and Hughes 1969); peak picking or other time-domain detections (Miller 1974); and autocorrelation analysis (Soullé 1968; Skinner 1973; Gillman 1975). Studies at Sperry Univac showed that the autocorrelation method worked better than cepstral analysis or use of the residual error function of a linear predictive analysis. With any Fo-tracking method, care must be taken; to eliminate octave errors that occasionally occur; to "smooth" the contour by correcting any single values that are out of line with neighboring values; and to set a minimum energy threshold that must be exceeded before Fo values are declared, since low energy
is usually a cue to unvoiced speech. The energy threshold can be used to insure that Fo values are not given after unvoiced consonants, until energy is high enough to be sure the voiced syllabic nucleus has begun. This will minimize the magnitude and duration of the large Fo jump and rapid fall in Fo after unvoiced consonants. Thus, phonetic influences on Fo contours may be partly controlled by energy conditions on the Fo-tracking algorithm.

In this section we will outline: how Fo contours and timing information can be used to detect syntactic boundaries or junctures (section 5.1); and how Fo, energy, and the delimitation of syllabic nuclei may be used to locate stressed syllables (section 5.2).

5.1 Syntactic Boundary Detection

A computer program has been developed based on the regular occurrence of Fo valleys at constituent boundaries. In experiments on over 1000 seconds of speech, Lea (1972, 1973a, 1973b) found that the boundary detector correctly detected over 80% of all syntactically predicted boundaries. Of the less than 20% of “missing” boundaries, about half were due to predicted boundaries between noun phrases and following verbs (auxiliary verbs or main verbs). There is considerable evidence (e.g., contractions like "I've", subject-verb inversion rules, and the general clitic phenomena) that such noun phrase-verb boundaries should not be expected in phonological structure. Thus, neglecting noun phrase-verb boundaries, about 90 per cent of all other boundaries are detected from Fo valleys.

It is important to realize that Lea's algorithm is not locating syntactic boundaries, but rather only detecting them. Sometimes the rise in Fo into a constituent may be delayed due to phrase-initial weakly stressed syllables or function words like a, of, into, from, etc. A predictable delay occurs before the bottom of the Fo valley occurs, but the delayed boundaries are still considered correctly detected. In recent studies (Lea 1975b), the detected boundary associated with an Fo valley was shown to occur immediately before the first stressed syllable in the following constituent.

Besides this regular acoustic manifestation of boundaries between major syntactic constituents, some boundaries between minor constituents (e.g., between an adjective and a following noun) were also detected by the fall-rise patterns in Fo. In his earliest studies, Lea (1972, ch. 3) investigated the effects of specific constituent categories (noun phrase, verb, prepositional phrase, etc.) on boundary detection. The lack of regular boundary marking between noun phrases and following verbs has already been noted. On the other hand, around 95% of all boundaries before prepositional phrases were detected by Fo fall-rise valleys. Also, coordinate noun phrases or coordinate adjectives were always accompanied by Fo valleys between the conjuncts. Sizes of Fo valleys were also studied for the various syntactic categories (Lea 1972; 1975b).
Sentence boundaries were always accompanied by fall-rise Fo contours. In fact, the rise in Fo (around 90% change) after a sentence boundary was substantially larger than the usual rises (about 40% or less) after non-sentential constituent boundaries. In addition, sentence boundaries were usually (in over 90% of all cases) accompanied by long (35 centisecond) stretches of unvoicing. Here "sentence boundaries" refer to both boundaries between matrix (unembedded) sentences and boundaries between embedded full-clausal sentences.

These earlier results with the texts studies at Purdue University were confirmed in Lea's later studies (Lea 1973a) with other texts. However, the spontaneous sentences taken from "man-computer interactions" showed more monotonic intonation and occasional thoughtful "hesitation" pauses that could be confused with sentential pauses.

Boundaries were occasionally falsely detected at Fo valleys that were apparently not in any way syntactically related. These resulted primarily from the Fo variations between vowels and consonants that were described in Section 4.3. Requiring a minimum fall and rise in Fo of around 10% eliminates most such phonetically produced false alarms in boundary detection.

There also are important cues to syntactic boundaries in the timing of speech events. Kloker (1975; cf. also Lea and Kloker 1975) showed that vowel and sonorant durations in phrase final positions were 20% above their median lengths for all positions, so that 91% of the phrase boundaries perceived in spectrally-distorted speech were detectable from groups of lengthened syllables (cf. also Lea and Kloker 1975; Klatt 1973). Phrase final lengthening is thus an independent cue to phrase boundary detections. If the end of the group of lengthened segments exactly corresponded with the end of a phrase, we would have a potential method for precisely locating phrase boundaries. Unfortunately, Kloker found that the end of the lengthened group did not correspond with the end of the syntactic unit, but instead could occur within the unstressed beginnings of the next constituent, just as the Fo detected boundary can. Despite this inability to precisely locate the boundary, these results do suggest that phrase-final lengthening of vowels and sonorants provides an independent timing cue to syntactic structure.

In another experiment (Lea 1975a), the question was whether or not one could detect major phrase boundaries from timing of prosodic features alone (namely the onsets of syllabic nuclei found from energy contours), without the need for a prior determination of the phonetic sequence or the detection of lengthening of phonetic segments. Lea found that, in read sentences and paragraphs, as well as simulated man-computer interactions, time intervals between onsets of stressed vowels ("disjunctures") clustered near mean values around 0.4 second, with standard deviations of about 0.2 second. Mean disjuncture durations doubled with spanning clause boundaries, and tripled when spanning sentence boundaries. Mean pause durations, as measured by durations of unvoicing, tended to be equal to or twice the mean interstress interval, for clause and sentence boundaries, respectively. Syntactically-dictated pauses thus appear to be one- or two-unit interruptions of rhythm, which can be used as acoustic cues to clause and sentence boundaries.
Durations of pauses provide only one of several ways in which syntactic boundaries relate to rhythmic structure. We also found that substantial syntactically-dictated valleys in fundamental frequency contours tend to occur at integral multiples of the average interstress interval (Lea 1974a: 44-45). In addition, longer time intervals between stresses occur when a syntactic boundary occurs between the stresses, even if no pause occurs. The durations of interstress intervals tend to be small when only a word boundary intervenes, but durations are increasingly larger for boundaries between phrases, clauses and sentences.

To further study the potential for automatically detecting syntactic boundaries from long disjunctures, the data from Kloker's experiment was re-examined to obtain the time intervals between vowels in syllables which were perceived as stressed by five listeners. Kloker's five listeners had been presented with spectrally inverted speech and asked to mark syllables which were perceived as stressed and to mark phonological phrase boundaries. Spectrally inverted speech was used because it preserves the prosodic information in the speech, but garbles the phonetic structure, so that the recognized words and English syntax cannot be directly applied to aid listeners in deciding where stresses and boundaries 'should' occur.

The results (Lea 1975) showed that 95% of all major phonological boundaries (which often correlate with underlying syntactic boundaries) can be reliably detected from long (one-half second) interstress intervals in the speech. Thus, long interstress intervals are another cue to the presence of major syntactic boundaries.

To use long interstress intervals (such as those above 50 centiseconds in duration) as another cue to phrase boundaries, one first must have an automatic procedure for locating stressed syllables in connected speech, and then be able to use those time intervals between locations to establish the presence of boundaries. Errors in stress location would obviously propagate into errors in phrase boundary detection in such a scheme. Perhaps this suggests that monitoring the interstress intervals would best serve as a further check, or refinement, on the detection of phrase boundaries by any other means, such as from Fo valleys.

Detecting phrase boundaries from lengthened vowels and sonorant consonants also would be a complicated procedure. First, decisions about the occurrences of the vowels and sonorants would have to be made, and their boundaries would have to be determined so the durations might be measured. This requires sophisticated procedures for phonetic segmentation and labelling, with perhaps a requirement for a better-than-usual method of establishing phone boundaries. Errors in phonetic categorizing and duration measurement would then propagate into potential errors in phrase boundary detection. As with the interstress intervals, lengthening of phonetic segments might, however, be useful in reinforcing or refining phrase boundary detection from other independent methods, such as Lea's constituent boundary detector which uses Fo contours.
5.2 Stressed Syllable Location

The assumption of Lea's constituent boundary detector is that sentences consisting of several major grammatical constituents will be broken into several Tune I- or Tune II-like subcontours, riding on the general tune for a sentence or clause. Thus, as was illustrated in Figure 6, Fo contours in sentences with several major constituents will have major Fo changes associated with the constituent structure. We might call these rapidly rising and gradually falling Fo contours as "archetype constituent contours". They resemble Lieberman's (1967) unmarked and marked breath groups, and Pike's (1945) primary contours plus precontours, and other contours associated with "sense groups" in the literature.

Lea devised a general algorithm for automatically locating stressed syllables, which assumed that the Fo rise that marks the beginning of the "constituent" found by his boundary detector was associated with the first stressed (or "HEAD") syllable of the constituent. In a sense, then, the constituent boundary detector may be said to be detecting some stressed syllables (but not locating them). Lea's stress-location algorithm then located stressed HEAD syllables within the nearest portion of speech which had non-falling Fo and high speech intensity for a long duration (yielding a large integral of energy within the "syllable"). Substantial decreases in energy level marked the ends or limits of this stressed syllable.

If each constituent had exactly one lexical word with a major-stressed syllable within it (as has been suggested for deep structures; cf. Chomsky 1965; Emonds 1970), one might well expect a one-one correspondence between a detected constituent boundary and the presence of a following stressed syllable. In fact, however, surface structure constituents (both as predicted syntactically and as found by Lea's boundary detector) sometimes contain more than one stressed syllable. In Lea's stress location algorithm, other stressed syllables in the constituent were assumed to be manifested by deviations in Fo above an archetype line (a straight line on a semi-log plot from the Fo peak to the Fo value at the end of the constituent). Again, such stressed syllables were delimited by decreases in energy.

Lea's algorithm for stressed syllable location succeeded in locating an average of 89% of all syllables perceived as stressed by the majority votes of a panel of listeners (Lea 1973a; Lea and Kloker 1975: 19). It was shown to succeed in locating about 10% more of the perceived stresses in continuous speech than did two simpler algorithms based on local Fo rises alone or long syllabic durations alone. The archetype contour algorithm was also least sensitive to the type of sentence being processed, while the other algorithms showed quite different performance in yes/no questions.

Sargent (1975) has developed a method for locating stressed syllables on the basis of high Fo, high intensity, and long syllabic duration, plus some contextual constraints such as an assumed alternation of stresses and unstresses, comparison with the Fo and energy in neighboring syllables, etc. His approach does not require a previous search.
for syntactic boundaries or the concept of an archetype Fo contour within phrases. However, it does require some complex comparisons of average Fo and energy integrals in neighboring syllables, including rhythmic constraints and some conditional probability decisions. Sargent obtained 92% of all the perceived stresses in some spoken texts for which Lea also obtained 92% correct location of perceived stresses.

Still another approach to stressed syllable location, developed by Cheung (1975), uses the concept of alternating stress, and "anticipatory and recency effects" but is not fully automatic since it requires the human to specify "primary stresses" in the speech, from which the stresses of all other (lower-stressed) syllables are derived. Both Cheung and Sargent provide complete "stress level" patterns, including distinctions between stressed, unstressed, and reduced syllables (or, in Cheung's case, levels of stress). Lea's algorithm, as initially reported on (Lea 1973f, 1974a), does not distinguish unstressed from reduced syllables. Syllables are either stressed or not. Refinements to include more levels of stress assignment are being considered.

Each of these algorithms for stressed syllable location show the value of combining Fo, energy, and duration cues, and demonstrate that considerable success is possible in locating stressed syllables from acoustic correlates.

6. CONCLUSIONS

We have briefly noted some weak secondary correlates of stress, with the following conclusions:

- Syllabic sonorants are a cue to the absence of stress (or vowel reduction);
- Glottal stops occur much more frequently before stressed word initial vowels;
- Aspirated plosives are more likely to occur before stressed vowels;
- Vowel quality is a potential cue to stress; and
- Long disjunctures (vowel-vowel intervals) are known to span stressed vowels.

These cues are listed in order of increasing likelihood of being useful in locating stresses. However, we have seen that the primary simple cues to stress are (in order of increasing effectiveness):

- High intensity in the stressed vowel;
- Long durations of stressed vowels or syllabic nuclei; and
- High Fo values in the stressed vowel.

Combinations of the cues, and patterns of change for the parameters, work even better for stress determination, with the following strategies found to be particularly effective (again, listed in order of increasing effectiveness):

- Regions of rising Fo are associated with stressed syllables;
- Long chunks of high-energy speech (or, equivalently, large values of energy integral) are usually stressed syllabic nuclei; and
Long chunks of high energy with local increases in Fo are very likely to be stressed syllabic nuclei, so that stressed syllables occur in the region of the initial rise in Fo to peak value within a constituent, plus near local rises above a gradually falling "archetype" line in the later part of a constituent.

Factors which interfere with the correct determination of stress from acoustic correlates include the following:

1. **Vowel identity** - Fo increases, and intensity and duration decrease, as vowel height increases; thus, as the first formant (F1) increases, Fo increases and intensity and duration decrease.

2. **Consonantal context** - Unvoiced consonants before a vowel cause the vowel's intensity to be lower and Fo to be high and rapidly falling in the initial portion of the vowel; unvoiced consonants after the vowel cause it to be shorter and less intense. Opposite effects accompany voiced consonants.

3. **Position in the sentence and phrase** - Vowels and syllabic nuclei are lengthened and lowered in intensity at the ends of phrases and sentences. Fo can be modelled as a function of superimposed sentence intonation, archetype rapid-rise gradual-fall contours in constituents, stresses, and phonetic influences. Successive stresses in constituents have lower Fo values.

4. **Rate of speech** - As the rate of speech increases, all vowels and syllabic nuclei tend to be shortened, but the durations of unstressed syllables decrease more than those of stressed syllables. Intensities may go down as articulations fall short of target positions. Fo tends toward higher average values at faster rates, but with no known systematic difference produced between stressed and unstressed syllables.

Based on the model of Fo contours as including superposition of a basic rise-fall sentence intonation contour and archetype rise-fall Fo contours within major syntactic constituents, an algorithm has been developed for finding boundaries between constituents at fall-rise valleys in Fo contours. This algorithm successfully detects (but doesn't precisely locate) about 90% of the major syntactic boundaries (i.e., syntactic "junctures") in connected speech. Experiments show that syntactic boundaries are also closely associated with phrase-final lengthening of vowels and sonorants, and long time intervals between the onsets of stressed vowels.

The archetype contour algorithm for locating stressed syllables uses constituents as delimited by the constituent boundary detector, searches for head stresses as long high-energy regions near the rise in Fo to its peak value in the constituent, and finds
other stressed nuclei as high-energy chunks of speech near local Fo rises above
a steadily-falling "archetype line". This algorithm, which obtains 89% of the stresses
perceived in connected speech, equals or exceeds the performance of other available
methods for stressed syllable location from acoustic data.

There thus are good available procedures for detecting stresses and syntactic
junctures. However, there still is some room for improvement in such procedures,
and a definite need to continue studies of acoustic correlates of stress and how to
use stress and juncture in various applications. A more controlled study is needed
to determine exactly when the procedures work best, and why they fail when they do.
With arbitrary uncontrolled speech texts, there were usually several alternative
explanations for differences between one successful result and another failure.
Refinements of the prosodic analysis algorithms are then somewhat handicapped by
any such inability to isolate the source of any problem. Pairs of sentences need to be
designed to exhibit minimal contrasts in structure, so as to carefully isolate the
effects of each change in structure from others. To maximize the number of questions
answered by the fewest number of spoken sentences, several occurrences of such
isolatable contrasts can be packed into each sentence.

A speech data base (including 1100 sentences, each spoken by three talkers) has
been designed and recorded (Lea 1974b: to appear) to permit such systematic studies
of the prosodic correlates of various linguistic structures. This data base permits
controlled studies of prosodic patterns accompanying such linguistic structures as:
noun phrases (with or without articles, quantifiers, possessives, adjectives, etc.);
auxiliary verb constituents; relative clauses; complements; conjoined noun phrases,
verb phrases, or clauses; subordinate constructions; noun/verb stress pairs; various
sentence types; etc. Published hypotheses about prosodic patterns that accompany
various sentence structures can be carefully tested with pairs of the sentences which
are identical except for one difference in linguistic structure (or a few isolatable
differences in structure). Such controlled studies will help determine the correct
form of rules relating prosodic patterns to linguistic structures, and will permit the
refinement of procedures for locating stresses and syntactic junctures from acoustic
correlates.

FOOTNOTES

This research was supported by the Advanced Research Projects Agency of

Timing information for the clauses, constituents, words, and phones must
also be given. Lengths of constituents and clauses, for example, will depend upon
the number of words or syllables, rhythm effects, and perhaps the relative importance
of each constituent in the sentence structure. In Figure 7, we shall assume that such
timing information is provided, and consider only how the properly timed Fo changes
accompanying events of phonetic sequence, stress patterns, and syntactic structure
combine to yield the overall Fo contours.
REFERENCES


Lea, W. A. 1975a. "Isochrony and Disjuncture as Aids to Syntactic and Phonological Analysis," presented at the 89th Meeting, ASA, Austin, Texas.


ON THE VALUE OF PHONOLOGICAL EXPERIMENTS
IN THE STUDY OF ENGLISH STRESS

Larry Nessly
The University of Texas at Austin

Compared to phonetics experiments and psychological experiments, phonological experiments are a recent, rare phenomenon. The earliest phonological experiment that I know of is Berko 1958; since then about twenty experiments have been conducted. Those that have investigated stress are even fewer: Ladefoged and Fromkin 1965, Jackson 1973, Nessly 1974, Ohala 1974, and Elman 1976. My purpose here is two fold: to report on the results of an experiment I ran on English stress, and to show the value of experiments in studying stress by focusing on two kinds of application (phonological and sociolinguistic).

The usual alternatives to phonological experiments are elicitation of occurring forms (as used in studying "exotic" languages or collecting dialect data), natural language observation (overhearing natural speech, as in Labov 1966: 99,100), archival studies (such as inspection of dictionaries or grammars), and introspection. A drawback of natural observation and archival studies is that they are usually not suited for studying rare or obscure phenomena, since these phenomena are not readily found in natural usage or archival sources. Introspection has the weakness that the investigator's introspections may conflict with the majority of other speakers' intuitions; some data on what is commonly shared among speakers is needed. A difficulty with any study that uses standard pronunciations as the only data (as elicitation methods normally do) is that the pronunciations are often standard by virtue of being "correct" rather than being regular; to determine what is regular it is more useful to know how speakers would pronounce similar items that have no established, "standard" form.

Phonological experiments are an especially useful means of collecting data because the experimenter can elicit pronunciations of a wide variety of occurring and novel forms in a controlled setting. Since the experimenter can choose (or devise) the test items, it is possible to study a large number of preselected phenomena. By testing a group of speakers the experimenter can collect data on generalizations shared by speakers. To test the validity of standard pronunciations it is possible to use novel items which
have a shape similar to that of occurring forms. By having control of the testing situation the experimenter can study phenomena that would be harder to study in any other way. Of course, experiments, like the other methods of collecting data, have their own drawbacks; for example, the testing situation is a controlled, rather than natural one (other potential drawbacks are discussed in section 4 below). I will try to focus here on the advantages of conducting experiments, in the hope of stimulating more experimental studies of stress.

1. Hypotheses

1.1 Phonologically based hypotheses. In designing the experiment I included test items to test a number of hypotheses about the phonological structure of English stress. Rather than report on all aspects of the test I will focus on three of the more interesting or more thoroughly investigated hypotheses.

1.1.1 The first hypothesis to be discussed is the following: three-syllable words with geminates at the end of the penultimate syllable receive penultimate stress, if the last syllable contains a reduced vowel (e.g. Aladdin, colossal, molasses, persimmon), while words with single consonants receive initial stress (e.g. paladin, codicil, nemesis, specimen). The stipulation that the last syllable contain a reduced vowel is made because in a prior experiment (Nessly 1974: 605) forms with geminates which had unreduced vowels generally received initial rather than penultimate stress: Maretlich, Phemaggor, Sárollitz, and Thólemnoch, only ocellus, with a reduced final vowel, received penultimate stress.

To test the above hypothesis I devised the test items falinner/faliner and maretten/mareten, and repeated the pair ocellus/ocelus. For each pair I expected the first word to receive penultimate stress, and the second word to receive initial stress.

1.1.2. The second hypothesis is that words whose penultimate syllables end in stops receive penultimate stress (e.g. epoxy, humectate, Monadnock, prospectus). This is a new hypothesis and is part of a broader, more complicated hypothesis that I call "stress obstructing." The stress obstructing hypothesis includes claims about the role of sonorants and fricatives in addition to stops, about the influence of different consonant types in open syllables, and about the effects of different vowels in the penultimate syllable (see Nessly 1974: 286-310, 337-45, 422-91 for details). The relevant part here concerns the role of stops at the end of the penultimate syllable.

All but one of the test items (caroptoo) we re devised with velar stops (since most occurring examples contain velars): adegnize, calegmite, calgmite, manectol,
nalecty, penaxy, remectate, and talacter. According to the stress obstructing hypothesis they were expected to receive penultimate stress (the item caroptove had a special function discussed below).

Besides testing the general hypothesis about the role of stops (especially velars) in stress placement, several items were intended to test more specific hypotheses. If stress obstructing is correct, then the initial-stress standard pronunciations character, galaxy, recognize, and stàlagmite (this variant is cited in Jones 1956, Random House 1966, and Merriam-Webster 1966) are irregular. To test the regularity of stress in these forms I devised the similar test items talacter, penaxy, adegnize, and calegmite/calogmite. The assumption was that speakers had memorized the initial stress of the standard pronunciations, but when given similar novel forms to pronounce they would respond with regular penultimate stress. I tried not to make the test items too similar to the target words, such as using character to test stress in character, since speakers might tend too strongly to pronounce the test item just like the known word. Instead I tried to make the test items somewhat removed, while still preserving what appeared to be the critical elements of the target words.

The standard pronunciation stàlagmite and the test items calegmite and calogmite deserve special comment. In Nessly 1974: 340-4 I observed that in both British and American dictionary entries stàlagmite has been gaining historically over stàlagmite. Such a shift in pronunciation suggests that the variant stàlagmite is more regular than stàlagmite, a result that is contrary to stress obstructing. I was therefore interested in testing the regularity of stàlagmite with the test items calegmite and calogmite. (The difference in vowels is intended to test a hypothesis that is not relevant here.) If stress obstructing is to be supported, the responses should be calegmite and calogmite.

The two items caroptove and manectoi are of special interest in determining the validity of stress obstructing. Halle and Keyser 1971: 70 claim that initial stress is regular for three-syllable words that have a tense final vowel (which is not part of a suffix); an example is âncédote. This analysis treats as significant the question of whether the final syllable contains a tense vowel and whether that vowel is part of a suffix, but treats as irrelevant whether the penultimate syllable ends in a stop. In contrast, my analysis treats the presence of stops as significant but ignores tense vowels and suffixes. To decide between the two proposals I constructed the test items caroptove and manectoi, which contain properties considered crucial in each analysis. Since both items have final tense vowels (of functions as a tense vowel in this context) and since neither can be said to end in a suffix, Halle and Keyser would predict caroptove and manectoi. Since both items have a stop at the end of the penultimate syllable I would predict caroptove and manectoi.
1.1.3. The third hypothesis tested was a claim I have made (Nessly 1974: 458-70) that divisibility of a word into morpheme-like components can affect stress placement. In particular, I claim that a three-syllable noun divisible between the second and third syllables (e.g. bandersnatch as bänder plus snatch) is strongly subject to initial stress (bändersnatch). When the word has a similar structure but is less divisible between the second and third syllables, such as hypothetical banpëngsnatch, the claim is that the word is less subject to initial stress. I call this phenomenon "pseudo-morpheme division."

To test the pseudo-morpheme division hypothesis I devised the pair miftenpet and miyltempet. The two forms are very similar to each other; they differ only in voice for f/y and for place of articulation for n/m. Yet miftenpet is divisible into miften plus pet whereas miyltempet is divisible into myl plus tempet. I expected the first item to be pronounced miftenpet and the second to be pronounced miyltempet, as for a compound, or else miyltempet, as for a nondivided form with a closed medial syllable and a final lax vowel (cf. uënsail, addëndum).

1.2. Sociolinguistically based hypotheses. Recent sociolinguistic studies have used observational and elicitational techniques to observe the effect of such variables as age, sex, geographical location, ethnic background, race, occupation, and education on such phenomena as the placement of vowels in the vowel space, nonstandard pronunciation of consonants, and deletion of grammatical forms. Other phenomena, such as stress assignment, have been largely ignored. Although it has been observed that some kinds of stress patterns are associated with specific areas (e.g. Southern Insurance, British corollary), these regional patterns have not been studied systematically by sociolinguists or phonologists. The interaction of age, sex, education, occupation, race, and ethnic background with stress assignment has been similarly neglected.

In conducting the experiment I had an opportunity to study the variables sex, age, occupation, race, and education on stress assignment. Also, by comparing the results in this experiment (conducted in Austin, Texas) with those of a prior experiment (conducted in Ann Arbor, Michigan) it was possible to compare geographical differences in stress assignment. I expected no differences between Austin speakers and Ann Arbor speakers, since the test items used were not the type known to have stress differences. Also, there was no reason to expect differences in age (given that the speakers were adults) or sex to influence placement of stress. Education can affect stress placement (note learned abdömen versus common abdomen) but I did not expect enough of a difference in education among speakers to have any effect (occupation is probably relevant only insofar as it reflects education). Finally, blacks often have distinct pronunciations, but I did not expect black subjects to use responses based on those pronunciations in the formal test setting, especially given the level of education (and presumed language standardization) in the subject pool (see 2.1). In summary, I either had no hypothesis about what to expect or else could frame a hypothesis but had no reason to expect the hypothesized result.
2. Method

2.1. Subjects. Subjects were residents of an apartment complex in Austin, Texas. Forty speakers were tested; the responses of three were discarded, two for not being native speakers of English, and one because she could not see well enough to read the test items. As noted in 2.2, speakers were divided into two groups, each of which responded to a different word list (lists A and B - see the Appendix). Of a total of 25 females and 12 males, ten females and nine males responded to List A, and fifteen females and three males responded to List B; the even distribution for List A was intended, since I wanted a balance for at least one of the lists. Seven speakers (four females, three males) were blacks; by coincidence the females were all given List A and the males List B.

Speakers were generally well educated; thirty had some college education (eight of these had some graduate training) and only one had not finished high school. Twenty-five had had a class in a foreign language or had some ability to speak or understand one, of which thirteen had a background in Spanish (one speaker was bilingual in English and Spanish). Twenty-three speakers were mainly from Texas, five from elsewhere in the South, and nine mainly from the North. Thirty-three speakers appeared to be younger than forty; I judged twenty-one of these to be younger than thirty (all were at least eighteen).

2.2. Materials. Two different word lists were devised, each containing 39 test items typed double space in elite type in two columns (see the Appendix). Some test items in one list were paired with similar items in identical positions in the other list. For example, mivtempet is the last item in the second column of List A, and miftenpet is the last item in the second column of List B. Since a speaker was asked to read only one of the lists, the responses to items in one list were independent of those to items in the other list. The items were arranged so that words with similar properties were not placed next to each other, such as malocly and manectoil, or talumbary, desulatory, and peturny. Also, I tried to interrupt strings of three-syllable items with longer or shorter items.

2.3. Procedure. I introduced myself to residents as being from the University of Texas and said I was doing research on reading. For those who asked for further information I added that I was on the linguistics faculty and was doing a follow-up study on a topic in my dissertation.

Once inside the apartment I explained that I had a list of words for the speakers to read. They were told that a few of the words might be familiar but that most of them were made up or were rare. I stated that since the words were largely made up there was no right or wrong way to pronounce them and that the speakers should read them as they see them. At this point I turned on a cassette recorder (Craig, Model 2621), and speakers read the test items.
Subjects were allowed as much time as they wanted in responding to each item. After they had read the items I asked them to reread some words, if it sounded as though they had misread the words the first time (what was counted as a misreading is described in 3.1). After speakers' rereadings I asked them whether any of the words looked familiar. If they chose some made-up words as being familiar, or if they identified real but obscure words as being familiar (such as desultory or modifieative), I asked them what the familiar words mean. Lastly I asked if any of the items reminded speakers of any other words, and asked them to state the word they were reminded of.

Next I asked a few questions about the speakers' background. One question was whether they had heard any other languages besides English, or whether they had seen any others written. Then they were asked what kind of job they had, how much education they had had, and whether they were mainly from Texas; if they were not from Texas I asked where they had lived most of their life. The answers to all of these questions were recorded.

The speakers' responses to the test items and the questions were transcribed by the author.

3. Results

In this section I discuss the results and my interpretation of them.

3.1. Validity of responses. Some pronunciations were discarded as misreadings. I counted as a misreading any response that could not have arisen from the test item by ordinary rule or convention. Common examples are [kalarasi] for colercy, [ai' anayz] for adegnize, [ms1 sabal ] for melicable, and [mənətnə] for mareten. Other responses that were disallowed were attempts to pronounce the item with a foreign accent (generally Spanish) and responses interrupted by a pause during their pronunciation.

Pronunciations that should be mentioned as having been considered valid are [an] for en for Texas speakers, nonconstriction of ɝ for speakers from the South where this pronunciation occurs (e.g. [faline ] for faliner from a North Carolina speaker), and [e] or an [ɛ]-like quality for spelled ɝ before [k] and before ɬ. The pronunciation [ɛk] for a[k] (in talacter, talachter, and penaxy) was found only for Southern speakers (Texas, Louisiana, North Carolina), so that I accepted it as a probable Southern dialect pronunciation. The rendering [ɛl] for al was found for Northerners and Southerners, for stressed and unstressed ɝ. Since it was so common, and since most instances of it (13 of 22) were for unstressed ɝ (which can take on many colorings), I decided to accept it. The only other instances of [ɛ] for ɝ were found in the following six responses, all of which were counted as misreadings: two menēctoi, two cenister, one [zenamaken] (for zenamacan), and one [talactar] (for talacter).
Another type of response that should be mentioned is the multiple response, in which speakers either respond with more than one pronunciation when they first encounter the item, or else they add another pronunciation when they encounter it again (as when they are looking for familiar words). I have omitted multiple responses, through either source, from the results. Some experimenters, when encountering multiple pronunciations, have chosen to select one of the pronunciations (such as the first one) as the official response, but I have avoided such a practice, since at this time we do not know the status of earlier versus later responses.

Finally, it should be noted that I have accepted rereadings as valid responses. In the presentation of raw data (Table One) they are explicitly identified, in case anyone would want to know which of the responses are rereadings.

3.2. Phonologically based results.

3.2.1. For the three pairs of test items used to test the role of geminates in stress assignment the results were as follows (the number of initial-stress responses is given first, then the number of penultimate-stress responses): ocellus 3:11 versus oculus 7:4; maretten 4:12 versus mareten 10:0; and falinner 9:7 versus faliner 11:7. With respect to significance of results, the three sets contrast markedly. The results for maretten/mareten are significant at the extremely strong 0.0002 level, whereas those for ocellus/oculus are barely significant at the 0.042 level, whereas the responses for falinner and faliner are virtually identical.

It should first be noted that whatever the role of geminates is in stress placement, it will not be descried by any simple hypothesis such as that given in 1.1.1. It will have to account for why the geminate in maretten strongly influences stress placement while that in the highly similar form falinner has no effect at all.

Another observation is that the responses for ocellus/oculus and maretten/mareten do follow the original hypothesis, disregarding the difference in the strength of result. In contrast, the responses for both falinner and faliner diverge from the hypothesis: falinner has too many initial-stress responses and faliner has too many penultimate-stress responses. I have no explanation for the responses to faliner; those for falinner are discussed again in 3.3.2.

3.2.2. For items testing the role of stops the results (omitting those for carpoptove for the moment) were adegnize 6:4 (40% penultimate), calegmitme 0:11 (100%), calogmitme 4:8 (67%), manectol 3:9 (75%), nalecty 2:12 (86%), penax 3:10 (77%), remectate 2:9 (82%), and talacter 2:13 (87%). Except for adegnize, the items received at least a 2/3 penultimate-stress response. Such results strongly support the validity of stress obstructing. As for adegnize, the initial-stress responses may be due to some additional aspect of the broader stress obstructing hypothesis (see 1.1.2).
Regarding the testing of individual forms, it appears that the standard pronunciations galaxy and character are irregular, since the parallel novel forms received strong penultimate stress (penaL by 77% and talacter by 87%). As for the regularity of stalagmite versus stalagm'ite, the variant stalagm'ite is more strongly supported, by 67% for calogmite and by 100% for calegmite. In addition, of the seven speakers who reported that calegmite or calogmite reminded them of stalagmite, all gave the pronunciation stalagmite. It therefore appears that stalagmite is supported by the test results, in contrast to the support for stalagm'ite by the dictionary evidence (see 1.1.2).

I prefer to accept the test results; why there is a discrepancy between the two sources of evidence needs to be studied, since one of the two sources seems to support an incorrect conclusion.

Turning now to the results for caroptove (and manectoi) both received a strong (75%) penultimate response (2:6 and 3:9, respectively). For caroptove, only responses with tense final o's were counted, since testing Halle and Keyser's analysis requires that the final vowel be tense. For manectoi I included two instances of [manächtwa]. Although [wa] for o1 is a French pronunciation, the stress in the two responses is not assigned according to French rules (*[manächtwa]), and therefore must be assigned according to the English pattern. Since [wa] qualifies as containing a tense vowel, and since there seems to be no suffix [wa] in English, I have accepted [manächtwa] as a test of the Halle-Keyser analysis. Although the number of responses is rather low (especially for caroptove), the two results are quite consistent with each other. They suggest that stops do play an important role in attracting stress, and that Halle and Keyser's analysis is in error, since it does not recognize that role.

3.2.3. To turn to miftenpet and mivtenpet as tests of pseudo-morpheme division, the results were miftenpet 7:7 and mivtenpet 0:14. These findings are consistent with the original hypothesis, which predicted that miftenpet would receive more initial stress than mivtenpet (unless mivtenpet received compound stress); the difference is significant at the 0.003 level. The unexpectedly high percentage of penultimate stress for miftenpet (50%) may be explained in the following way. Although it is somewhat indistinct on the tape, some speakers seem to have assimilated the n to the following p, in effect responding to the test item miftenpet. Since the dissimilar place of n and p is important for dividing miftenpet into miften plus pet, speakers who treated the form as miftenpet could be expected to assign penultimate stress, as for mivtenpet. A strong test of the pseudo-morpheme division hypothesis would be to repeat the test, and observe the speakers' lips when they pronounced miftenpet. If those who pronounced n tended toward initial stress while those who pronounced m had penultimate stress, then the hypothesis would be even more strongly confirmed.

3.3. Sociolinguistically based results. Differences in age, education, occupation, and race were not represented widely enough to lead to any correlation
with differences in responses. However, one sex-linked difference did appear, as did several instances of geographical differences.

3.3.1. The sex-linked result concerned the responses to the item *falinher*. For men that response was 2:6 while for women it was 7:1. The difference is statistically significant at the 0.038 level. This is a surprising result, since there is no known reason why men should favor *faliner* while women favor *falinher*. Furthermore, female speakers who responded to *ocellus* and the highly similar item *marettan* on the other list favored *ocellus* by 9 to 3, and *marettan* by 9 to 4; it is a mystery that one set of women would prefer *falinher* while another preferred *ocellus* and *marettan*.

Whatever the explanation is for the female preference for *falinher*, that preference seems to be the source of the unexpectedly high initial-stress response for *falinher* (see 3.2.1). As the female preference is explained, this aspect of the results for *falinher*/*falinher* should also be explained.

3.3.2. Responses that split along regional lines were found for the items *capister*, *nolosc*, *alezstey*, and *malidogram*. The regional split was usually found between the subjects in the Austin experiment on the one hand and the subjects in the Ann Arbor experiment on the other. Before discussing the results I will briefly describe how the Ann Arbor experiment was run.

The experiment was run in Ann Arbor, Michigan, in 1972, and used substantially the same procedure as the one reported here; two groups of speakers were given two different word lists to pronounce, and when they were done they were asked a series of questions. The major differences in the two experiments are that for the Ann Arbor test I did not ask speakers to reread any of the items, and did not ask subjects what part of the country they were from. Speakers generally had more education and more contact with other languages, and were on the average a little older than in the Austin study. Fewer speakers were tested, thirty, and males outnumbered females seventeen to thirteen. Detailed information about the experiment can be found in Nessly (1974: 892-921).

One difference between the Ann Arbor results reported in Nessly (1974) and those reported here is that for the results here I have excluded the responses of three speakers: a nonnative speaker, a junior high school student, and a person with a hearing difficulty. The junior high school student is excluded because I have learned through more recent unpublished research (see 5.2) that junior high students sometimes differ from adults in their responses. The number of subjects in the Ann Arbor sample is therefore twenty-seven.

For most of the items that were used in both the Austin and Ann Arbor studies the responses were quite similar. The differences for the following four items are the only ones that are significant.
The most dramatic results concern the test item *capister*. Unlike the other three items, for this one there was not only a difference between Austin speakers and Ann Arbor speakers, but there was even a difference within Austin speakers, between those who had spent most of their lives in Texas ("Texas residents") and those from outside the state ("non-Texas residents"). Texas residents preferred the pronunciation *capister* (by 8 to 2) while non-Texas residents preferred *capister* (5 to 1). The difference is significant at the 0.047 level. The other non-Texans, the speakers in the Ann Arbor study, also favored *capister* (by 11 to 2); the difference between Texas residents and Ann Arbor speakers is significant at the 0.006 level. It therefore appears that speakers from Texas have one kind of pattern for *capister* while speakers from other parts of the country have another pattern.

The second result to be described is that for *nobico* (for Ann Arbor speakers, *Nobico*). For this and the other two items I will not break down Austin speakers into Texas residents and non-Texas residents, since the Austin speakers did not differ along these lines. Austin speakers preferred *nobico* (by 12 to 5) while Ann Arbor speakers preferred *Nobico* (by 8 to 2). The difference in responses is significant at the 0.030 level.

The third result is for *nalesty*. Austin speakers favored *nalest* (by 13 to 2) while Ann Arbor speakers favored *nalest* (by 6 to 4). The difference is significant at the 0.044 level.

The final item is *malidogram*. For Ann Arbor speakers there were four instances of *malidogram*, three of *malidogram*, and three of *malidogram*. Austin speakers, in contrast, had thirteen instances of *malidogram*, and only one each of *malidogram* and *malidogram*. Since the two groups differed mainly in whether they assigned antepenultimate stress or some other stress, one can test the difference in responses by comparing antepenultimate-stress responses with other responses. By this measure the difference is significant at the 0.013 level.

In assigning interpretations to these results I will first discuss the results for *nobico*/*Nobico*. One possible interpretation is that the difference in responses is due to the difference in capitalization of the items. I do not accept this interpretation. I believe instead that the explanation has to do with a proposal made in Nessly (1974: 559, 560) that words that end in vowels receive penultimate stress as the regular pattern, except for a few subcases such as words in -ico, -ea, and -era. What may have happened is that the special antepenultimate-stress pattern of words in -ico was lost, so that *nobico* received the penultimate stress of the general pattern. In fact, all words that ended in vowels received mainly penultimate stress: *ampeco* (14 of 16), *crotera* (14 of 15), *fesola* (14 of 16), *logea* (11 of 16), *lomea* (11 of 17), *meugla* (10 of 16), *nobico* (16 of 17), *penida* (14 of 18), *placina* (17 of 18), *plotina* (15 of 15), *ralido* (11 of 13), and *ralino* (14 of 16). Some of these (crotera, logea, lomea, meugla)
could have expected to receive more antepenultimate-stress responses (see Nessly 1974: 521-523, 529, 530, 551). It appears, then, that for some reason the Texas speakers have lost the special subcases of words which receive regular antepenultimate stress.

What would cause Texas speakers to overgeneralize penultimate stress for these words? I believe that it is their exposure to Spanish. Texas speakers encounter Spanish loanwords, phrases, city and river names, street names, and proper names, most of which end in vowels and have penultimate stress. Given enough exposure to these words speakers may generalize that all words that end in vowels receive penultimate stress, wiping out the special subcases. Since non-Texas residents responded with penultimate stress as well, I propose that they have lived in Texas long enough to be subject to the Spanish influence. According to this explanation the subgeneralizations would be lost in any area where Spanish words or names are widely encountered, as in the Southwest or in southern Florida. More generally, they would be lost in any place where there are large numbers of words that end in vowels and have penultimate stress, from such sources as Italian (e.g., in parts of the Northeast) or a Polynesian language (e.g., in Samoa or Hawaii).

A similar explanation may also be given for the Austin speakers’ preference for nalesť. In Nessly (1974: 435, 440) I proposed that words in -ý (such as made-up nalesť) tend to receive antepenultimate stress rather than the penultimate-stress pattern of words in -i. If Texas speakers were exposed to enough words that end in -ý and have penultimate stress, they might lose the antepenultimate-stress pattern for any new words in -ý and treat them as words in -i. Nalesť would be treated as nalesť and stressed nalesť. Whether such an influence can override the cue of spelled -ý is uncertain. At least, however, the proposed explanation suggests why Austin speakers and Ann Arbor speakers would differ on nalesť.

Whatever accounts for the Texas resident response capister, it is different from what accounts for the Austin speaker responses nobico and nalesť. For nobico and nalesť both non-Texas residents and Texas residents have the same pattern, to the extent that the few non-Texas resident responses can be interpreted (see Table One). For capister, however, non-Texas residents have a response distinct from the Texas residents and like that of the Ann Arbor speakers. Since whatever affects Texas residents’ responses to capister does not affect non-Texans, that influence must be more subtle or less pervasive than the Spanish influence suggested above. I have no proposal for what that influence is.

Concerning the Austin and Ann Arbor responses to malidogram, it is not clear which set of responses requires a special explanation. Since the Austin speakers and the Ann Arbor speakers both heavily favored antepenultimate stress in the similar form malinogram (by 13 to 0, and 10 to 0, respectively) and since the Austin
speakers also heavily favored antepenultimate stress in malidogram (by 13 to 2),
it would appear that the Ann Arbor responses need the special explanation. Indeed,
it might appear that the Ann Arbor responses are due to some type of error, and
the regular pattern is malidogram. On the other hand, the item malidogram may
have some special property that has been lost among the Austin speakers. How
this result is to be explained is not certain.

To summarize, the widespread occurrence of Spanish words and names in
Texas may influence Texas speakers' responses to nobico and nalesity. The cause
of divergence between Austin speakers and Texas speakers for capister and malidogram
is unknown.

4. Discussion

In this section I discuss possible complications or drawbacks in the way the
experiment was run. Topics are the role of spelling, the size of the sample, and
the role of placement of items on the list.

4.1. Spelling. A potential problem in the experimental design is that since
such heavy use is made of spelling in the presentation of test items to subjects, perhaps
the results reflect generalizations about spelling rather than about stress. Such a
conclusion is logical, but I think incorrect.

One way to arrive at the conclusion just described is to take the position
that speakers do not convert spelling into phonological structure and then assign
stress, but they rather assign stress directly to the spelling as spelling. By
this position, for example, speakers who encounter the item capister assign stress
not to /kapister/ but rather to the sequence of letters c-a-p-i-s-t-e-r. The result
is that speakers have (presumably) one set of rules to assign stress to phonological
structure and, apparently, a completely independent set of rules to assign stress
to pure spelling. I see no reason to accept this position. Rather, I think it is
more reasonable to assume that speakers convert spelling into structure and then
assign stress to the structure by ordinary phonological rules. If speakers pronounce
the items by applying rules to structure, then it follows that the generalizations
governing the responses are generalizations about phonological rules rather than
about spelling rules. Accordingly, generalizations about words that end in -ister,
for example, are generalizations about words ending in /stər/, not about words
ending in /stər/.

It is not entirely correct that the only function the spelling can have is to
represent phonological structure. There are, of course, spellings that signal, or
cue, certain kinds of pronunciations. For example, in 3.3.2 I contrasted the two
made up words nalesity and nalesi and claimed that their regular stress patterns are
nalesity and nalesi. Information about where to stress the two words is not contained
in the structure, which is /nalesi/ for both, but rather in the spelling. The
conclusion (mentioned above) that the test results reflect generalizations about spelling rather than stress can therefore be arrived at by claiming that the test items contain unintended spelling cues, so that the stress is inadvertently assigned according to spelling rather than to structure.

This second position can be countered by observing that spelling cues were anticipated and even made a part of the larger analysis. For example, I used the test item *manectoi*, with a final *i*, rather than *manectoy*, with expected *y*, because *manectoy* is subject to a division into *manec* plus *toy*. By the pseudo-morpheme division hypothesis (see 1.1.3) *manec* + *toy* would be subject to the stress *manectoy*. Since the purpose of using the item *manectoi* was to test stress obstructing rather than pseudo-morpheme division, I used the *i* rather than the *y*. As another example, the letters *e* and *i* in closed syllables often represent reduced vowels (e.g. *minister*). Whether *manectoi* is spelled *manectoi* or *manuctoi* can therefore affect where the stress is placed. This consideration has been taken into account in constructing the items; in fact, some pairs of items were constructed to illustrate just such a difference: *odontoid/odontoid*, *capister/capaster*, *talembary/talumbary*, and *calegnite/calegnite*. My reply, then, to the claim that the test items contain unintended spelling cues is that for the items discussed here, the cues were intended or else were not relevant to the phenomenon tested.

A final way to take the position that the results concern spelling and not stress is to hold that spelling cues (intentional or not) should not be an important aspect of a test of the validity of rules that operate on structure. By this position, for example, consideration of whether *nalesty* and *manectoi* contain final spelled *i* or *y* does not belong in a phonological test, nor do spellings like *mareten* if the geminate is a non-phonological cue for placing stress. With respect to *mareten* it is uncertain whether the geminate is a spelling cue or a phonological geminate, but before deciding that question it is necessary to determine whether stress does reliably fall on the syllable with the geminate; it is the second issue that is being tested. As for *nalesty* and *manectoi* I believe that speakers incorporate into their mental representations distinctive aspects of the spelling just as they incorporate distinctive aspects of the phonological structure. Whether this belief, or any opposing belief, is correct is more a matter of faith or temperament at this time rather than one to be argued.

4.2. Sample size. A second potential problem with the experiment is that the samples are small. Since the results reported here are all within the 0.05 level of acceptability, smallness of the sample does not prevent the results from being significant. Rather, what it does affect is the reliability of the results.

When samples are large it is possible that several favorable responses could have been reversed and still not affected the significance of the result; there is ample
cushion for idiosyncratic or bizarre responses. With small samples, however, a shift of even one favorable response can considerably affect the level of significance; there is little or no cushion for bizarre responses. The issue becomes one of how reliable each response in the sample is.

To make the idea of response "cushion" more concrete I have developed what I call a "significance margin." The idea is to have a measure of the degree to which responses could be reversed while still preserving significance. For example, as shown in Table Two, the responses for mivtempet (0:14) and miftenpet (7:7) are different at the 0.003 level. If one speaker in each group had given the opposite response then the new distribution would be 1:13 and 6:8, and the results would still be significant, at the 0.038 level. The original distribution (0:14, 7:7) has a significance margin of one, since one person in each group could have given the opposite response and the difference would still have been significant. In contrast, when two speakers in each group reverse their responses, the new distribution (2:12, 5:9) is no longer significant, being at the 0.192 level.

Another notion that can be used is that of the single "switch." In this instance the response of a single speaker from only one group is reversed. An example can be taken from the responses of Texas residents (2:8) and Ann Arbor speakers (11:2) for the item capister. The original significance level is 0.007. By switching a Texas resident response the new distribution becomes 3:7, 11:2 and the new level is 0.024, still significant. By switching an Ann Arbor speaker response the new distribution is 2:8, 10:3, and the new level is still significant, at 0.020. By switching both, at 3:7, 10:3, the new level is 0.074, and is no longer significant. For the original distribution a switch in either group still leads to significance but a switch in both does not; we might say that the original distribution has both of the possible "switches" but no "margin."

As mentioned above, there is a margin of one for mivtempet/miftenpet. When a switch for either group is added to that margin the results are no longer significant, being at 0.032 and 0.104. It is possible to add switches to margins to determine more closely the extent of cushion in the result.

Having established some terminology for discussing response cushions, note in Table Two that only two results (for maretten/maretten and mivtempet/miftenpet) have any margin, and that only one (for maretten/maretten) has a margin of two. Furthermore, of the nonmargin results only two (for capister (Texas residents versus Ann Arbor speakers) and for malidogram) allow switches. For the remaining four results any reversal of responses would render the results no longer significant. For such small samples, then, significance of difference is valid only as long as the responses are reliable.
Having established that reliability of responses is the central issue, I will now consider two aspects of that issue. First I will try to minimize the effects of bizarre responses on the results, and then I will try to gauge how reliable the responses are.

In considering the effect of "errors," or bizarre, idiosyncratic responses, on the results, it is important to keep several considerations in mind. One consideration is that gross errors, or misreadings, have been eliminated from the results. While one might expect some mistakes when speakers try to pronounce novel words in a formal situation, many of those mistakes are filtered out when the misreadings are discarded. As a second consideration, one should distinguish between significant errors and trivial errors. While some speakers might make subtle errors, ones that would not be counted as misreadings, not all of those errors will be relevant to the phenomenon being tested. Finally, one should distinguish between what might be called "random" errors and what might be called "inflating" errors. A random error is one that is indifferent to any prior generalization or hypothesis; a speaker might for some reason say maretten when maretten is closer to that person's intuitions, or might say maretten for some reason when maretten is more appropriate. In contrast, an inflating error is one made in the particular direction of the most common response. For example, if a subject in the present experiment says maretten when maretten is more appropriate, that error inflates the penultimate-stress count for maretten and artificially increases the significance of difference for maretten/maretten. On the other hand, if the speaker says maretten (when maretten is appropriate) then the error artificially reduces the significance; since, however, the main issue with respect to small samples is that they can represent as significant some difference that is not, an error that reduces the significance is not a problem. The only problematic errors are inflating errors, and these are only a small percentage of all the various kinds of errors that speakers could make.

An additional point regarding random errors and inflating errors is that some results can be discounted only by assuming a staggering preponderance of inflating errors. For example, in 3.3.1 it was observed that the female preference for falinner is inconsistent with the male preference for falinner, and with the female preference for maretten and ocellus. For the female preference to have the same distribution as the other three sets of responses it would have to be 2:6 instead of actual 7:1. But notice that five errors in eight responses are needed to produce discovered 7:1 from assumed appropriate 2:6; worse, ALL FIVE errors have to be inflating errors; still worse, that's a NET five inflating errors over any errors made by males for falinner or by females for maretten and ocellus. To discount the women's preference for falinner by this means is to invoke a deus ex errata. Some restraint is needed in invoking errors to explain away discomforting findings.

The second point to be made in minimizing the effect of errors on the results is that not all of the five main types of results are vulnerable to misrepresentation by
the smallness of the sample. The results of testing stress obstructing are clear and consistent regardless of the size of the numbers. Regarding the role of geminates, the findings for maretten/mareten have a safe margin of two, while larger numbers would probably not affect the outcome for falinner/faliner (except with respect to the sex-linked result); the test of ocellus/ocelus would, however, likely profit from a larger sample. The results for mivtempet/miftenpet do have a margin of one; whether that is safe enough is unclear. The major results that are vulnerable are those that concern the variables sex and geographical area. Even for these it is probably better to demote them to the status of "suggestive" results rather than discard them altogether.

The second aspect of the reliability of the responses, besides minimizing the effect of the sample size, is to consider ways of gauging that reliability. A straightforward means is to repeat the experiment and see if the results are the same. For some items I have been able to do this, by giving them to both Austin speakers and Ann Arbor speakers. Of twenty-three comparable items, fifteen had responses that agreed within 15% of each other (e.g., Austin speakers had a 79% penultimate-stress response (II of 14) for ocellus while Ann Arbor speakers had a 69% response (9 of 13), for a difference of 10%); four diverged significantly from each other (the four discussed in 3.3.2); and four were in between at 21% to 33%. It is not clear how significant this agreement is. Indeed, it will not be possible to accurately determine the agreement until it is known how much divergence is due to differences between speakers, how much to differences in tests, and how much to residual variability. At least these figures can be used to inform speculation about replicability of the results.

Other approaches can be used to judge the reliability of the test in disclosing actual phenomena. It is unlikely that the results of testing stress obstructing (see 3.2.2) are artificial or random, since those results are so systematic, given the highly similar responses to each of the test items (except adegnize). Also, the results for nobico/Nobico and miftenpet/mivtempet are rendered more credible because they can be given a systematic, plausible explanation. Again, the less certain results concern those linked with sex and geographical location.

4.3. Placement in the list. A third potential complication in the experimental design is that the responses to some of the items may have been influenced by responses to surrounding items. As a result, the placement of items in the list may have been a variable in the experiment. The best way to determine the effect of placement in the list is to repeat the test with the potential influence removed. The next best method is to locate instances where there may be influence.

I have inspected the responses and have found a number of sequences where adjacent responses have at least a two-thirds agreement in stress placement. However, it is also possible to find adjacent responses with at least a two-thirds disagreement (stress dissimilation?). Since it is impossible at this time to distinguish fortuitous stress agreement from real stress influence (if there is any) there is limited value in identifying all the instances of agreement. I will instead note that the stress obstructing items calegnite, caroptove, manectoi, nalecty, remectate, and talacter have at least
a two-thirds agreement with preceding items, but believe that most of this agreement is coincidental. Also, of the four items for which Austin and Ann Arbor speakers differed, I could find no agreement with stress in preceding items for the Ann Arbor responses. Among Austin speakers, responses to nalesy agreed with those to nobico for ten of fourteen (72%) speakers who gave valid responses, while two Texas residents (non-Texans are irrelevant) had capister preceded by odontold, talacter, ristten, ampoco, colercy, and saloden, while one had preceding ampoce, colercy, and saloden only. I believe that the agreement for nalesy is coincidental; the agreement for capister is more plausible, but not certain.

In summary, there is not much constructive that can be said about the role of placement in the list without further testing.

5. Conclusions

At the beginning of this paper I identified two goals: to report the results of a specific experiment and to show the value of experiments in general in studying stress. Much of the paper so far has focused on the description of the Austin experiment and the validity of the results. I now use that discussion to argue for the value of experimental studies of stress. Afterwards I discuss possibilities for further research.

5.1. The value of experiments. The contention I would like to support is that experimental testing is highly useful (at times indispensable) for studying stress phenomena. Sometimes experiments offer the only practical means of testing a hypothesis. Concerning Halle and Keyser's claim that stress in words like anecdote is regular (see 1.1.2), there are only a few words like anecdote which support the claim, and only a few words like Trubétzkoy which do not. By using an experiment it is not only possible to explicitly test the analysis, but it is possible to study the generalizations involved in great detail.

Other phenomena probably would never have been imagined, much less assigned a hypothesis, using ordinary means of evidence. Examples are the differences for maretten/mareten versus falinner/faliner, the sex-linked difference for falinner, and the geographically linked differences for nobico/Nobico, nalesy, capister, and malidogram. Once these phenomena have been identified through testing, the only practical means of studying them is through further testing.

Experiments also offer a direct, convenient way to test the regularity of standard pronunciations, such as galaxy and character. Other forms of evidence on the regularity of stress, such as historical change and dialectal differences, are often not available for a given form.

Finally, experiments can provide a striking, economical demonstration of a given phenomenon. One can cite, for example, various occurring forms which
suggest the validity of pseudo-morpheme division, but by using novel test items like near-minimal microtemp and microtemp one can demonstrate that validity simply and effectively.

In summary, experiments offer unique opportunities for studying stress because of the possibilities for devising novel test items which can effectively demonstrate obscure or subtle phenomena in systematic ways. In some instances there may be no other opportunity for studying the phenomena.

5.2. Further research. Most of the types of needed follow-up research have already been mentioned. Such phenomena as the effect of sample size, the effect of placement in lists, and the validity of sociolinguistically based phenomena can be studied by repeating the same items in another test. Other phenomena, the role of exposure to Spanish words and the role of geminates, require gathering new information, such as testing in other regions where there is Spanish influence, and testing other items that contain geminates. How one would specifically conduct these tests involves tactical decisions that need not be discussed here.

One topic that I would like to touch on is a third application (in addition to phonological and sociolinguistic issues) for experiments: acquisition. By using the same methods as those used here it is possible to study the acquisition of stress among literate nonadults (from third grade onward). I have conducted such a test among third, seventh, and eighth graders, and have found that this kind of study is indeed feasible.

Another use is to study acquisition among adults: foreigners learning English, and speakers of one "dialect" (e.g. Ann Arbor speakers) learning another (e.g. that of Austin speakers). The first study can be made by asking speakers with varying proficiency to pronounce a constant set of test items and see what stages they go through in acquiring the native pattern. The second study can be made by having different speakers who have lived in the new area for varying periods pronounce items relevant to the speech of that area; one can determine the stages of acquisition, the length of time it takes to acquire the new pattern, and which of the new elements are acquired.

In this third application of experiments, as in the other two, testing would be a valuable, perhaps indispensable, means of study. Its value lies in its being a flexible, systematic means of studying subtle stress phenomena.

FOOTNOTES

1. The experiment reported here was conducted in Austin, Texas, in August of 1975. Copies of tapes and transcripts are available (at cost) for those who desire
them. I would like to thank the following people for comments that have been useful in writing this paper: Alan Bell, Cathy Flores, Steve Herman, Wayne Lea, Leland Lester, and, especially, Larry Hyman.

2 The claim that words with single consonants receive initial stress is too strong; words that end in vowels tend to receive penultimate stress (see 3.3.2), as do noun and adjective forms of verbs (e.g. compliant, refusal, dextrous).

3 I chose to make these four items look like names and capitalized them. It is probably not significant that all the forms that received initial stress were capitalized while the one form that received penultimate stress was not.

4 These items were also given to third, fifth, and seventh graders (see 5.2); they were not capitalized, so that third graders would have one less complication in the spelling to deal with.

5 Adjective was included in the experiment but not to test stress obstructing; since it is an established word, speakers may respond with the established pronunciation without having to decide what kind of regular pattern it should receive. Talachter, another test item, could be given a velar rendering for the ch, but is also subject to other pronunciations ([ɛ] and [ʃ]) and is therefore not a straightforward item for testing stress obstructing.

6 Calegnite and caloggnite could also be test items of the Halle-Keyser hypothesis, except that speakers may interpret -ite as a suffix.

7 The relevant criterion in judging whether a speaker had "some ability" in a language was whether it appeared that the speaker had enough exposure to the language to draw generalizations about how it is pronounced. One speaker (possibly two—the information is vague) had had a class in linguistics. Having a class in linguistics or being bilingual appeared not to have an effect on the responses.

8 Three oclus responses with a [k] were discarded, for the following reason. As noted in 1.1.1, oclus and ocellus were used in the Austin and Ann Arbor studies. Five speakers in the two studies pronounced the c before e as a [k], which is unusual enough in speaker responses, but all five were for oclus; none were for ocellus. The discrepancy suggests a special property of oclus. Indeed, from misreadings and speakers' statements of words they were reminded of, it appears that the [k]elus responses were based on the stem ocul- 'eye', as in ocular.
For the reason *nobico* is not capitalized in the Austin test see footnote 4.

I did not ask non-Texas residents how long they had lived in Texas, so it is difficult to judge the plausibility of this explanation.

Other issues are discussed in Nessly 1975.

Due to a number of complications in the results I have omitted discussion of the results for these items.

REFERENCES


<table>
<thead>
<tr>
<th>List A</th>
<th>List B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column One</td>
<td>Column Two</td>
</tr>
<tr>
<td>plotina</td>
<td>talembary</td>
</tr>
<tr>
<td>magary</td>
<td>ocelus</td>
</tr>
<tr>
<td>nobico</td>
<td>pemida</td>
</tr>
<tr>
<td>nalesy</td>
<td>modificateive</td>
</tr>
<tr>
<td>malidogram</td>
<td>ralino</td>
</tr>
<tr>
<td>adjectival</td>
<td>tepilature</td>
</tr>
<tr>
<td>socum</td>
<td>calogmite</td>
</tr>
<tr>
<td>falinner</td>
<td>bulonesk</td>
</tr>
<tr>
<td>mareten</td>
<td>filecum</td>
</tr>
<tr>
<td>fetoos</td>
<td>luperinary</td>
</tr>
<tr>
<td>colapis</td>
<td>penaxy</td>
</tr>
<tr>
<td>logea</td>
<td>nobiko</td>
</tr>
<tr>
<td>comparable</td>
<td>caroptove</td>
</tr>
<tr>
<td>odintoid</td>
<td>ompustry</td>
</tr>
<tr>
<td>talachter</td>
<td>fesola</td>
</tr>
<tr>
<td>ristinen</td>
<td>remectate</td>
</tr>
<tr>
<td>mengula</td>
<td>celotis</td>
</tr>
<tr>
<td>adjective</td>
<td>adegalize</td>
</tr>
<tr>
<td>salonen</td>
<td>mivtempet</td>
</tr>
<tr>
<td>capaster</td>
<td></td>
</tr>
</tbody>
</table>
### Table One: Raw Data

1. adegnize-6 adegnize-2 adegnize-2; adegnize-1; other-8.
2. calegnite calegnite-7 calegnite-1(C) Qcalegnite, Qcalegnite-1 calegnite-1 calegnite-1 calegnite-1 calegnite-1 calegnite-1; calegnite, calegnite-1; other-6.
3. calogmite calogmite-3 calogmite-1 calogmite-5 calogmite calogmite calogmite-1 calogmite-1; calogmite-1; calogmite-1; other-5.
4. capister T res: capister-2; capister-5 capister-2 capister-2; capister-1; other-1. non T res: capister-3 capister-1 Qcapister, capister-1; capister-1; capister-1; other-1.
5. caroptove caroptove-1 caroptove-1 caroptove-5 caroptove, caroptove-1; caroptove-1 caroptove-2; caroptove, caroptove-1; caroptove, caroptove-1 caroptove-1 caroptove-1 caroptove-1 caroptove-1(C) G caroptove, caroptove-1; other-4.
6. faliner faliner-9 faliner-1 faliner-1; faliner-1 faliner-1; faliner-1 faliner-1; other-0.
7. faliner (by speaker numbers) faliner-5; 11, 14, 17, 31, 34 faliner-9, 40 faliner-1; faliner-21, 35 faliner-7 faliner-19 faliner-25 faliner-24 faliner, faliner-9; other-3, 22, 23.
8. faliner (for responses corresponding to numbers, see 7.) Male: A-5, 14; P-7, 19, 20, 21, 24, 25; other-22. Female: A-1, 9, 11, 17, 31, 34, 40; P-35; other-3, 33.
9. malidogram T res: malidogram-1 malidogram-4 malidogram-2 malidogram-1 malidogram-1 malidogram-1 malidogram-1 malidogram-1; other-1. Non T res: malidogram-2 malidogram-1 malidogram-1; malidogram-1; other-2.
10. manectoi manectoi-3 manectoi-3 manectoi-3 manectoi-3 manectoi-3 manectoi-3 manectoi-3 manectoi-3 manectoi-3; other-5.
11. mareten mareten-7 mareten-1 mareten-1; other-8.
12. mareten mareten-2 mareten-1 mareten-1 mareten-8 mareten-2 mareten, mareten-1; other-2.
13. miftenpet Qmiftenpet-2 miftenpet-1 miftenpet-1 miftenpet-1(R) G miftenpet-1 L miftenpet, QG miftenpet, miftenpet-1; miftenpet-2 G miftenpet-1 miftenpet-3(R) G miftenpet-1; G miftenpet-1; other-3.
14. miftenpet miftenpet-3 miftenpet-2 miftenpet-3(R) miftenpet-3(R) miftenpet-3(R) miftenpet-3(R) miftenpet-1; other-5.
15. nalesty nalesty-1 nalesty-1; nalesty-10 nalesty-1 nalesty-1; other-4.
17. nobico T res: nobico-2 nobico-1 nobico-7 nobico-2; other-0. Non T res: nobico-2 nobico-2 nobico-2 nobico-7 nobico-2; other-2.
18. ocellus ocellus-3 ocellus-7 L ocellus-9 ocellus-2 ocellus, ocellus-1; ocellus, ocellus-1 ocellus, ocellus-1; other-2.
19. ocellus ocellus-2 L ocellus-9 ocellus-2 ocellus, ocellus-1; ocellus-2 ocellus-1 ocellus-1 ocellus, ocellus-1; ocellus-1, ocellus-1; ocellus-1 ocellus-1; other-3.
20. okelus okelus-1 okelus, okelus-1; okelus-1 okelus, okelus-1; okelus-1 okelus-1; other-3.
Explanations: 1. Responses are given in the following order: initial stress, stress on the second (third, etc.) syllable, multiple responses, valid responses not relevant to the hypothesis, 'other' (=misreadings). 2. Abbreviations - A = antepenultimate stress; G = geminated (G abcd indicates that the b is geminated; in G abcd both b and d are geminated); L = lengthened (L abcd indicates that the b is longer than for a geminate); LG = lengthened & geminated (in LG abcd the b is lengthened and d is geminated); non T res = non-Texas resident; P = penultimate stress; Q = with question intonation (can be combined with other symbols, as in QG); nR = n number of rereadings included; R = the response is a rereading; T res = Texas resident.

3. Vowel notation - \( \hat{a} \) = tense native vowel (e.g., \( \hat{a} = [\acute{a}] \)); \( \hat{V} \) = tense foreign vowel (e.g., \( \acute{a} = [\acute{a}] \)); \( \check{V} \) = lax vowel; \( \check{y} \) = reduced vowel; \( \hat{\check{y}} = \hat{y} \); \( \hat{\check{o}} = [\check{a}] \). 4. Underlined response = a rereading. 5. N in miftenpet = a nasal indeterminate on the tape between n and m. 6. For falinner, numbers indicate the speaker-number of the subject (e.g., '7' is the seventh speaker tested).

Table Two: Statistics

<table>
<thead>
<tr>
<th>Table</th>
<th>T res</th>
<th>AnAr</th>
<th>T res switch</th>
<th>AnAr switch</th>
<th>1 margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>capister</td>
<td>2</td>
<td>6</td>
<td>0.006</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 tail</td>
<td>AnAr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>capister</td>
<td>2</td>
<td>5</td>
<td>0.049</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 tail</td>
<td>non T res</td>
<td>5</td>
<td>1</td>
<td>0.119</td>
<td></td>
</tr>
<tr>
<td>1 margin</td>
<td>3</td>
<td>7</td>
<td>0.364</td>
<td></td>
<td></td>
</tr>
<tr>
<td>fallinner</td>
<td>M</td>
<td>2</td>
<td>6</td>
<td>0.040</td>
<td>3</td>
</tr>
<tr>
<td>2 tail</td>
<td>F</td>
<td>7</td>
<td>1</td>
<td>0.013</td>
<td>12</td>
</tr>
</tbody>
</table>
Explanations: 1. In all cases a Fisher exact test was used (see Blalock 1960: 221-5 for description of the test). A '1 tail' test is one for which the direction of imbalance in the distribution of responses is predicted beforehand (e.g. that miftenpet receives more initial stress than mivtempet). A '2 tail' test is one for which no prior prediction is made. Probabilities are accepted as significant if they are at 0.05 or less. 2. Abbreviations -- AnAr = Ann Arbor; Aus = Austin; F = female; M = male; see also "Abbreviations" section in Table One Explanations. 3. For explanation of 'margin' and 'switch' see 4.2 in the text.
INTRODUCTION

Although "stress" has been observed in many of the world's languages, its physical correlates have been widely studied only in European languages and of these, one language, English, has been the most extensively investigated. Therefore, the points made in the following review of the physiological correlates of stress should not be taken as necessarily applying to any language outside this restricted range. Moreover, I will make no attempt to give systematic coverage to those aspects of the phonetics of stress and accent covered in some of the extensive reviews already available, e.g., Vanderslice (1968), Broad (1968), Lehiste (1970), Ohala (1970), Benguerel (1970), Gärding (1973), Netsell (1973), van Katwijk (1974), and Di Cristo (1975). I will instead concentrate on some of the more controversial aspects of the subject.

EARLY VIEWS ON STRESS

Early assumptions on the articulatory correlates of what was called "stress" in such languages as English, German, and Russian generally had it that stressed syllables were produced with greater articulatory, especially expiratory, force and thus greater intensity than unstressed syllables (Sweet 1890, 1911, Passy 1895, Jones 1960, Grannmont 1965). This type of accent system, so-called (among other things) "expiratory accent" or "accent of intensity," was held to be different from the type of accent found in such languages as Swedish, Serbo-Croatian, and Chinese, which was called "pitch accent". The latter type was supposedly realized by distinct pitch modulations (Roudet 1910). The term "stress" was generally reserved for the first type of accent system. It was also allowed that stressed syllables were generally longer than unstressed syllables (Parmenter and Treviño 1935, Muysken 1931).

There were, however, a few who thought a pitch change was an essential element of "stress"; Stetson (1923) cites Mitford (1804), Coleman (1914), Abas (1925), and Morris (1925). Stetson himself, while allowing that there might be a pitch change on accented syllables, thought that even so it would be a secondary effect of the increased expiratory force since

... the heavy stroke of the accent involves the chest pressure and is apt to change the pitch because the laryngeal musculature is often affected by tensions in the other musculatures of speech (141).
However, he cites no evidence supporting this claim (of the special sensitivity of the laryngeal muscles to other muscles' contractions) and I know of none even today.

Some of the earliest objective analyses of speech supported the minority view on the role of pitch in the production of stressed syllables. Muyskens (1931) and Scott (1939), for example, found significant pitch variations on stressed syllables and the latter even claimed:

...there seems to be a strong indication that stress [i.e., intensity] unaided, is not very efficient as a distinguishing feature of English...

Some influential research on the physiological correlates of stress was that conducted by Stetson (1928) who presented evidence (e.g., records of oral and subglottal air pressure and chest movements) that there were separate expiratory pulses ('breath pulses') for each spoken syllable as well as extra heavy pulses for stressed syllables. Many of Stetson's recording techniques were rather primitive and some of his experimental techniques problematic (for example, the fact that many of his syllables showing discrete breath pulses also happened to be almost isolated utterances) and many of his findings were not replicated by Ladefoged (1962, 1967) and his colleagues or by Liberman, Griffiths, Mead, and Knudson (1967), both using more modern techniques. Nevertheless, Ladefoged did find a momentary increase in the activity of the expiratory muscles (sampled via electromyography) and in the subglottal air pressure during and sometimes immediately before a stressed syllable, whether emphatically stressed or not. He calibrated the effect that variations in subglottal pressure, $P_s^2$, would have on both intensity (Ladefoged and McKinney 1963) and $F_o$ (Ladefoged 1953). The $P_s$ was found to be directly related to the major variations in intensity of voice (independent of those due to the resonance characteristics of the vocal tract). However, the major part of the $F_o$ variations including those on stressed syllables, were accomplished by the larynx via changes in the tension of the vocal cords.

RESULTS OF MODERN RESEARCH

As a prelude to the discussion of subsequent work on the physiology of stress it may be appropriate to mention some of the extensive experimental data on the perceptual correlates of stress. Using such techniques as synthetic speech, Bolinger (1958), Lawrence (1953), and Fry (1955, 1958, 1960, 1965) found that in English there was actually a hierarchy of cues that listeners used to identify stress on a syllable; listed in the approximate order of greater to lesser importance: pitch modulation, duration, intensity, segmental quality (including especially vowel quality, but probably also such things as degree of aspiration, etc.). As for pitch modulation, it was shown that it was changing pitch, not necessarily higher pitch, which served most effectively to signal stress.
Acoustic analysis of natural speech (again in English) generally reinforced the findings from these perceptual studies (e.g., Bolinger 1958, Lieberman 1960, Lehiste and Peterson 1961, Morton and Jassem 1965, Lisker and Abramson 1967). Generally similar results were obtained from other languages identified as having stress: Polish (Jassem, Morton, and Stefan-Batog 1968), German (Isacenko and Schadlich 1966), Czech (Janota and Liljencrans 1969, Ondráčková 1972, Janota and Ondráčková 1975), Dutch (van Katwijk and Govaert 1967, van Katwijk 1969). (It should be acknowledged, however, that quite different results from the above have also been obtained on occasion. Fonagy (1958), for example, reports that in his studies of Hungarian stress increased respiratory muscular activity is a more reliable correlate of stress than variations in intensity, Fo, or duration.)

Perhaps the ultimate verification of the correctness of the above findings, especially the primary importance of pitch and duration as the essential cues for stress, has been the fact that highly successful synthetic speech has been generated by rules based on these findings (Mattingly 1966, 1968, Rabiner 1968, Matsui, Suzuki, Umeda, Omura 1968, Umeda 1976).

Given these results in the perceptual domain, it is not surprising that the more recent physiological investigations of stress have confirmed that all the characteristics of the speech signal shown to be perceptually important are accomplished actively by the articulators in the vocal tract (and the muscles which serve them) which are responsible for producing those features. For example, cine x-ray studies show greater extent and force of articulator movement (and consequently greater duration) when stressed as opposed to unstressed syllables are produced (Broad 1968, Harris, Gay, Scholes, and Lieberman 1968, Kent and Netsell 1972). This is, however, mainly true of what would be classified as emphatically stressed syllables, less so of normal, non-emphatic, non-contrastive, word stress.

F<sub>0</sub> VARIATIONS IN STRESS: IS THE LARYNX OR LUNGS RESPONSIBLE?

The only moderately controversial topic in the whole area of the physiology of stress is that surrounding the question of how the pitch variations associated with stress are controlled, specifically, whether by the laryngeal or the pulmonic systems. This controversy was precipitated by the publication in 1967 of the M.I.T. doctoral dissertation of Philip Lieberman in which it was claimed that except for the terminal pitch rise in yes-no questions, the pitch variations in speech were regulated primarily by the subglottal air pressure which, in turn, was controlled by the pulmonic system.\textsuperscript{3}

The basis for his claim was recordings (from three speakers) of the acoustic speech signal and Ps during a variety of utterances. A sample of his data is given in Figure 1. In this data he noted that there was generally a close temporal coincidence between a momentary F<sub>0</sub> rise on stressed syllables (marked 'A' in Figure 1) and a momentary increase in Ps (marked 'A'\textsuperscript{1}) and likewise the fall in F<sub>0</sub> at the end of declarative sentences and the fall in the Ps (not manifested in this figure).
Figure 1. Sample of data obtained by Lieberman (1967). Top: Fundamental frequency (Fo); bottom: subglottal pressure (Ps). The data shows the close temporal coincidence of the momentary rises in Fo (A) and Ps (A').
Such observations, of course, had been made before by others who recorded Ps variations in speech, e.g., Smith 1944.) Lieberman concluded from this that the Ps changes caused the Fo variations; he assumed the laryngeal muscles played no essential part in varying pitch in these cases. He measured corresponding values of Fo and Ps at points in the data where he assumed the laryngeal muscles were inactive (i.e., in declarative sentences) and arrived at "calibrations" of the effect of Ps on Fo that ranged from 16 to 22 Hz/cm H2O.4

The crucial assumption in Lieberman's claims, that the laryngeal muscles maintain a constant level of activity except during the Fo rises of questions, was quite unsupported by independent evidence. Of course, at that time neither Lieberman nor anyone else had directly measured the level of activity of the laryngeal muscles during connected speech. Nevertheless up to that time almost all phoneticians and speech scientists held the view opposite to Lieberman's, viz., that the Fo of voice during speech is in all cases primarily regulated by the laryngeal muscles and that the effect of Ps on Fo was too small to account for the major part of the observed Fo variations, including those on stressed syllables (Sweet 1877, Scripture 1902, Staton 1928, Pressman and Kelemen 1955, Ladefoged 1963, Fry 1964, Öhman and Lindqvist 1966, Zemlin 1968, Proctor 1968). There were several reasons for this dominant view:

1. It was commonly noted that the larynx moved up and down in the neck during the Fo changes in speech and it was therefore assumed that (somehow or other) the larynx contributed to these changes (Herries 1773, Scripture 1902, Critchley and Kubik 1925).

2. It had been a common clinical observation that paralysis of the intrinsic (or some extrinsic) laryngeal muscles frequently led to defects in the control of pitch in speech (Critchley and Kubik 1925, Luchsinger and Arnold 1965, Arnold 1961, Sonninen 1956). On the other hand, respiratory paralysis did not lead to any difficulty in pitch regulation (Peterson 1958).

3. Direct electromyographic recordings of the activity of the laryngeal muscles during steady-state phonation or during singing showed the intrinsic and extrinsic laryngeal muscles, especially the cricothyroid, to be very active during Fo variations (Katsuki 1950, Faaborg-Andersen 1957, 1965, Sawashima, Sato, Funasuka, and Totsuka 1958, Arnold 1961, Kimura 1961, Hirano, Koike, and von Leden 1967). It was therefore assumed (explicitly so by Arnold) that these results could be extrapolated to speech conditions.

4. Calibrations of the effect of Ps variations on Fo, where care was taken to insure that the vocal cords maintained a constant level of tension, were done on excised larynges as early as the first half of the 19th century (Mueller 1851) and on intact speakers by Ishikii (1959), Ladefoged (1963), and Öhman and Lindqvist (1966). In the studies involving the living subjects the procedures involved having the subject produce a steady-state vowel at a constant pitch while receiving slight pushes on the chest at unexpected moments.
These pushes produced brief involuntary increases in $P_s$ which in turn produced brief increases in $F_o$. In general such studies yielded ratios of $\Delta F_o/\Delta P_s$ in the range 2 to 5 Hz/cm H$_2$O in the pitch range used in speech. Ladefoged and Öhman and Lindqvist then applied these values to the analysis to their own records of $F_o$ and $P_s$ during connected speech (i.e., working with the same kind of data Lieberman did) and factored out that part of the $F_o$ contour that could be attributed to variations in $P_s$. From this the major part of the observed $F_o$ variations in speech had to be attributed to the action of the laryngeal muscles. Öhman and Lindqvist cited additional reasons for their conclusions:

...The $\Delta P$ change [i.e., the change in the pressure drop across the glottis] which is due to stress is always much smaller than that due to stop consonants for instance, and ...the $F_o$ changes during the stressed syllables do not correlate well with the stress-induced $\Delta P$ changes either in phase or in amplitude.

(This latter point, in fact, also applies to the data in Fig. 1).

Lieberman failed to note most of this evidence and made no attempt to reconcile his claims with the part of it he did review except to dismiss Öhman and Lindqvist's results as applying to singing not to speech.

Lieberman's hypotheses were quickly disproved. New evidence was obtained as well as refinements of the existing counterevidence which was known before the publication of Lieberman's provocative claims. These took several forms:

1. Vanderslice (1967) recorded the vertical movements of the larynx and $P_s$ during connected speech and showed the former to be in better synchronization with $F_o$ than was $P_s$. It is reasonable to conclude, then, that the larynx, at least in part by its vertical movements, actively participates in the control of $F_o$. It should be mentioned, though, that the larynx-hyoid apparatus (unlike the lower mandible, for instance) is quite complex as regards its muscular structure and attachments to the skeletal frame and accordingly has many degrees of freedom in its movements (Ohala and Hirose 1969, Ohala 1972). It is not unexpected then that there should be some inter- or even intra-subject variability as regards use of gross vertical movements of the larynx in $F_o$ regulation. Thus, although there may be some speakers for which the vertical movements of the larynx do not show clear correlation with $F_o$ during speech (Gandour and Maddieson 1976), this in no way undermines the conclusion that this mechanism is used by many if not most speakers during speech, especially for large variations in $F_o$ such as occur on stressed syllables (Ohala 1972, Ewan and Krones 1974, Shipp 1975a, b, Ewan 1976).
2. Additional calibrations of the effect of Ps on Fo were done using refined techniques (including monitoring laryngeal muscle activity during the induced transglottal pressure changes) and studying a greater range of voice qualities, initial Fo levels, and intensities (Ohala and Ladefoged 1969, Ohala 1970, Lieberman, Knudson, and Mead 1969, Hixon, Mead, and Klatt 1971). A summary of the values of $\Delta$Fo/$\Delta$Ps from these and earlier studies are given in Table 1. In general, the studies using excised larynges yield higher values for this ratio than do the studies involving intact larynges, Lieberman's (1967) study excepted. This may very well be an artefact due to the excised larynx lacking normal muscle tonus (van den Berg and Tan 1959). Lieberman, Knudson, and Mead's (1969) upper limit for normal voice of 10 Hz/cm H$_2$O is still on the high side although it is substantially lower than the values claimed by Lieberman (1967). Curiously, Hixon, Mead, and Klatt (1971), using the same technique and one of the same subjects involved in the study of Lieberman et al. (1969), could not replicate those authors' findings.

There are, of course, methodological problems with all these calibration techniques. Nevertheless, there is little evidence to support Lieberman's claims that the Ps could be responsible for all or even most of the observed Fo variations on stressed syllables. This conclusion is not surprising since Lieberman made no serious attempt to insure that the tension of the vocal cords (as affected by the laryngeal muscles' activity) remained constant during the portions he computed $\Delta$Fo/$\Delta$Ps. 5

3. It was pointed out that Ps variations can be caused in part by changes in glottal and oral Impedance, i.e., by anything which would cause reduced air flow, e.g., obstruent closure, reduced mean glottal area (such as happens during increases in Fo (Sonesson 1960, Ishizaka and Flanagan 1972)) or increased percentage of closed time in the glottal area function (such as occurs during increases in voice intensity) (Isshiki 1969, Ohala 1970, 1975a, 1975b, van Katwijk 1971, 1974). This is not a controversial point; it is well known among speech scientists (Stetson 1928, Peterson 1957, Isshiki 1961, 1964, Campbell, Murtagh, and Raber 1963, Yanagihara and von Leden 1966, Öhman and Lindqvist 1966, Ladefoged 1968, Koyama, Kawasaki, and Ogura 1969, Rothenberg 1968, Zemlin 1968, Benguerel 1970, Netsell 1973, Lofqvist 1975).

A related fact which undoubtedly reflects the increasing impedance of the glottis with increasing Fo is that the minimum pressure drop necessary to maintain voicing is greater for high Fo than low Fo (Mueller 1851, Isshiki 1959).

Indirect evidence of increased glottal resistance during stressed syllables (in comparison to unstressed syllables) would be the finding of less oral air flow on stressed syllables (Ohala 1975a, 1975b). Such reductions in
Table 1. Values of $\Delta F_o/\Delta P_s$ (in Hz/cm H$_2$O) from various studies.

<table>
<thead>
<tr>
<th>Source</th>
<th>Normal voice</th>
<th>High pitch &amp; Falsetto</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mueller (1851)*</td>
<td>4.3 - 4.5</td>
<td>10 - 16</td>
</tr>
<tr>
<td>Isshiki (1959)</td>
<td>3.3</td>
<td></td>
</tr>
<tr>
<td>van den Berg and Tan (1959)*</td>
<td>5 - 13</td>
<td>17 - 20</td>
</tr>
<tr>
<td>Ladefoged (1963)</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Öhman and Lindqvist (1966)</td>
<td>2.5</td>
<td></td>
</tr>
<tr>
<td>Furukawa (1967)*</td>
<td>8**</td>
<td></td>
</tr>
<tr>
<td>Anthony (1968)*</td>
<td>6 - 8</td>
<td></td>
</tr>
<tr>
<td>Lieberman, Knudson, and Mead (1969)</td>
<td>3 - 10</td>
<td>9 - 18</td>
</tr>
<tr>
<td>Ohala and Ladefoged (1969)</td>
<td>2 - 4</td>
<td>7 - 10</td>
</tr>
<tr>
<td>Ohala (1970)</td>
<td>2 - 8</td>
<td></td>
</tr>
<tr>
<td>Hixon, Mead, and Klatt (1971)</td>
<td>2 - 4</td>
<td></td>
</tr>
<tr>
<td>Lieberman (1967)</td>
<td>16 - 22</td>
<td></td>
</tr>
</tbody>
</table>

* Used excised larynges.

** Average slope of one $F_o$ vs. $P_s$ plot which ranged from 0 to 16 Hz/cm H$_2$O.
air flow, although probably not to be found on all stressed syllables, are evident in some of the published data of Klatt, Stevens, and Mead (1988; see their figure 5) and Broad (1968).

Thus not only is it improbable that the observed Ps variations could cause much of the observed Fo changes, it is probable that to some extent the Ps fluctuations were themselves caused by the laryngeal muscles as they adjusted the state of the vocal cords for Fo and intensity variations.

4. Finally, and most conclusively, the activity of the laryngeal muscles was sampled directly during connected speech using electromyography (Ohala and Hirano 1967, Hirano, Ohala and Vennard 1969, Hirano and Ohala 1969, Fromkin and Ohala 1968, Ohala 1970, 1972, Lieberman, Sawashima, Harris, and Gay 1970 Atkinson 1973, Netsell 1973, Collier 1975, Maeda 1975, Shipp 1975a, Erickson and Atkinson 1975, Kakita and Hikl 1975) and it was found that Fo was varied in speech in much the same way as it was in singing. There were muscles active for raising pitch: the cricothyroid, thyrohyoid, lateral crico-arytenoid, and vocalls; and muscles active during lowering of pitch: the sternohyoid and sternothyroid. (The involvement of the strap muscles in Fo change is not surprising given the evidence reviewed above that larynx height varies with Fo.)

Typical data is shown in Figs. 2 and 3 (from Ohala 1970). In Fig. 2 the records of Fo, Ps, activity of the cricothyroid and lateral crico-arytenoid, and the voice signal are given for three utterances. From such data it is evident that although there is a noticeable increase in Ps on stressed syllables (compare the middle sentence with the first in Fig. 2), there is also a considerable increase in the activity of the cricothyroid muscle at the same time. Fig. 3 shows an increase in the activity of the sternohyoid muscle during the lowering of Fo. (It should be mentioned, however, that the sternohyoid also participates in purely segmental gestures as well (Ohala and Hirose 1969, Ohala 1972) and may in some subjects only reveal its involvement in Fo lowering in the indirect way of not being active during high Fo.)

The increase in the activity of the cricothyroid muscles during the stressed syllable in Fig. 2 (middle sentence), where Lieberman would have predicted the change in laryngeal muscle activity to be nil, is of about the same magnitude as that during the final pitch rise in the question (right-most sentence), where the involvement of the laryngeal muscles has never been in question.

To summarize the evidence reviewed: the magnitude of the effect of Ps on Fo is not great enough to account for the major part of the observed Fo changes on stressed
Figure 2. From top: fundamental frequency (F0), subglottal pressure (Ps), electromyographic signal from the cricothyroid muscle, electromyographic signal from the lateral crico-arytenoid muscle, and microphone signal of voice. The figure shows how these parameters vary during the three utterances (from left to right) "Bev bombed Bob," "BEV bombed Bob," and "Did Dev bomb Bob?" (nonsense sentences constructed to have mostly labial and voiced consonants and non-high vowels).
Figure 3. From top: fundamental frequency (Fo), electromyographic signal from the cricothyroid muscle, electromyographic signal from the sternohyoid muscle, and microphone signal of voice. The utterances are: "Mom bombed Bob" (on the left) and "Mom bombed Bob?" (on the right). As the two utterances are phonetically identical except for the Fo contour, the observed changes in the muscle activity from one utterance to the other must be due to these muscles' involvement in Fo regulation, not segmental gestures. The cricothyroid shows increased activity during increases of Fo, no activity during lowering of Fo; conversely, the sternohyoid shows, in addition to participation in segmental gestures (jaw opening for labial consonant + low vowel sequences), increased activity for Fo lowering, lessened activity for Fo rises.
vowels (nor the fall in Fo at the end of declarative sentences); some of the Ps variations, in fact, are probably dependent upon the action of the larynx itself, not the pulmonic system; there is indirect and direct evidence that the laryngeal muscles cause the Fo variations in speech. The answer to the question, so epigrammatically phrased by Vanderslice (1967), 'is it the larynx or lungs that controls pitch?', is the larynx.

THE PULMONIC CONTRIBUTION TO STRESS

Even though it is clear now that any pulmonic contribution to the Fo change observed on stressed syllables must be small this does not rule out the possibility of a pulmonic contribution to other aspects of stress, in particular, increases in intensity. Ladefoged (1963, 1967, 1968) citing his own extensive electromyographic investigations of the activity of the expiratory muscles during speech, maintains that there is a momentary increase in expiratory activity during or slightly preceding each stressed syllable, whether emphatically stressed or not. Van Katwijk (1974), however, obtained surface electromyographic recordings of the activity of some of the expiratory muscles during speech (of two native speakers of Dutch) and found appreciable increases in expiratory muscular activity only on emphatically stressed syllables, not on conversationally stressed syllables.

It is difficult to evaluate such electromyographic records by themselves, however, for two reasons: First, there is no simple way of estimating the magnitude of lung volume decrement such bursts of expiratory muscle activity cause and, after all, it is only by decreasing the lung volume at a faster-than-normal rate that these muscles can have any influence on Ps. It is possible these fluctuations in the level of expiratory muscle activity produce only negligible variations in Ps. Second, it is difficult to know why the expiratory muscles show increased activity near stressed syllables. Aside from the pulmonic system's possible role in the production of individual speech segments or stressed syllables, we do not know whether its long-term function is to maintain a constant pressure in the lungs or to maintain a constant rate of lung volume decrement, or some combination of the two. If its task is to maintain a steady rate of lung volume change, then the momentary increase in expiratory activity on stressed syllables could simply be a compensatory reaction to the increased glottal and supraglottal resistance to air flow expected during the production of stress—which increased resistance would momentarily slow the rate of lung volume decrease. If this were so, the increased activity of the pulmonic muscles would not, strictly speaking, be an independent feature of stress, it would be a feature dependent upon glottal and supraglottal activity.

It would be helpful to obtain some independent measure of the pulmonic system's contribution to Ps changes during speech. This can be done most directly by recording variations in lung volume during speech. Such records can be obtained using a plethysmograph such as has recently been constructed in the Phonology Laboratory at Berkeley.
Ours is a whole-body pressure plethysmograph. It consists of 1 and 2/3 50-gal. oil drums welded together and tightly fitted with a plexiglass lid. The subject stands in the plethysmograph and breathes and speaks through a face mask mounted in the side of the lid (see Fig. 4) which vents to the atmosphere. As the flange of the face mask creates an air-tight seal around the subject’s nose and mouth and as the rest of the drum is air-tight, there is a fixed mass of air inside the plethysmograph. Changes in the lung volume during breathing and speaking (see chest profile in solid and broken lines in Fig. 4) therefore cause corresponding changes in the air pressure inside the plethysmograph. These pressure changes are sampled continuously through a port in the side of the plethysmograph, transduced by a sensitive pressure transducer, and recorded on tape along with the voice signal for later analysis.

Preliminary studies have been done with three speakers (including myself): 2 English speakers and 1 Swedish speaker. Using speech samples that would eliminate or minimize jaw movement (in order that the only volume changes recorded by the plethysmograph would be those of the chest wall), the following results have been obtained (illustrated in Figs. 5 and 6):

1. There are relatively large rapid decreases in lung volume during moments of high air flow, e.g., during aspiration, [h], and fricatives. See Figs. 5b, c, and d. These presumably represent a passive collapse of the lungs due to the rapid flow of air out of the lungs and the consequent decrease in lung pressure (= Ps). This contrasts with the case of sonorants (nasals, laterals, etc.) where the rate of lung volume decrement is the same as that for surrounding vowels. See Fig. 5a.

2. There were also moments of lesser-than-normal lung volume decrement during periods of reduced air flow, i.e., during the closure phase of some stops. See arrows in Fig. 6a and b.

3. There were momentary greater-than-normal decreases in lung volume during emphatically stress syllables. See portion delimited by broken vertical lines in Fig. 6b.

4. There need not be any obvious change in the rate of lung volume decrement on non-emphatically stressed syllables. This is easiest to see in sentences consisting entirely of sonorants since the segments themselves will cause no perturbation in the rate of air flow out of the lungs. See portion delimited by broken vertical lines in Fig. 6a and also note lack of change on stressed vowel in Fig. 5a.

In general, these results are completely compatible with those obtained by van Katwijk (1974) in his electromyographic study of the expiratory muscle activity in the speech of two Dutch speakers.
**Figure 4.** Schematic representation of the pressure plethysmograph. The subject breathes and speaks through a face mask which makes an air-tight seal around the nose and mouth. The pressure of the air trapped inside the drum varies directly with changes in chest volume. The pressure is sampled via a port, transduced by a sensitive pressure transducer, and recorded.
Figure 5. Representative data obtained using the plethysmograph. 5a through d show the lung volume (bottom) along with the voice signal (top) for the consonants [m, s, h, tʰ], respectively. Each of the consonants, whose onset and offset are indicated by broken vertical lines, were spoken by a male speaker of American English in the frame "deem _oon real" [dim'_unril]. (Subject LB.)
Figure 6. Representative data obtained using the plethysmograph. a and b show the lung volume pattern for the utterances "dom unilluminable real" [dimunihbliril], spoken conversationally on the left and with emphatic stress on the right. The broken vertical lines delimit the stressed syllable in each utterance. The arrows mark instances of lesser-than-normal lung volume decrement during stops. (Subject J.B.)
Of course, the number of subjects studied so far with this technique is quite small, and it is possible that a simple visual analysis of the plethysmograph output is not capable of revealing small but consistent changes in the slope of the lung volume curve (computer averaging of the signals is planned for the near future). Therefore these results should be viewed cautiously. Nevertheless, the data obtained so far do not support the notion that there must always be an appreciable expiratory pulse accompanying the production of ordinary (non-emphatic) stressed syllables. For emphatic stress, however, the involvement of the respiratory system is quite apparent.

ACKNOWLEDGEMENT

This research was supported by grants from the National Science Foundation and from the Holbrook Fund, University of California, Berkeley.

FOOTNOTES

1 I use 'pitch' and 'Fo' (fundamental frequency) interchangeably in this paper.

2 Actually, the transglottal pressure drop which, during vowels, is usually equal to Ps.

3 In fact, Lieberman's claims were stated with so many qualifications and occasionally in such self-contradictory ways, that it may be seriously questioned whether they represented empirical (i.e., potentially falsifiable) hypotheses at all. In this paper I have attempted to interpret his hypotheses as if they were serious scientific statements. On this point, cf. also the detailed criticisms in Vanderslice 1968, 1970, Ohala and Ladefoged 1969, Ohala 1970.

4 Even aside from what later investigations have revealed (see below), the significance and reliability of these calibrations is questionable due to the confusion over how they were obtained. As noted in Ohala and Ladefoged (1969) Lieberman claimed to have measured the corresponding values of Fo and Ps at points in the data curves where it was assumed that the tension of the laryngeal muscles was unchanged, so that the fundamental frequency was a function of the subglottal air pressure [Lieberman 1967:95-6].

Nevertheless, a few pages later he remarks:

The points in Figure 4.35 [the data in question] have a fair amount of horizontal dispersion, which indicates that the laryngeal tension is not always constant throughout the non-terminal portion of each breath-group [102, emphasis added].
And a few pages later:

Quantitative calculations of the relationship between fundamental frequency and the subglottal air pressures were made... These calculations showed that the tension of the laryngeal muscles was relatively constant during the non-terminal portions of the breath-group [107, emphasis added].

As Ohala and Ladefoged commented:

This kind of self-contradiction abounds in Lieberman's work; and it makes it very difficult for anybody else to contradict him, since he can always claim that he has stated the opposite to what is attributed to him... Besides the contradictions between these quotes there is also the rather strange procedure of first assuming X (i.e., the inactivity of the laryngeal muscles) as part of the basis of Y (i.e., the calculated $\Delta F_o / \Delta P_s$ ratio), and then later taking Y as proof of X without, as far as we can see, any other independent supporting evidence.

Moreover, there is evidence that these calibrations were done carelessly and inconsistently. Though trivial by itself, this evidence is important in the context of Lieberman's strong claims since it reduces even the prima facie plausibility of some of these claims. In his book Lieberman presents measurements forming the basis of the estimates of $\Delta F_o / \Delta P_s$ in four figures, three of which give the individual data from his three subjects and one of which presents the combined measurements from all the subjects. For unexplained reasons the data in the fourth figure does not completely agree with that in the other three although it was supposedly based upon them. In addition, although Lieberman is (commendably) perfectly explicit about how he did the measurements on his raw data--all relevant bits of which he published--it turns out that a replication of these procedures fails to yield exactly the same data points as those given in his $F_o$ vs. $P_s$ plots. The discrepancies are not small; they would lead to ratios of $\Delta F_o / \Delta P_s$ as high as 28 or 33 Hz/cm H2O--absurdly high values that would have clearly revealed the assumption underlying the calibration, i.e., that the level of activity of the laryngeal muscles was constant, was erroneous.

Lieberman (personal correspondence) acknowledged the mix-up in the construction of the figures referred to but denied that any of the data in them was unjustified; that the data points I couldn't find were in those sentences which in his book he said he hadn't used for the calibration, namely, the yes-no questions, where it could not safely be assumed that the laryngeal muscles were inactive!
This is not an isolated case. His Figure 4.34, though in no way crucial to the issue under discussion, also bears no relationship to the published data it was alleged to be based on. This mix-up was also acknowledged by Lieberman (personal correspondence).

Liberman (1973) repeats this fallacy in his analysis of some of my own data (presented in Fromkin and Ohala 1968 and Ohala 1970): from records of utterances showing Fo, Ps, and the activity of the cricothyroid and the lateral crico-arytenoid muscles (Fig. 2 of this paper is a sample of that data) he estimates that Fo varies as a function of Ps at the rate of 12.5 Hz/cm H2O during periods where the two muscles recorded showed little variation in level of activity. However, Liberman chooses to ignore the fact that in Ohala 1970 other recordings of this same subject show the sternohyoid muscle active during these kinds of pitch changes (the drop in pitch at the end of declarative sentences). Moreover, it is well known that to some extent almost all of the dozen or so intrinsic and extrinsic laryngeal muscles can play a role in regulating Fo; one cannot rule out their participation either. Liberman seems to operate according to some extreme form of positivism: those factors not represented in the instrumental records he examines (even if they may be represented in other records) simply don't exist.

REFERENCES

Benguere, A. P. 1970. Some physiological aspects of stress in French. Natural Languages Studies No. 4, Phonetics Laboratory, Univ. of Michigan.
Broad, D. J. 1968. Some physiological parameters for prosodic description. Speech Communication Research Laboratories, Santa Barbara, Monograph No. 3.


Mitford 1804. Inquiry into the principle of harmony, etc. London.


Shipp, T. 1975a. Vertical laryngeal position during continuous and discrete vocal
frequency change. JSHR 18:707-718.


The Five Nations of the Iroquois—i.e., the Mohawk, Oneida, Onondaga, Cayuga, and Seneca—speak distinct but quite closely related languages (see Chafe forthcoming for a linguistic survey). Conspicuous among the differences which separate these languages are variations in the placement of the morphological accent, as well as in several other phenomena conditioned by or conditioning the accent. For example, in all the languages some accented vowels are lengthened, but the conditions for lengthening vary. All the languages show epenthetic vowels, but not always the same vowels and not always under the same conditions; the effect of epenthetic vowels on accent placement also varies. In Cayuga there is metathesis and devoicing in certain unaccented syllables. The counting of syllables is basic to accent placement in all the languages, but again in different ways. From an original pattern in which syllables were counted backward from the end of a word there developed in the western languages, particularly in Cayuga and Seneca, a new pattern in which syllables were counted forward from the beginning of a word. Relics of the earlier pattern still exist in these languages, however.

I will present here a survey of these developments, beginning with the Proto-Northern-Iroquoian (PNI) situation and then discussing each of the Five Nations languages in turn. In general the number of modifications of the original pattern increases as we move westward from Mohawk to Seneca, so that the developments are most easily described in that order. Two other Northern Iroquoian languages on which we have extensive information—Huron and Tuscarora—seem not to have taken part in the particular complex of changes which affected the Five Nations languages, and they will not be included in the discussion.

The term "accent" is meant to include both higher pitch and stronger stress (amplitude) on the syllable affected. In all the languages except Onondaga accented syllables are evidently both higher pitched and louder than other syllables. That is true for most words in Onondaga also, but under certain conditions these two features have become separated so that the syllable with highest pitch is different from the syllable with strongest stress.
The patterns I will be discussing are those morphological accent patterns which appear on words in clause-final position or in isolation. These patterns are often simplified when a word is in clause-medial position. It is the clause-final form which is most complex, and most interesting. The discussion will also be limited to the more regular patterns to be found in these languages. Almost all the statements I will make have certain exceptions, a few of which I will mention but many of which are too particular to have a place in a paper of this sort. Some of the exceptions appear to be retentions from irregularities already present in the proto-language, while others seem to have resulted from analogic changes applied to particular items. Since not all these phenomena have been thoroughly studied for all the languages, furthermore, some of what I say will have to be tentative and incomplete.

Proto-Northern Iroquoian

For the most part, clause-final words in reconstructed PNI were accented on the penultimate syllable. Each vowel constituted a syllabic nucleus, so that it would be equally correct to say that the accent was on the penultimate vowel. Examples are *ohra'ta? 'feather', *yakwayethwas 'we plant'.

Because of morphological constraints, all nouns in PNI would naturally have contained more than one syllable, so that a penult was always available for accenting. The morphology of verbs was such, however, that occasionally it would have been possible for a verb to consist of only one syllable if nothing except its contentful morphemes were included in it. In such a case an empty syllable i- was added at the beginning of the verb. The only function of this syllable was to produce a word that consisted of more than one syllable, so that a penult would be available to take the accent. Thus, instead of *twak 'let's eat it' (consisting of the pronominal prefix twa- 'inclusive plural' plus the root -k 'eat'), the word was *(twak. It is possible that this "empty" initial i- has been retained from a past as remote as Proto-Iroquoian-Caddoan, since an i- with a somewhat similar function appears in Caddo (Chafe 1973: 1196-1197).

In PNI it was usually the case that when a noun root ending in a consonant was followed by another morpheme beginning with a consonant, an intrusive vowel a was added to separate the two consonants. If this "stem-joiner" vowel formed the nucleus of the penultimate syllable, it was not accented, but the accent was placed on the antepenult. In other words, this vowel was "weaker" than other vowels in the sense that it was unable to assume the penultimate accent, although it was counted in determining what other syllable was the penult. In the reconstructed word *kanōhsoako 'in the house', for example, the second a joins the noun root -nohs- 'house' to the locative suffix -k0, and it passes the accent back to the antepenult.
If the accented syllable was an "open" one, the vowel was lengthened. In general, "open" meant that there was only one consonant intervening between the accented vowel and the vowel following, as in *kahse-naʔ 'name' or *kanə-takoʔ 'in the town'. In the second example the penultimate syllable contains the stem-joiner vowel, and is thus not eligible for accenting. The two so-called laryngeal consonants, ʔ and h, functioned to close a syllable even when they were not in clusters. Thus we find no lengthening in *hrotəh 'he's asleep' or *kanəheʔ 'seed'.

In the extant languages there are sporadic cases in which lengthening does occur nevertheless before a single ʔ or h, as in Cayuga owivəjäh 'infant' or Onondaga kahnya-ąhaʔ 'it barks'. Some such cases were probably already present in PNI, and some of them can probably be attributed to analogic extension of lengthening to accented vowels occurring before single consonants in general, and not just non-laryngeal consonants.

Another systematic exception to the lengthening process as stated involves the cluster kw. When both elements of this cluster occurred within a single morpheme (when there was no morpheme boundary between the k and the w), the cluster functioned as if it was a single consonant and did not close the preceding syllable. For that reason we find lengthening of the accented vowel in *hr3`.k-weh 'person', for example. It has been suggested by Lounsbury and others that these cases of kw stem from an earlier p, hence from what was originally a single consonant (Postal 1969:245-254).

We can now look at the modifications of the PNI accent pattern which are found in each of the Five Nations languages. My information on Mohawk has come largely from Bonvillain (1973), Michelson (1973), and Roy A. Wright (personal communications). Information on Oneida is from Lounsbury (1942, 1953). For Onondaga I have relied partly on my work (Chafe 1970a) and partly on that of Woodbury (1975 and personal communications). Cayuga data have been provided by Michael K. Foster (1974 and personal communications). Information on Seneca is entirely from my own work (Chafe 1967).

Mohawk

For the most part, Mohawk has preserved the accent and vowel length patterns of PNI. Two special developments are worth noting, however: one involving epenthetic vowels and the other the emergence of a falling pitch accent.

In Mohawk an epenthetic e was inserted within certain consonant clusters. Among them were clusters ending in one of the "resonant" consonants n, r, or w. (The fourth resonant, y, did not produce this epenthesis.) If the resonant was preceded by either an "oral obstruent" (t, k, or s) or by another resonant, the epenthetic e was introduced. However, it was not introduced into kw unless this cluster contained a morpheme boundary (cf. Postal loc. cit.; Chafe 1970b: 121-123). An epenthetic e was also introduced in Mohawk before a final ʔ in order to separate the ʔ from an immediately preceding consonant.
Of interest here is the fact that these epenthetic vowels did not assume the accent themselves if they were in penultimate position, nor were they counted (as the stem-jointer a was counted) in the determination of which syllable was the penult.

In other words, the Mohawk accent has remained on the syllable where it was located in PNI, although this syllable is often no longer penultimate in Mohawk. Examples in which the penultimate vowel is epenthetic are ıkę́ ki 'two' (*tékni) and wä̱keras 'it smells' (*wa̱kras). In ḋekhren 'I will cut it' (*ekhren?) it is the vowel in the final syllable which is epenthetic, and which is therefore not counted in the determination of where the accent falls. (Mohawk usually drops final ?.)

If the PNI accent was on a vowel immediately followed by ?a, in Mohawk such a vowel was lengthened and pronounced with a falling pitch, which I will show with a circumflex accent, as in raksâʔa 'boy' (*hraksâʔah). If, furthermore, the ?a was followed by another consonant, it was itself optionally lost, leaving behind the falling pitch and length as a clue to its former presence. Thus, from *ohrâʔta? 'feather' we have Mohawk ohrâ̊ta, although speakers will sometimes preserve the ?, saying ohrâ̊ʔta. Similarly, from *kâ̊terʔ 'I dwell' we have kâ̊̃terâ̊, but sometimes also kâ̊̃terâ̊.

A similar development took place with posttonic h, but only when the h was followed by a resonant (n, r, w, or y). That is, under these conditions the vowel acquired length and a falling pitch, and the h was (always) lost: oʔwâ̊́râ̊ 'meat' (*oʔwâ̊hroʔ).

Oneida

Oneida shares with Mohawk the retention of the original accent and the development of the falling pitch accent where there was originally a posttonic ? or h. It has epenthetic e in fewer environments than Mohawk, but where it does show this epenthesis it is like Mohawk in ignoring this e both in counting syllables and in accent placement. There are, however, two accent-related phenomena which are peculiar to Oneida. The most important of these has traditionally been called the "Oneida accent shift", although Lounsbury's present view (personal communication) is that it is better regarded as a repetition or copying of the accent onto the syllable immediately following that which was originally accented. This extension of the accent to the originally posttonic syllable took place when either no consonant or a single oral obstruent (t, k, s, or c) intervened between the two syllables; PNI *cyâʔalk 'seven' became Oneida cyâ̊̃alk, and *kanâ̊-tấkâ̊ 'in the town' became kanâ̊-tấkâ̊.

(This second example shows another characteristic feature of Oneida: the devoicing of clause-final vowels under certain conditions (Lounsbury 1942: 1953: 33-34).) If the newly accented vowel was word-final, length on the preceding vowel was lost, as in ḍókwe̊̊̊ 'person' (*kó̊̊̊-kwế̊̊̊). This last example shows again that kw, when it was internal to a morpheme, behaved as a single obstruent, since accent-copying took place across it.
A minor change associated with the accent in Oneida was the replacement of ? by h in post-tonic syllables, as in the change from PNI *wahateswate? 'he played' to Oneida wahateswahte?. Lounsbury notes that this word fell together in Oneida with the word wahateswahe? 'he smelled it', in which the h was original.

Onondaga

Some of the most interesting changes in accent placement and associated lengthening have taken place in Onondaga. As of this writing not all of them have been systematically explained, and the following is to some extent preliminary and speculative. It represents, nevertheless, a reasonable synthesis of the present state of our knowledge.

It will be recalled that in PNI the stem-joiner vowel a was not eligible for accenting, and that if it appeared in a penultimate syllable it passed the accent along to an earlier syllable. Thus we can reconstruct a PNI form *skany&tah 'on the other side of the lake', in which the accent is on the antepenult because the penultimate vowel is the stem-joiner. Onondaga has added something new to such words. As noted first by Woodbury, Onondaga has retained the original antepenultimate accent as a high pitch on that syllable (and has retained the length which was present when the accented syllable was an open one). It has, however, moved the strongest stress in the word to the penult, and also lengthened the penult when it was open. I will mark pitch and stress in Onondaga words in the following way. The acute accent mark will indicate highest pitch, the grave accent mark second highest; both indicate pitches higher than any unmarked syllables. It will then be understood that the last marked syllable in a word has strongest stress, regardless of whether it is marked acute or grave. Thus, the Onondaga reflex of the word cited above can be written skany&tah. Pitch is highest on the second syllable and next highest on the third. Stress is strongest on the third syllable.

By separating highest pitch from strongest stress, Onondaga has been able to have its cake and eat it too. From the point of view of stress, it has generalized penultimate stress to all words, eliminating the exceptions (due to stem-joiner and epenthetic vowels) that were present in PNI. But by keeping the highest pitch where it always was, it has been able to preserve allegiance to the PNI accent as well. This separation, of course, took place only when the PNI accent was on a prepenultimate syllable; in words where there was a penultimate accent to begin with, Onondaga simply retains the original pattern intact.

The Onondaga generalization of stress to all penults has applied not only to words in which the original accent preceded the penult because the latter contained the stem-joiner vowel a. It has also applied in those cases where an antepenultimate accent was produced through the development of an epenthetic vowel. Epenthetic
vowels arose in Onondaga under fewer conditions than in Mohawk, and there was
greater variety in the specific vowel which appeared. It is interesting, for example,
to compare the Mohawk and Onondaga reflexes of PNI *waʔkahroʔ? 'I hear/understand
it'. In Mohawk an epenthetic e was inserted before the final ʔ, but this e is ignored
in the determination of which syllable is the penult. Thus the Mohawk word is waʔkahroke.
(Note the replacement of h by vowel length and the falling pitch accent.) The cognate
word in Onondaga is waʔkahëkeʔ or waʔkahëkëʔ. (The consonant r first changed a
following ʔ to ʔ and then was lost in Onondaga.) Where Mohawk introduced an epenthetic
e, Onondaga has introduced either a or e, the former being the more commonly used
epenthetic vowel in this environment, and the latter evidently the result of vowel
harmony. As with the stem-joiner vowel, Onondaga treats these epenthetic vowels
like all others in the determination of penultimate stress, although it continues to
ignore them in the assignment of highest pitch.

Onondaga speakers are not always consistent in using the pattern just described.
In fact, a change may be taking place in the language whereby highest pitch is being
transferred to the syllable with strongest stress in words where the two became
separated. This is a variable feature of current Onondaga speech which might eventually
lead to a reuniting of pitch and stress on the penult, if the development were allowed
to run its course.

But in addition to words of the type ooʔtoʔtaʔ 'standing tree' which conform to
the pattern described above, Onondaga also has many words like očiʔwaʔkëh 'it's bitter,
sour', in which the stressed syllable is as high-pitched as the syllable preceding it.
In other words, there are two equally high-pitched syllables, the second of which
has the strongest stress. These words seem generally to have had an original penulti-
mate accent; that is, both the high pitch and the strong stress on the penult are retentions
from the proto-language. But there is no basis in PNI for the high pitch or the length
on the antepenult. Words of this type were discussed in Chafe 1970a:76-78 in terms of
'pretonic lengthening', but usually at least there has been pretonic pitch-raising as
well. Briefly, a vowel is treated in this way if it fulfills all of the following conditions:
(1) it is in an open syllable (separated from the following accented vowel by a single
consonant); (2) it is in an even numbered syllable, counting from the beginning of
the word; and (3) it is not a stem-joiner vowel. Furthermore, the following accented
vowel must itself be either long, or a short vowel followed by a syllable-final ʔ or h.
One further peculiarity is that in the determination of even syllables, a word-initial
consonant cluster beginning with a stop (t or k) counts as if it were a syllable. Examples
in addition to that given above are akëʔnohtsʔ 'I know it' and thāʔtyeʔs 'he's standing
there'. In the second word the pretonic vowel is counted as being in an even-numbered
syllable because of the initial cluster th; cf. hataʔtyeʔs 'he's standing', without the
cluster and without the length or high pitch.
It is possible that this pretonic lengthening (and pitch-raising) took place in Onondaga as a generalization of the pattern described earlier. The generalization of stress to all penults created words in which the stressed syllable was preceded by one in which there was high pitch and often length. It may be that subsequently the habit of lengthening and raising the pitch of syllables preceding the one with strongest stress was generalized to other words of the type just mentioned, these remaining distinguishable from the earlier type by the presence of highest pitch on their stressed syllable as well. It remains problematic, however, just why this generalization should have been restricted to the rather special environments in which it is found.

What has been said so far does not account for all the cases of vowel lengthening found in Onondaga which are associated somehow with syllable count. One other, seemingly bizarre case has been accounted for by Woodbury in the following terms. In a word consisting of five or more syllables, the second vowel is lengthened if it is followed by a cluster of stop or affricate plus resonant. Thus we have ḍık-ci-ko-taché-ni? 'I will find the fish'. The same explanation holds for the length in the second syllable of wá:ká-ri-ndó-tó? 'I sang', where there was earlier a cluster tr after the second vowel. (The length of the ò in this word is not because of its pretonic position, since it is in an odd numbered syllable. Rather, it is compensatory for the loss of ð.)

So far I have ignored the fact that the main stress in Onondaga sometimes falls on syllables other than the "real" penult. The best one can do is to mention two classes of what appear to be arbitrary exceptions (see Chafe 1970a:74-75). First, there are some suffixes—most notably certain alternants of certain aspect suffixes—which cause the accent to fall on the final syllable. This is true of the punctual aspect suffix with the root -ýe- 'put down', for example, as in hē:ñi-ýe-7? 'he will put it down over there' (note the pretonic lengthening and pitch-raising). The explanation for these cases is uncertain, but it seems likely that they are a retention of some feature that already existed in PNI. It is noteworthy, for example, that most such cases in Onondaga correspond to cases in Seneca where the word-final vowel is long; cf. Seneca hē:ñí-ýe-7? 'he will put it down over there'.

On the other hand, there are a few word-final morphemes, again mostly variants of aspect suffixes, whose effect is to repel the accent to the antepenult. Such is the case, for example, with the stative aspect suffix which occurs with the root -a-í- 'be in', as in kahwë-stí-tah 'the money is in it'. Again, the explanation of such cases is uncertain, and in this case I know of no corresponding phenomenon in any of the other languages.

Cayuga

In Cayuga the PNI accent pattern has been preserved provided the accent fell on an open syllable, and therefore on a syllable which was lengthened, as in ohnekí-ron 'good water' or kan-ó-ta-te? 'hill'. In the second example the accent and length are
However, the PNI pre-short, that the accent was long, that length to the original was surrendered.

The non-final syllable immediately preceded syllable, and therefore placed on the d, producing the syllable, it was syllable was in this case one.

The conditions will contain no closed syllable no accent at all. "Wild animal". The syllable, but the only by the new rule. syllable as the ends of this general, and would there--the last nonfinal in with. The Cayuga.

For it has been influenced clear. The conditions tonic lengthening evidently th (as well lengthening and the "it's burning, fire" akatë no-ta?, with penthetic a before the the final ? has
been dropped, the original accent has remained intact, and the pretonic vowel has been lengthened. The precise conditions for pretonic lengthening in Cayuga require further study.

In Mohawk and Oneida, it may be recalled, the placement of accent and vowel length was determined only by counting syllables backward from the end of a word. In general it was the second syllable from the end, the penult, which was crucial. Probably counting backward was the only process of this kind that existed in PNI. In Onondaga we noticed one phenomenon which also required counting syllables forward from the beginning of a word: pretonic lengthening takes place only on even numbered syllables counting in that direction. Now, in Cayuga, we have seen that when the accent does not fall on the penult, it is recomputed by counting forward to find the last eligible even-numbered syllable. Evidently at some point in the development of Onondaga, Cayuga, and (as we will see) Seneca, the alternation between odd and even syllables, counting from the left, became important, whereas this development never affected Mohawk and Oneida.

In this context it is interesting to observe another phenomenon in Cayuga, one which applies to all the odd-numbered syllables in a word—those syllables which were presumably in some sense phonologically weaker than the even-numbered ones. We have seen that odd syllables were not eligible for the new accent. In addition, they underwent a change which is referred to as "laryngeal metathesis". This change affected odd-numbered syllables which originally ended in ʔ or h. It is necessary to say something slightly different about each of these two consonants. So far as ʔ was concerned, it simply shifted its position around the vowel, occurring now before the vowel rather than after it. For example, from an earlier *ocihsʔtaʔ 'star' we have in Cayuga ocihsʔtaʔ. It may be noted that in the earlier form the penult was closed and therefore short, so that it lost its accent in Cayuga, the new accent appearing on the second syllable. But in addition the earlier sequence ʔʔ in the third syllable was metathesized to ʔʔ.

Laryngeal metathesis had the effect of changing the shape of various morphemes in significant ways. For example, the PNI word *waʔkoʔkaʔ ʔ 'I see you' began with the 'factual' prefix waʔ-. This prefix commonly fulfills the conditions for laryngeal metathesis, since it is more often than not in the first syllable of a word (an odd syllable), and ends in ʔ. The cluster in the form wʔa- which was produced by the metathesis was not viable, however, and the w was lost, leaving a prefix ʔa-. There is no distinction between words beginning with ʔ plus vowel and words beginning with a vowel, so that the Cayuga reflex of the word cited above is usually written αkʔkʔa.

It is interesting to consider also clusters of stop plus ʔ which resulted from this development. From PNI *kʔkwats `I'm digging it', for example, we have Cayuga kʔokwas. In most American Indian languages one might expect the resulting initial sequence kʔ to be pronounced as a glottalized stop. In the Iroquoian languages, however, no such sounds are present, and the k and the ʔ are in fact articulated separately.
Laryngeal metathesis with h rather than ? differed only in that the result is simply a voiceless vowel, rather than a sequence of h plus vowel. Thus, PNI *wakakáhnye? 'I'm playing' became Cayuga akátkAnye?, where the capital letter indicates a voiceless vowel. From PNI *kehuyasá'keh 'on my neck' we have Cayuga kEnyà'sá'kehe, with devoicing in the first syllable and metathesis of the ? in the third syllable.

For the sake of completeness, it should be added that the laryngeal metathesis did not take place when the syllable in question itself began with a laryngeal, as in Cayuga kekáha?keh 'my eye'. The ? in the third syllable remains in its original position because the syllable begins with h.

Seneca

It will be recalled that in Cayuga the original PNI accent was lost when the accented vowel was short, but retained when that vowel was long. In Seneca the original accent was lost everywhere. The length which resulted from the PNI accent was nevertheless retained. Thus, from *kahsé-na? 'name' we have Cayuga kAsé-na? with the original accent, but Seneca kahsé-no? without it. And in Seneca the pattern of accenting the last even numbered non-final syllable that is either closed or immediately followed by a closed syllable has been extended to all words in which such a syllable is present. Thus, from *wakahshé-téh 'I'm counting it' we have Cayuga akahshe-teh with the original accent, but Seneca akashe-téh with the old accent gone and a new accent on the second syllable. Seneca does not, as Cayuga does, lengthen these newly accented vowels in open syllables. Thus, while Cayuga has akIta7h 'I'm asleep', Seneca has akIta?h. In Seneca, then, the original accent pattern, based on counting from the right, has been entirely replaced by a new pattern based on counting from the left. The old pattern is still reflected, however, in the preservation of length in previously accented vowels.

Seneca has also extended lengthening to the vowels of final syllables that end in a single oral obstruent. For example from *kanôshes 'longhouse' we have in Seneca kanôshes. The reason for this development is not clear, although possibly it is related to the fact that in clause-medial position final syllables are accented. One can speculate that this accent produced a lengthening of final syllables which was generalized to clause-final words. In any case, this lengthening developed prior to the simplification of certain consonant clusters. For example, we find no lengthening in the final syllable of okto-kès 'I straightened it' because this word earlier ended in st, the final t having later been dropped.

There are a few further points which are of interest with respect to the occurrence of accent and vowel length in Seneca. We have seen that the lengthening in PNI which resulted from the accenting of an open penultimate vowel has been preserved in Seneca, although the accent itself was lost. So far as odd numbered syllables in Seneca are
concerned, the lengthened vowels correspond quite closely to those which were lengthened in PNI. That is, they are vowels which are followed by a single consonant other than ʔ or h. But, whereas in PNI the vowel a failed to be lengthened only when it was a stem-joiner, in Seneca this exception has been generalized. In odd numbered open penultimate syllables the vowel a is in general not lengthened, whether it is a stem-joiner or of any other origin. Naturally, there are exceptions: when the a is in the first syllable of a word, it is lengthened. Since the vowel in a first syllable cannot be a stem-joiner, this lengthening does reflect the PNI situation. However, as an exception to the exception, a is never lengthened in Seneca in the first syllable of an imperative verb. In summary, so far as odd numbered syllables are concerned, the pattern of lengthening in Seneca reflects that in PNI except for the lengthening of a. Whereas in PNI a failed to be lengthened only when it was a stem-joiner, in Seneca a fails to be lengthened regardless of its origin, except in initial syllables. In imperatives it fails to be lengthened even there (Chafe 1959: 491).

In even numbered syllables in Seneca the situation is somewhat different, again apparently as a reflection of the fact that even numbered syllables are phonologically stronger. Essentially, the definition of "closed syllable"—the kind of syllable which prevents penultimate lengthening—has been narrowed for even numbered syllables, so that in Seneca there are more even numbered penults in which lengthening occurs than there were in PNI. Whereas originally a closed syllable was one whose vowel was separated from the following vowel by any cluster of two or more consonants (or by a single ʔ or h), in even numbered syllables in Seneca the cluster must be of two or more oral obstruents. (ʔ and h still function to close a syllable, even when they are not in clusters.) Thus, as compared with Onondaga hatákheʔ, "he's running", in which the accented syllable is not lengthened because of the kh cluster, Seneca has hatá-kheʔ, with lengthening, since kh is not a cluster of two oral obstruents. It is the case, then, that although many penultimate lengthenings in Seneca are retentions of lengthenings caused by the original accent, there are others which are not. As in the example just given, these other cases show an extension of the lengthening pattern to a broader set of open syllables.

The distinction between "closed" and "open" syllables is important in Seneca not only in the determination of which penultimate vowels are lengthened, but also in the placement of the new accent. Here again we find that the definition of "closed syllable" has been narrowed, but not in quite the same way. The new accent, it will be recalled, occurs in Seneca (as in Cayuga) on the last even numbered non-final syllable that is either closed or immediately followed by a closed syllable. But "closed" in this case means that the vowel in question must be separated from the following vowel by ʔ, by h, by two obstruents (either oral or laryngeal), or by s plus n or w. Thus, for example, although the cluster ty did function to close a syllable in PNI, it does not do so in Seneca. A common ceremonial word usually translated as 'the Thunderers' can be reconstructed as 'hratiwēnotāyeʔs. The ty cluster prevented lengthening in the penult. In Cayuga the new accent pattern has taken over, and the fourth syllable is accented because the following syllable is closed.
hətɨwənətətyə's. But in Seneca the fifth syllable no longer counts as closed, and
the word appears with no accent at all: hətɨwənətətyə's.

In general, then, for phenomena which have to do with even-numbered syllables
in Seneca—both lengthening and accenting—there are fewer syllables which count as
closed than there were in PNI. The definition of a closed syllable has been narrowed.
As a consequence, there is more lengthening and less accenting, since lengthening
occurs in open syllables and accenting depends on the presence of a closed syllable.
But the definition of a closed syllable has been narrowed in somewhat different ways in
the two cases. We noted above that kh, for example, does not close a syllable so far
as lengthening is concerned; hence the lengthening in hətə-kheʔ. But for accenting,
any two obstruents, whether they be oral or laryngeal, will close a syllable. We
have, then, hətətəkheʔ 'they are running', in which the accent is on the second syllable
because the following syllable is closed. The same kh which fails to close a syllable
for lengthening does close it for accenting. We have here still another example of
some rather clear principles which are constrained by some rather fussy and seemingly
quite arbitrary restrictions.

Summary

We have seen how the reconstructed pattern of PNI, with penultimate accent and
lengthening of the accented vowel in open syllables, has been preserved essentially
intact in Mohawk and Oneida, although in both languages the definition of "penultimate"
has been complicated by the emergence of epenthetic vowels. In both these languages,
too, a falling pitch accent has developed out of syllable-final ? and h. Oneida has
sometimes extended the accent also to originally posttonic syllables. Onondaga, with
a few odd exceptions, has extended stress to all penultimate syllables, but has retained
the highest pitch on the antepenult when it was there originally. It has also extended
a new high pitch and lengthening to many antepenultimate syllables that were not
accented in PNI. Cayuga has kept the PNI accent on lengthened vowels, but has lost
it on short vowels. In the latter case it has placed a new accent on certain vowels
defined in terms of even numbered and closed syllables. It has also metathesized
? and h in odd numbered syllables, though with h the result has been the creation of
voiceless vowels. Seneca has lost the original accent altogether, though it has
preserved the lengthening which resulted from it. It has applied the new even numbered,
closed syllable accent wherever it will apply. In addition, it has narrowed the
definition of "closed syllable" for even numbered syllables, with the result that it
has fewer accented but more lengthened vowels than would otherwise be predicted.

References

SYNTACTIC ACCENT IN NORWEGIAN MORPHOLOGY

Alice Wyland Grundt

University of California, San Diego

The tonal accents of Norwegian and Swedish dialects have been studied for many years, both from a phonetic point of view to establish their parameters and from a historical/comparative point of view in an attempt to account for their origin and distribution. In this paper, I will be concerned with both the distribution and the phonetic realization of the tonal accents in East, or Standard, Norwegian. I will argue that certain instances of the contrasting accents which are currently accounted for by lexicon listing can be syntactically motivated. Specifically, I will consider the class of compound verbs and their associated nouns which do not differ in their segmental forms but contrast in accent. For example, uttale 'pronunciation' has Tone II while its associated verb uttale 'to pronounce' has Tone I, yet their segmental structure and their syllable structure do not differ.

The usual explanation for the tonal contrast is that the contrasting accents are reflexes of former polysyllables and were probably present allophonically at least in Old Norse. However, those modern Scandinavian languages which are most conservative, namely, Icelandic and Faroese, do not have the tonal contrast, nor do the dialects of Swedish spoken in Finland, the Norwegian dialects spoken in Northern Norway nor most of the Danish dialects. Martin Kloster Jensen also claims that at least the Bergen dialect of Southwest Norwegian does not have a real tonal contrast. If mere polysyllabicity were the source of the contrast, we should expect to find it in all the Scandinavian dialects and we would further expect that word pairs such as uttale and uttale should not differ in accent.

The two tonal accents of East Norwegian are contrastive in words or phrases of two syllables or more. Tone II cannot occur in monosyllables and the tone which does occur in monosyllables is usually assigned to Tone I. All accents of foreign words which retain their original position are also realized as Tone I. Phonetically, Tone I is a low pitch and Tone II is a falling pitch, but Haugen and Joos have shown that the difference between the two accents can also be described as one of phase difference: Tone I falls immediately in the syllable and Tone II falls late. Martin Kloster Jensen argues that the second syllable contains no contrasting information except that a tonal accent has occurred in the preceding syllable. I have made some preliminary investigations, using the pitch extractor on spoken tokens from Einar Haugen's tapes for his beginning course in Spoken Norwegian, and the results confirm
both the phase difference and the lack of contrast in the second syllable in terms either of the timing of the pitch rise or the height of the pitch rise. However, native speakers say that Tone II without the second syllable feels 'unfinished.' Many scholars have proposed that the second syllable in Tone II words has a secondary accent even though the phonetic analysis in terms of timing and pitch change indicates that such a secondary accent would be hard to prove.

The tonal accents occur on the first syllable of a tonal contour which may embrace an unstable number of syllables, usually from two to seven. While the tonal contours in Swedish are word-bound, those in East Norwegian may include more than one word and are not constrained to begin at a word boundary: they may start in the middle of a word. These tonal contours are called tonelagsgrupper in Norwegian but, following Haugen and Joos, I will refer to them as 'measures.' Each measure contains one, long accented syllable, the initial syllable. All other syllables in the syllable train of the measure are unaccented and short; and, in any utterance containing more than one measure, the measures tend to be isochronous.

East Norwegian measures usually contain members of syntactically bound phrases but not necessarily. For example, in the phrase sykdommens forløp 'the course of the illness' the two measures coincide with the two words of the phrase. However, in etter en tids forløp 'after a space of time' the first measure, which is etter en tids for- contains as its last syllable the first syllable of the following word while the second measure contains only the last syllable of the phrase-final word forløp. Thus, the measure is not constrained by either word boundaries or syntactic phrase boundaries. It appears to be a rhythmic pattern whose function is to serve as a timing unit for the utterance.

The strong effect of rhythmic force in East Norwegian can be seen in the compression or expansion of words and phrases to fit the measure: words containing liquids or nasals tend to be reduced to a two syllable measure as in the various forms based on menneske 'human being.' Alternatively, phrases or sentences which would otherwise end on an accented syllable will often be expanded with empty words to complete the trochaic rhythm as in morgen da [mörn na] 'morning, then' and han er død, han 'he is dead, he'.

Haugen (1967) notes that the Norwegian tones are not on a par with segmental phonemes but at the same time are distinct from the sentence intonational curves. He points out that the tones usually function as a means of disambiguating morphemic contrasts such as in tanke 'the tank' vs. tanke 'the thought' where the former has an underlying base tank and the latter tanke, each with a postposed definite article. He further suggests that the Norwegian tonal accents are similar in their distribution to the tonal accents of Lithuanian and Serbo-Croatian (Haugen, 1967, p. 185).
Halle (1971), among others, has pointed out the 'well known fact... that in cognate words the accented vowel with the "falling" tone in Standard Stokavian corresponds to the stressed vowel in Russian, whereas the accented vowel with "rising" tone in Standard Stokavian corresponds to the pre-tonic vowel in Russian.' (Halle, p. 6). Halle's remark suggests that retraction of accent will produce a contrasting tonal accent and, in fact, Halle's description of the Stokavian accents recalls the southern Swedish dialect studied by Hadding (1961) in which Tone I is falling and Tone II is rising. (Lehiste, 1963, p. 356).

There is evidence in East Norwegian sentences, phrases, and words which suggests that Tone II is often derived from accent retraction. In sentences such as jeg vil gå hjem 'I want to walk home', when the verb går appears with Tone I, the sentence is interpreted to mean 'I want to walk home' while if it occurs with Tone II, the sentence means 'I want to go home.' That is, if this were an English sentence, we would shift the stress from the verb to the following noun when we wish to emphasize the goal of movement, but in Norwegian the accent remains on the verb and shifts to Tone II. Similarly, in the sentence hun er så snill 'she is so nice' the neutral comment would have Tone I on the verb er but if the adverb så is to be emphasized as in 'she is so nice' the accent remains on the verb er but changes to Tone II.

In phrases such as sykdommens forløp and etter en tids forløp the word forløp in citation form has Tone II on the first syllable but, when the measure begins with the second syllable, the accent is Tone I. Prefixes like för are unaccented in Low German loans and when they remain unaccented, the second syllable is in Tone I as in forbruker 'consumer' but when the accent occurs on the prefix in non-verbs, the accent is Tone II as in forbruk 'consumption'. In verbal phrases such as å ta på noe, if the verb ta has Tone I, the following word på is interpreted as a preposition governing noe and the phrase means 'to touch on something' but if the verb ta is in Tone II, the following på is interpreted as a verbal particle and the phrase means 'to take something on.'

Finally, there are a considerable number of loan words which take Tone I when they retain their proper accent placement, but if the proper accent is on a non-initial syllable and the accent is nevertheless realized on the initial syllable, it becomes Tone II as in avis vs. åvis 'newspaper', manér vs. måner 'manner'. This accent retraction and shift to Tone II is said to be characteristic of less well educated and country people.

If we wish to derive Tone II by means of accent retraction from an underlying secondary accent, we can motivate a great many instances of Tone II in East Norwegian. However, if we wish to motivate the accent contrast in words like åttale vs. ûttale by means of accent retraction in the noun but not in the verb, we must posit root
accent in the noun but not in the verb. I propose to use two assumptions in order to motivate this difference in underlying root accent: 1) that in Germanic, as usually assumed, the primary accent occurs on the root syllable, and 2) that the assignment of primary accent is based on sentence accent.

If we consider the sentence accentuation of Old Norse as determined by an examination of prosody, it is possible to motivate the lack of root accent in verbs: according to Gordon (1957), primary accent in sentences fell on nouns, adjectives, and adverbs. A secondary accent fell on the second members of compounds and on derivational suffixes and, according to Benediktsson (1968), the suffix also had to occur in either a heavy syllable or the first of two light syllables to receive secondary accent. Weak or unaccented words included pronouns, prepositions, and verbs. Such a pattern of sentence accent is very old and widespread: Kurylowicz (1959) has shown that the same pattern of accented nouns and unaccented verbs was also true of Sanskrit, is attributed to reconstructed Indo-European, and can be shown to be true of Hebrew as well. In addition, Kurylowicz points out that, for the combination of a verb and a verbal particle as in 'to take something on', the particle was accented and the verb was not, whether the two elements were separated in the sentence or were compounded with the particle preposed. Popperwell (1963) essentially confirms this pattern for modern East Norwegian except that he says nothing about secondary accent and indicates primary accent for 'principle verbs' but not 'auxiliary' verbs without further definition.

A number of suffixes must be lexicon-marked for East Norwegian as inducing Tone II in a preceding syllable but sound changes have often caused nominal and verbal suffixes to merge phonetically: both the plural marker -er and the present tense marker -er induce Tone II unless the plural marker occurs with a noun of the umlaut class as in tómer 'teeth' (Tone I) vs. dagar 'days' (Tone II), or the verb is a strong verb also belonging to the umlaut class as in knökkjer '(it) breaks' (strong, intransitive) vs. knëkkjer '(he) breaks (it)' (weak, transitive). Similarly, the weak masculine nominal suffix, the infinitive marker, and the postposed definite article have merged in -e [ɛ] but only the postposed neuter article does not induce Tone II: lån [lɔnɛ] 'to lend' vs. lånnet [lɔnɛt] (lån + et) 'the loan'. Since such Tone II-inducing suffixes must be lexicon-marked, they have become arbitrary and have lost their status as carrying secondary accent. The case is similar for compounds: compounds take the accent of their first member unless that member is monosyllable in which case we are faced with the anomalies such as āttal vs. āttal, being considered in this paper.

If we assume that secondary accent of earlier times is now lexicon-marked and that primary accent of earlier times is still assigned to the root but rather by sentence function instead of morphological shape, we must still account for the
relation of these arbitrary and syntactic accents to the measure accent which, as we have seen, is not grammatically conditioned, in order to derive the correct surface accent in both words and phrases. To express this relation, I propose that two descriptions of accent are necessary: a lexical/syntactic description which assigns accented syllables arbitrarily or grammatically and a phonetic description in terms of the initially-accented measure which functions as a timing unit of the utterance. The intersection of the initially accented syllable of the measure with the grammatical accent will produce a phonetic Tone I and the intersection of the measure accent with a syllable preceding the grammatical accent will produce a phonetic Tone II.

The rhythmic measure, once introduced into an utterance, is fairly rigid, but the Norwegian speaker has several means at his disposal to manipulate his utterance in such a way as to ensure that the measure accent will fall on the syntactically important words as far as possible. He may delay the beginning of the first measure with two or three neutral syllables called a 'neutral anacrusis' by Haugen (1967, p. 198) as in etter en in the phrase etter en tds fortælp; he may add empty words when necessary to fill out the measure as in han er død, han; he may drop syllables in words which will permit it as in menneskene; and he may shorten and compress a syllable train as in bare på ham så lenge, but he must maintain the rhythm based on the isochronous measures in order to sound like a Norwegian.

By positing only three ordered phonological rules for the phonetic description of the accent which apply after the lexical/grammatical accent has been assigned, it is possible to account systematically for both the anomalous noun/verb class we have been considering and, in addition, all the words which are usually accounted for by the polysyllabic illustration argument. This can be clearly seen in the sample derivations for uttale 'pronunciation' vs. uttale 'to pronounce', lyse 'illumination' vs. lyse 'to shine', lænet [lænt] 'the loan' vs. lene [lænt] 'to lend', knækkere 'it breaks' vs. knækker 'he breaks it'. The first rule assigns measure-initial accent, the second rule will derive Tone II only if a lexical/grammatical accent occurs in the following or second syllable, and the third rule deletes all the accents in the measure except the measure-initial accent. Thus, if Rule P-2 does not apply, Tone I will be phonetically realized, but if Rule P-2 does apply, Tone II will be realized.

Although my solution requires a double description of the accent in terms of grammatical function on the one hand and in terms of phonetic timing units on the other, I think that this solution can be justified on several grounds. First, the solution not only reduces arbitrariness in the grammar, but it also motivates the assignment of grammatical accent in terms of function in a sentence. Simple Germanic root accent is ultimately as arbitrary as lexically assigned accent and, furthermore, cannot account for the anomalous noun/verb class. Second, the separation of the
descriptions of function and phonetic structure provides phonetic motivation for loss of accent and length which cannot be accounted for by grammatical considerations. Third, the separation of functional and structural descriptions of the same phenomena agrees with the findings of all other scientific disciplines, both the hard sciences such as physics, chemistry, and biology, and the social sciences such as sociology, psychology, and economics, all of which have found that functional and structural descriptions of the same phenomena are different and non-comparable. Therefore, unless we separate functional elements of speech utterances from the phonetic means with which we may express them, we will not be able to determine the relationship between them.

APPENDIX

**East Norwegian Tonal Accents - Phonological Characteristics:**

<table>
<thead>
<tr>
<th>Tone I</th>
<th>Tone II</th>
</tr>
</thead>
<tbody>
<tr>
<td>low pitch</td>
<td>falling pitch</td>
</tr>
</tbody>
</table>

(written: [á])

Accent domain: two syllables

Accented syllable: long (CVVC, CVCC)

Unaccented syllable: short

_Tonelagsgruppe_ (‘measure’): tonal contour occurring over an unstable number of syllables (approx. two to seven). May include more than one word and may begin in the middle of a word. First syllable of train is long and accented but no other accents may occur within the measure. Measures of an utterance tend to be isochronous.

Measures - permitted syllable trains:

- 2 syllables: høre
- 3 syllables: høre på
- 4 syllables: høre på ham
- 7 syllables: høre på ham så lenge

‘to listen’

‘to listen to’

‘to listen to him’

‘to listen to him for the time being’

Effects of trochaic (strong + weak) rhythm:

Compression:

<table>
<thead>
<tr>
<th>Norwegian</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>menneske [menn̩skə]</td>
<td>[mønskə] 'human being'</td>
</tr>
<tr>
<td>mennesker [menn̩skə:kə]</td>
<td>[mønskə:kə] 'people'</td>
</tr>
<tr>
<td>menneskene [menn̩skənə]</td>
<td>[mønskənə] 'the people'</td>
</tr>
<tr>
<td>menneskelig [menn̩skəlɪɡ]</td>
<td>[mønskəlɪɡ] 'human'</td>
</tr>
</tbody>
</table>
Expansion (addition of empty words):

(God) morgen, da! [mørn na] '(Good) morning, then'
Han er død, han. 'He is dead, he'

Evidence for underlying secondary accent:

Sentences:
Tone I: jeg vil gå hjem 'I want to walk home.'
Tone II: jeg vil gå hjem 'I want to go home.'
Tone I: hun er så snill. 'She is so nice.' (neutral)
Tone II: hun er så snill! 'She is SO nice!' (emphatic)

Phrases:
Tone I (2 measures): etter en tids forløp 'after a space of time'
Tone II (2 measures): sykdommens forløp 'the course of the illness'
Tone I: å ta på noe 'to touch (on) something' (verb: ta)
Tone II: å ta på noe 'to take something on' (verb: ta på)

Accent vacillation and retraction in loan words:

<table>
<thead>
<tr>
<th>Tone I</th>
<th>Tone II</th>
<th>Gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>avis</td>
<td>avis</td>
<td>'newspaper'</td>
</tr>
<tr>
<td>måner</td>
<td>måner</td>
<td>'manner, behavior'</td>
</tr>
<tr>
<td>kontor</td>
<td>kontor</td>
<td>'office'</td>
</tr>
<tr>
<td>telefon</td>
<td>telefon</td>
<td>'telephone'</td>
</tr>
<tr>
<td>instruere</td>
<td>instruere</td>
<td>'instruct'</td>
</tr>
<tr>
<td>arbeidsom</td>
<td>arbeidsom</td>
<td>'hard-working'</td>
</tr>
<tr>
<td>mistenksom</td>
<td>mistenksom</td>
<td>'suspicious'</td>
</tr>
</tbody>
</table>

Note: Tone I in words retaining loan-word accent
Tone II in words with retracted accent
Syntactic (sentence) accent:

Indo-European, Sanskrit:
- nouns = [ + accent ]
- verbs = [ - accent ]
- verb + particle =
  1) when separated in sentence,
     particle = [ + accent ]
     verb = [ - accent ]
  2) when compounded with particle
     preposed:
     particle = [ + accent ]
     verb = [ - accent ]

Hebrew:
- nouns = [ + accent ]
- verbs = [ - accent ]

Old Norse:
- Primary (root) accent: nouns, adjectives, adverbs
- Secondary accent: B-members of compounds
derivational suffixes
  (if in heavy syllable or first of
two light syllables)
- Weak (unaccented): finite verbs, endings, prepositions,
  pronouns, conjunctions

East Norwegian:
- Primary accent: nouns, adjectives, adverbs,
demonstrative and interrogative
  pronouns, 'principle verbs'
  (not further defined)
- Weak (unaccented): auxiliary verbs, endings, conjunc-
  tions, prepositions, personal
  pronouns, conjunctive adverbs

Homophones except for accent:

<table>
<thead>
<tr>
<th>Nouns (Tone II)</th>
<th>Verbs (Tone I)</th>
</tr>
</thead>
<tbody>
<tr>
<td>inn-rele</td>
<td>ñn-rele</td>
</tr>
<tr>
<td>'arrangement, layout'</td>
<td>'to build, to lay out'</td>
</tr>
<tr>
<td>av-reise</td>
<td>áv-reise</td>
</tr>
<tr>
<td>'departure'</td>
<td>'to depart'</td>
</tr>
<tr>
<td>an-klage</td>
<td>án-klage</td>
</tr>
<tr>
<td>'accusation'</td>
<td>'to accuse'</td>
</tr>
<tr>
<td>òt-tale</td>
<td>ùt-tale</td>
</tr>
<tr>
<td>'pronunciation'</td>
<td>'to pronounce'</td>
</tr>
</tbody>
</table>
Homophones:

<table>
<thead>
<tr>
<th>Nouns (Tone II)</th>
<th>Verbs (Tone II)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rejse</td>
<td>'journey'</td>
</tr>
<tr>
<td>rekke</td>
<td>'line, file, rank'</td>
</tr>
<tr>
<td>kægge</td>
<td>'keg, barrel, float'</td>
</tr>
<tr>
<td>løjse</td>
<td>'a loosened mass'</td>
</tr>
<tr>
<td>lyse</td>
<td>'illumination'</td>
</tr>
<tr>
<td>drive</td>
<td>'drift (of snow, sand)'</td>
</tr>
</tbody>
</table>

Nouns (definite and indefinite):

tanken (tank + en) 'the tank' en tank 'a tank'
tanken (tanke + en) 'the thought' en tanke 'a thought'

Verb vs. definite noun:

lån (lån + et) [lɔːn] 'the loan' et lån 'a loan' [ɛt lɔːn]
låne [lɔːnɛ] 'to lend'

Strong vs. weak verbs:

knækker '(it) breaks' -strong verb, intransitive
knækker '(he) breaks (it)' -weak verb, transitive

Suffixes inducing Tone II (partial list):

Nouns: plural marker -er weak masc. ending -e

Verbs: present tense marker -er infinitive marker -e

Proposed change: -er, -e → -ér, -é

Unstressed formatives:

Postposed definite article: Singular Plural

Common gender -en -éne (-ér + -ene)
Neuter gender -et [-ɛ] -éne (Ø + -ene)
Unstressed prefixes:

for' forbygning 'retaining wall; bad building'

er' erfare 'to experience, to find out'

be' bevisse 'to prove'

ge' geberde seg (reflexive) 'to behave'

Rules and derivations:

Lexicon marking:
1) Foreign words - intrinsic accent
2) Certain suffixes - intrinsic accent
3) Certain prefixes - intrinsic lack of accent

Syntactic accent:
1) Accented root; nouns, verbal particles
   (partial list)
2) No accent assigned: verbs
3) Intrinsic lack of accent: postposed definite articles

Phonological rules:

Redundance: [+ accent] = [+ low pitch]
               [+ Tone II] = [+ falling pitch]
               S = syllable
               # = 'measure' boundary
               ## = end of utterance

Ordered rules:

P-1 [- accent] [+ accent] / [# + S] S_o # [+ accent]

P-2 [+ accent] [+ Tone II] / # [+ S] [+ accent] S_o #

P-3 [+ accent] [- accent] / # [+ accent] S_o [+ S] #

Sample derivations: 'tale 'pronunciation' lyse 'illumination'
                     'to pronounce' 'to shine'
Lexicon:  
<table>
<thead>
<tr>
<th>Word</th>
<th>Syntactic accent</th>
<th>Surface</th>
<th>Pronunciation</th>
</tr>
</thead>
<tbody>
<tr>
<td>lønt</td>
<td>(N)</td>
<td>ut-tal-e</td>
<td>'to lend'</td>
</tr>
<tr>
<td>lån</td>
<td>(N)</td>
<td>ut-tal-e</td>
<td>'the loan'</td>
</tr>
<tr>
<td>lys</td>
<td>(V)</td>
<td>lys-e</td>
<td>'illuminations'</td>
</tr>
<tr>
<td>knór</td>
<td>(V)</td>
<td>lys-e</td>
<td>'to shine'</td>
</tr>
</tbody>
</table>

P-rules:  
<table>
<thead>
<tr>
<th>Rule</th>
<th>P-1</th>
<th>P-2</th>
<th>P-3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>#ut-tal-e#</td>
<td>#ut-tal-e#</td>
<td>#lys-e#</td>
</tr>
<tr>
<td></td>
<td>#ut-tal-e#</td>
<td>-</td>
<td>#lys-e#</td>
</tr>
<tr>
<td></td>
<td>#ut-tal-e#</td>
<td>#ut-tal-e#</td>
<td>#lys-e#</td>
</tr>
</tbody>
</table>

Surface:  
- 'pronunciation'  
- 'to pronounce'  
- 'illumination'  
- 'to shine'  

Note: Strong verbs like knór must be marked to take unaccented suffix.

REFERENCES


Grundt, Alice Wyland. 1971. Tonal accentuation in East Norwegian and the place of accent in the grammar. Mss.


TONAL ACCENT IN CREEK

Mary R. Haas
University of California, Berkeley

1. INTRODUCTION

Creek is a Muskogean language formerly spoken in Alabama, Georgia, and other parts of the Southeast. At the present time it is spoken in eastern Oklahoma and by one of the groups of Seminoles in southern Florida.1

This language has tonal accent. In other words, there is a pitch-defined key syllable which determines the pitches of surrounding syllables unless another key syllable follows. The working out of the nature of this tonal accent proved extraordinarily difficult at first. The usual method of recording pitches with diacritics tends to obscure rather than reveal the true pattern. The reason is that the usual grid achieved by this device is too coarse since it is not possible to operate in terms of a predetermined number of levels. Another surprising aspect of the problem is that thousands of words have largely predictable patterns and, as a consequence, minimal pairs are extremely rare in a list of lexical items, no matter how extended. Indeed the intricate nature of the Creek tonal accent does not really become apparent until careful attention is paid to the tense-aspect system of the verb. Then a wealth of minimal and near-minimal contrasts occur. But unless translations which accurately express tense-aspect have been obtained these contrasts can be easily overlooked at first. Take, for example, the following pairs:

(1) 
\[
\begin{align*}
\text{a. } & \text{la`ni} & [\ldots] \text{yellow, green} \\
\text{b. } & \text{lani} & [\ldots] \text{turning yellow, green}
\end{align*}
\]

(2) 
\[
\begin{align*}
\text{a. } & \text{a`yaks} & [\ldots] \text{he is going (II)}^{3} \\
\text{b. } & \text{a`yark} & [\ldots] \text{he went (II)}
\end{align*}
\]

If examples like these are not carefully distinguished in both pitch and translation, the presence of tonal accent will simply be missed. The average native speaker will tend to define both members of such a pair in terms of the basic lexical meaning.
Thus the most natural translation for both la and lb is simply 'yellow, green'. This appears to be at least one of the reasons why tonal accent was overlooked by such early workers as Gatschet, Swanton, and Speck.

Another important feature of Creek tonal accent is that there can be one or more than one key syllable in a word. And in the event there are more than one, each succeeding key syllable will be one step lower than the preceding key syllable. In other words, we have what may be called downward drift. Examples:

(3)

a. na·fki's [− −] he is hitting him
b. náfka·kis [− −] they are hitting him.

This and other characteristics are more fully described in the sections which follow.

2. DEFINITIONS

In order to state the rules about Creek tonal accent it is necessary to define the syllables of the language on two axes, (1) the structural and (2) the functional. Structurally there are three types of syllables, light (L), heavy (H), and aphaeretic (A). Light syllables consist of C or # (word-onset) plus a short vowel. Heavy syllables consist of C or # plus a long vowel, C or # plus a short or long vowel plus C. Heavy syllables must in turn be divided into a sonorant type, C or # plus a short vowel plus a sonorant, and an obstruent type, C or # plus a short or long vowel plus an obstruent. Aphaeretic syllables are lost initial syllables of either the light or heavy type whose loss is revealed by the pitch of the new initial syllable provided it is light. Summing up we have the following:

Light: CV

Heavy: CV
  CV(·)C
  Sonorant: CVM
  Obstruent: CV(·)P

Aphaeretic: ' (Apostrophe)

Examples:

(4)

a. ayíta LLL to go  b. a·yéys HH I am going
c. hátki· HH white  d. hoki· HH woman
e. bonánwa LHL man  f. ¹·kəna' HLL land
On the functional level there are key or determining syllables distinguished from subordinate or nonkey syllables. Once the key syllable is marked, the pitches of all subordinate syllables are predictable in terms of their structure (L, H, A). A key syllable may have any one of the following three 'tones':

- level (\(\uparrow\))
- falling (\(\downarrow\))
- extra high or slightly rising and crescendo (\(\uparrow\))

Every word except a light monosyllable will have at least one key syllable bearing a tonal accent. However, there may be several more key syllables within a word and they will often be consecutive, though this is not necessarily so.

3. DESCRIPTION OF PITCHES

3.1 At this point a more detailed description of the pitches is needed. Although all are relative and not absolute, it will be convenient to describe them on a descending scale of 1 (highest) through 5 or more to d ('deep'). (The 'deep' pitch is the lowest in the word, no matter how many syllables precede.) But before giving the rules for the pitches of key syllables, it is important to understand the nature of the pitches ascribed to three types of nonkey syllables, (1) initial, (2) medial, and (3) final.

3.1.1 Except for monosyllables and certain kinds of disyllabic words, an initial light nonkey syllable will have a pitch that is slightly lower than the pitch of the following syllable, regardless of whether that syllable is key or nonkey. Such an initial syllable will be marked as i ('initial') so that the need to assign a definite pitch may be obviated. Examples:

(5)

a. ifólo i-2-d screech owl
b. apatana' i-3-3-3 bullfrog
c. osahwa i-2-d crow
d. cawanay'ta i-3-3-2-d one person to tie me

e. hatâm 1-2 d again

f. acolak' i-3-3-3 (or) 'colak' 3-3-3 old timers, elders

This last example illustrates how the pitch of the syllable -co- remains the same whether or not the initial syllable has been lost.

Monosyllabic words comprising a light syllable have no tonal accent, e.g. ma 'that'. But monosyllabic words comprising a heavy syllable will always have a tonal accent, e.g. l̂ 'arrow', l̂yks 'he is seated'.

In disyllabic words beginning in a light nonkey syllable this syllable will normally have the pitch of the following key syllable if level, or it may be slightly lower. Examples:

(6)

a. coffi LL 3-3 rabbit

b. osá' LH 3-3 (or) i-3 poke weed

And see also 5e above.

3.1.2 Just as a L initial syllable may be slightly lowered, so a medial L syllable may also be slightly lowered in certain circumstances, viz., when between two heavy syllables or when preceding a nonfirst key syllable. Such a syllable is marked m ('medial'). Examples:

(7)

a. na'fkay'ajks i-3-m-3 (or) i-3-3-3 I was hitting him (II)

b. atōkay'ajks i-24-m-3 I worked (II)

3.1.3 A final or ultimate nonkey syllable in a word has the deep pitch. Such a syllable may be L, as in 5a, 5c, and 5d above, or it may be H. Examples:

(8)

a. aha'gátki- 1-3-2-d counted (past participle)

b. hökô-lin 3-24-d two (in counting)

c. hi'ni. 21-d good
3.2 We are now ready to describe the pitches of the key syllable or syllables in a word. The first or only key syllable in a word will have the following pitches according to its type, as described in the following paragraphs.

3.2.1 If first, LEVEL, and nonultimate, the key syllable will have the pitch of 2, e.g.

(9)  
   a. is'ita  1-2-d  one to take one  
   b. cawanay'ita  1-3-3-2-d  one person to tie me  
   c. fsti  2-d  person  
   d. lá-ni.  2-d  yellow, green  
   e. nisáhkis  1-2-d  they bought it  

But if first, LEVEL, and ultimate, it will have the pitch of 3, e.g.

(10)  
   a. apataná  i-3-3-3  bullfrog  
   b. nafkitá  3-3-3  one to hit one  
   c. fó.  3  bee  
   d. la-tká.  3-3  June bug, June beetle  
   e. okylhá.  3-m-3  (or)  3-3-3  mosquito  
   f. hofonnana.  1-3-3-3  long, long ago  
   g. nafkitiká-ylta  3-3-m-3-3-3  (or)  3-3-3-3-3  one to hit several

3.2.2 If first, FALLING, and nonultimate, it will have the pitch 24, e. g.

(11)  
   a. ná'ki  24-d  something, anything, what  
   b. ci-kólko  3-24-d  purple martin  
   c. a-fackálki  3-3-24-d  the ball players
d. hokkô·lin 3-24-d two (in counting)
e. ântawa 24-d-d the wilderness

If first, FALLING, and ultimate, it will have the pitch 2d, e.g.

(12)
   a. lâyks 2d he is seated
   b. hatâm 1-2d again
   c. wana·yâyô·f 1-3-3-2d when I am tying him
   d. cana·fâ·t 1-3-2d he who is hitting me

3.2.3 If first, EXTRA HIGH or RISING, and nonultimate, the key syllable will have the pitch 2l or 1, e.g.

(13)
   a. hokkô·nla 3-21-d (or) 3-1-d both
   b. nâ·nkêys 21-d (or) 1-d he kept hitting him (I)
   c. yahâ·ynkeys i-21-d (or) i-1-d he kept singing (I)
   d. hâlnwôsî 21-3-d (or) 1-3-d extremely high

3.3 We now come to the most interesting feature of Creek tonal accent, namely the description of what happens when there is more than one key syllable in a word. As was briefly mentioned in the introduction, we have what we may call downward drift. This means that nonfirst key syllables will have progressively lowered pitches according to the scheme described in the following paragraphs.

3.3.1 Each nonfirst key syllable will be pitched one step lower than the immediately preceding key syllable. Thus if the first key syllable is 2, 24, or 21, the second will be 3, 35, or 3d (ultima). Examples:

(14)
   a. pó·fâ·ks 2-3 they are blowing (with mouth)
   b. wanâ·hyô·cko·f i-2-3-d when you tied him (I)
   c. nisâ·ña·nâks i-2-3-3 I am not about to buy it
d. náfka·këys 2-3-3d they were hitting him (I)
e. fóll·plyâ·t 2-3-m-3d when we were about

(15)
a. náf·këys 24-3 he hit him (II)
b. sâ·sâkwa 24-3-d goose
c. canafêyêckîs 1-3-24-3-d you hit me (I)
d. nâ·fìkayantâs 24-m-3-3 he hit him (IV)
e. haìntacokwocì 1-24-m-3-3-3 hummingbird

(16)
a. atô·ntkâ·ti·s 1-21-3-d I will keep on working
b. lá·nstósì 21-3-d one completely black, black all over

And in the same way, the third, fourth, etc., key syllable will in turn be one step lower than its immediately preceding key syllable. Examples:

(17)
a. atôtkkkârkās 1-2-3-4 he didn't work (II)
b. atô·tkîckantâs 1-24-3-4-4 you worked (IV)
c. lá·nslâ·tósì 21-3-4-d several completely black
d. wanâylocokomâc 1-23-4-4-5 you didn't tie him (III)
e. atô·ntkîccêkâ·ti·s i-21-3-4-5-d you will not be working regularly, constantly

There now remain the subordinate or nonkey syllables whose pitch is predictable in terms of (1) their position in the word and (2) their position in relation to surrounding key syllables. The rules are given in the sections which follow.

3.3.2 An initial L syllable will have pitch 1 (see 3.1.1); any other pretonic L syllable (including one preceded by A) and any pretonic H syllable will have pitch 3. Examples:

(18) a. hícîta LLL 1-2-d to see, look at
b. wanayłta LLLL 1-3-3-3 one to tie one

c. cawanayłta LLLLL 1-3-3-2-d one to tie me

d. acahar'katita LLLL 1-3-3-3-2-d one to count me

e. acolak' LLLL 1-3-3-3 (or) 'colak' ALLL 3-3-3 oldsters, old timers

f. 'tacs·koci ALHLL 3-3-3-3 Sunday (earlier nittaca·koci)

3.3.3 A nonultimate intertonic (between key syllables) H syllable will have the pitch of the next following level key syllable while a L syllable in the same position will have that pitch or m pitch (see 3.1.3). Examples:

(19) a. ahar'ka·teymac 1-3-24-3-3 I counted (III)

b. ahár'ka·tayárks i-2-3-m-3 I was counting (II)

c. ató-nñtiyális i-21-m-3-d we will keep on working

On the other hand, a nonultimate intertonic L or H syllable will be one step below the high part of a falling key syllable. Example:

(20) iwanhka·keys i-2-4-3d they were getting thirsty (I)

A posttonic L or H syllable has the lowest pitch, marked d, in the word, and if there is more than one posttonic syllable, it will have the same pitch. Examples:

(21) a. nafeykičcis i-24-3-d you hit him (I)

b. atoteykičcica i-3-23-3-d-d you have worked (I)

4. PREDICTABLE AND NONPREDICTABLE TONAL ACCENT

4.1 Some morphemes contain a syllable with a fixed tonal accent and some do not. However, there are many words made up completely of morphemes having no fixed tonal accents. In the latter case the placement of the tonal accent is automatic according to the rules described in the following paragraphs.

4.1.1 In any string of L syllables containing no fixed accents the tonal accent will be placed on the last even-numbered syllable.10 In other words, if the total
number of syllables in the word is odd, the tonal accent is on the penult, but if the
total number is even, the tonal accent is on the ultima. An A syllable counts as one
L syllable. Examples:

(22)  
  a. ma that (no accent because there is no penult)
  b. ifá dog
  c. ifóci puppy
  d. amífocí my puppy
  e. hicíta one to see one
  f. ahicíta one to look after, watch one
  g. imahiríta (or) 'mahiríta one to look after for (someone)
  h. isimarícitá (or) 'isimarícitá one to sight at one
  i. itiwanayípíta (or) 'tiwanayípíta to tie each other

If the penult is a H syllable and the ultima a L one, the tonal accent will be placed
on the penult. Examples:

(23)  
  a. cá·lo trout, bass
  b. sókca sack, bag
  c. pocós ça axe
  d. famí·ca canteloupe, muskmelon; perfume
  e. ałakkóycka appreciation

Furthermore, if there is a string of L syllables following a H syllable, the rule of
placing the tonal accent on the last even-numbered L syllable will apply. Examples:

(24)  
  a. aktopá bridge
  b. wa·kocí calf
c. hoktañi women

d. μγκοσαπίτα one to implore

e. αλπατόκι baby alligator

f. yapaphoyn̄a two to walk

In the examples shown in (22), (23), and (24), the ultima is always a L syllable. However, if the ultima is a H syllable, then the placement of the tonal accent is no longer automatic. Instead morphological rules come into play at this point. Hence the tonal accent may be on the penult or on the ultima depending on nonphonetic factors. Examples:

(25)

a. hάτκι white

b. hοκτί woman

c. cáhtki stuck up (as posts)

c. cá kutki sticking up (as posts)

4.1.2 There are many morphemes which have a fixed tonal accent. This means that under all circumstances this morpheme will retain this accent no matter how many other accents also occur within the word. Certain aspectual forms of the verb have fixed accents and the same is true of some of the pronominal and tense and modal suffixes. Examples:

(26)

a. nίhsís he bought it (I). nίhs- completive aspect (I) of root nís-.

b. nίhsíc̄kís you bought it (I). -ćk- you (subject).

c. nafč̄ykís he hit him (I). nafč̄yk- completive aspect (I) of root nafk-.

d. nafč̄ykc̄kís you hit him (I).

e. ní sác̄jkís he bought it (II). ní s- completive aspect (except I), -ác̄jk- (II) past tense marker.

f. ní stıc̄jkís you bought it (II)
g. ná-nfkis he is constantly hitting him. ná-nfk- continuative aspect.

h. ná-nfkı́čkis you are constantly hitting him.

i. nisíkomác he didn't buy it (III). -íko- 'not', -mác (III) past tense marker.

j. nisíccúkomác you didn't buy it (III). -ícc-, dissimilated form of -íck- you.

In addition to morphemes that have fixed accents, there are a few suffixes which have variable accents, e.g. -o·f N -o f 'when, if' and -ey- áy- (I past tense marker used with the progressive aspect).

(27) a. ni-sayō·f when I am buying it. ni s- progressive of root nis-.

b. ni-sícko·f when you are buying it

c. ni-sayėys I was buying it (I)

d. ní-nseys he was continually buying it (I)

e. nisákęys he didn't buy it (I). -ík(o)- not.

f. na-fikęys he was hitting him (I). na-fık- progressive of root nafk-

g. náfka-kęys they were hitting him (I). nafka-k- progressive of nafkak-, plural of root nafk-

In general the unmarked accent (e.g. -o·f) occurs when there is an immediately preceding accent in the word (27b, 27d, 27e), whereas the marked accent occurs otherwise (27a, 27c, 27g).

5. CONCLUSION

The type of tonal accent system that has been described in this paper is not a commonly occurring type. It is particularly interesting because of the possibility of multiple accents in the word accompanied by the downward drift that characterizes successive accents. Other Muskogean languages also appear to have tonal accent systems but these have not yet been thoroughly described so it is not yet possible to determine in what ways they may differ. The problem of how
such a system has evolved in the several daughter languages and out of what manner of protosystem awaits future investigation.

FOOTNOTES

1. Most of my work on the Creek or Muskogee language was done during the course of several field trips to Oklahoma (in the old Creek Nation) between 1936 and 1940. More recently, in the course of a brief visit to Oklahoma in 1969, I had the opportunity to recheck some of my conclusions about the tonal accent. My earlier work was supported by the Department of Anthropology of Yale University, the American Philosophical Society (Phillips Fund), and the Committee on Research in American Native Languages.

2. A list of the Creek tonal accents is included as part of a brief statement on Creek phonemics in an early paper on Creek ablaut (Haas 1940). A couple of years earlier I had used a different and, as it turned out, inadequate method of indicating the tonal features of Creek (Haas 1938). Several different ways of indicating these features were experimented with before I finally arrived at the solution presented in some detail in the present paper.

3. The roman numerals I, II, III, and IV, when placed in parentheses after the gloss, indicate first, second, third, and fourth past tenses, respectively. (See Haas 1940.)

4. For examples of the work of these earlier investigators, see Gatschet (1884), Speck (1907), and Swanton (1924). Although haphazard attempts were made by all of them to indicate some kind of stress, many significant differences simply could not be handled in this way.

In the mid-nineteenth century, Rev. R. M. Loughridge, missionary to the Creeks, and David M. Hodge, his interpreter, developed an alphabet for Creek. In their work, too, an attempt to cope with the tonal accent was made by the use of an acute accent mark, presumably to indicate what they thought was stress. However, inconsistencies were inevitable and most Creeks who learned the alphabet simply omitted the accent. However, at least one literate (in the Creek alphabet) speaker with whom I worked in the late 1930s recognized the fact that some words which had to be written alike were actually different, though he had no way to handle the problem.
This feature was briefly noted but not so named in the short phonetic statement found in Haas (1940:150). In recent years a somewhat similar phenomenon, as it applies to "syllable-tone" languages, has been described for a number of African languages (cf. Hyman 1973).

The symbol M is used as a cover symbol for any sonorant and P for any obstruent.

In some of my earlier papers I wrote disyllabic words of LL structure with the tonal accent on the first rather than the second syllable. This is similar to an earlier writing of LLLL words with the tonal accent on the second rather than the last as at present. It just happens that by oversight the LL words were the last to be changed over to the present system of marking the accents.

When the symbol n appears between vowel length and a consonant, as in hihn 'good', it indicates nasalization of the preceding vowel. Similarly, when n appears between a sonorant and another consonant, as in 13c and 13d in 3.2.3, it indicates nasalization of both the preceding vowel and the sonorant. Hence the sequence ay in 13c and the sequence al in 13d are nasalized.

See footnote 8.

I owe this felicitous generalization about automatic placements of the tonal accent in a string of L syllables to P. Ladefoged and V. Fromkin in the course of the discussion during the symposium. My original statement had been solely in terms of an odd or even totality of syllables.

REFERENCES

Haas, Mary R. 1938. Geminate consonant clusters in Muskogee [Creek]. Lg. 14. 61-65
Haas, Mary R. 1940. Ablaut and its function in Muskogee [Creek]. Lg. 16. 141-150. [See especially pp. 149-150.]


Swanton, John R. 1924. The Muskogean connection of the Natchez language. IJAL 3.46-75.
TOWARDS A RECONSTRUCTION OF THE INDO-EUROPEAN ACCENT

Paul Kiparsky
Morris Halle
MIT

Our thesis is that the accentual systems of Slavic, Lithuanian, Vedic and Classical Greek share the following general abstract characteristics, which are to be reconstructed also for Proto-I. E.:

(A) They are pitch accent systems as opposed to tone systems. Specifically there is only one kind of primary accent, and phonetically every word normally has one and only one such accent.

(B) In underlying representations, however, words may have more than one accented vowel, or no accented vowel.

(C) The underlying representations of words are realized phonetically by the Basic Accentuation Principle:

(1) If a word has more than one accented vowel, the first of these gets the word accent. If a word has no accented vowel, the first vowel gets the word accent.

(D) There are three basic classes of morphemes:

unaccented morphemes

accented morphemes, which have an accent on one vowel in their underlying representation

preaccenting morphemes, which induce an accent on the immediately preceding vowel. We assume that such morphemes trigger the rule of Metatony (2)

(2) \( V \rightarrow [h]_V / -c_O + c_OV \) where \([h]_V\) represents an accented vowel.
Because of the Basic Accentuation Principle any inherent accent on a preaccenting morpheme will be eclipsed by the accent that it causes to be assigned on its left.

(E2) Membership in one of the above three classes is an inherent property of each morpheme; there are, however, certain morphological processes which change class membership for a given morpheme in a given context. In particular, a given morpheme may trigger the rule of Deaccentuation (3) which deaccents the entire preceding string. Whether or not a morpheme triggers Deaccentuation is in part specified in its lexical entry, and in part predictable morphologically.

(3) Deacc. \( V \rightarrow [-h] / _{-} X + C_{O} V \)

Below we illustrate briefly how systems with these general properties account for the accent pattern of several branches of Indo-European that most directly continue the original. In the concluding part of the paper we sketch the main outlines of the historical evolution of this system and comment on parallels between the system we have found and a number of accentual systems discovered by other scholars in non-Indo-European languages.

Sanskrit

Our exposition begins with Sanskrit as this language according to our reconstruction reflects most faithfully the original Indo-European accentual system.

In Sanskrit, nouns of the athematic declension fall into the two main accentual types illustrated in (4):

\[
\begin{array}{lll}
\text{(i) movable:} & \text{(ii) fixed:} & \\
\text{A} & \text{B} & \text{C} \\
\text{Dat. Sg. (and other weak endings)} & \text{Acc. Sg. (and other strong endings)} & \text{(Deaccenting)} \\
\text{duhitr} + e & \text{duhitār} & \text{dyhitār} \\
\text{pad} + e & \text{pad} + am & \text{pad} \\
\text{bhrāt} + e & \text{bhrāt} + am & \text{bhrāt} \\
\text{marūt} + e & \text{marūt} + am & \text{marūt} \\
\text{gāv} + e & \text{gāv} + am & \text{gāv} \\
\end{array}
\]

Stems of the movable class (see (4i)) have no accent in their underlying representation, while stems of the fixed class (see (4ii)) have an inherent accent on the appropriate syllable. The behavior of the three classes of case forms can then be characterized as follows. The suffixes of the dative and the other so-called weak cases (column A in (4)) are themselves inherently accented. When added to unaccented stems, the suffixes contain the only underlying accent of the word and therefore get the word accent by BAP. When these endings are added to accented stems, the stem accent, being now the leftmost accent of the word, prevails over the suffix accent, again by the
BAP. The suffixes of the strong cases, column B, are inherently unaccented. This is a natural assumption since the strong case endings are in fact never accented on the surface. It furthermore makes possible a simple formulation of the Metatony rule (2), the rule that assigns presuffixal accent in strong cases. For this rule can then be conditioned completely generally by any unaccented inflectional case suffix. It correctly derives the string of duhitaram and similar forms in the movable class.

Column C in (4) comprises only the Vocative. If we suppose that this case is subject to the Deaccentuation rule (3), then the BAP accents the initial syllable as required.

Further support for our analysis is offered by the small class of inherently unaccented stems in which the accent alternates between the desinence in the weak cases and the initial syllable in the strong cases, e.g., the noun puminus (5). The proposed treatment of strong case suffixes as unaccented makes it possible to integrate these stems into the system in a natural way: we need only suppose that they do not undergo Metatony. They will then emerge unaccented at the point at which the BAP applies, and therefore get the word accent on their first syllable.

(5)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weak</td>
<td>pum-č</td>
<td>puminus-č</td>
<td>puman</td>
</tr>
<tr>
<td>Strong</td>
<td>purr ams-am</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deaccenting</td>
<td></td>
<td></td>
<td>&quot;man&quot;</td>
</tr>
</tbody>
</table>

The nouns in (6) exemplify nouns that take the theme vowel +a+ before the case endings, in contrast to pad and marut, which are athematic. If we further assume that theme vowels are normally accented and that, moreover, åva is inherently accented whereas deva is not inherently accented, the surface forms are obtained without complication by the application of the Basic Accentuation Principle.

(6)

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Accentual class in the declension</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
<td>h</td>
</tr>
<tr>
<td>N., A. Dual</td>
<td>pad+au &quot;foot&quot;</td>
<td>marut+au &quot;wind&quot;</td>
<td>dev+au &quot;god&quot;</td>
<td>asv+au &quot;horse&quot;</td>
<td></td>
</tr>
<tr>
<td>Loc. Sg.</td>
<td>pad+i</td>
<td>marut+i</td>
<td>dev+e+ dev+a</td>
<td>asv+e+ asv+a+i</td>
<td></td>
</tr>
<tr>
<td>Voc. Sg.</td>
<td>pad</td>
<td>marut</td>
<td>dev+a</td>
<td>asv+a</td>
<td></td>
</tr>
</tbody>
</table>

This analysis sheds new light on the accentuation of nominal compounds in Indo-European. For bahuvrihi compounds and a large, morphologically well-defined subclass of tātpurusa compounds, the oldest accent pattern, amply attested in Vedic and confirmed by Greek (Kuryłowicz 1968, 55ff.), is illustrated in (7):
When the first member of the compound is inherently accented on some syllable—the first, as in (7a), the second, as in (7b), and so forth—the whole compound is accented on that syllable. When the first member of the compound is inherently unaccented, the whole compound is accented on the inherently accented syllable of the second member. But this basic accent pattern of compounds follows straight from the BAP.1

Verb inflection works like noun inflection, with this difference: that here Metatony is not predictable for all unaccented suffixes, but only for some, which must therefore be designated in the lexicon as triggering this rule.

It is true of all the Indo-European languages we have studied that the derivational system presents a more elaborate set of accentual categories than the inflectional systems. This tendency is manifested in Sanskrit in the following two characteristic ways. First, while inflectional suffixes normally do not trigger Deaccentuation of the stem by rule (3), the Vocative, as we mentioned, is a lone exception in the inflectional system here), derivational suffixes quite commonly do. Secondly, while Metatony is partly predictable for inflectional suffixes (unaccented case endings always trigger it), this property is entirely unpredictable for unaccented derivational suffixes. Therefore, derivational suffixes may be accented, unaccented metatonizing, and unaccented non-metatonizing, and, independently of this, they may or may not trigger Deaccentuation of the stem to which they are added. Of the six combinations that are thus logically possible, as least four are actually instantiated in the language, as shown in (8).

(8) All forms given in g. pl.; accented stems in left column, unaccented stems in right column.

Class II - Suffixes that trigger Deaccentuation of the base.

la) Metatonizing: abstract noun -ta:

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>bandh+u-tā+nam→bandhūtānām</td>
<td>agn+i+tā+nam→agnītānām</td>
<td></td>
</tr>
<tr>
<td>&quot;relationship&quot;</td>
<td>&quot;firehood&quot;</td>
<td></td>
</tr>
</tbody>
</table>

lb) Accented: adjective forming -in:

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>h</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>asv+in+i+nam→asvīnīnām</td>
<td>pakṣ+in+i+nam→pakṣīnīnām</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;having horses&quot;</td>
<td>&quot;having wings&quot;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Class I - Suffixes that do not trigger Deaccentuation of the base.

1a. Metatonizing (this type attested only in inflectional suffixes).

1b. Accented: noun declension theme -a:

\[
\begin{align*}
\text{h} & \quad \text{h} & \quad \text{h} \\
\text{ašv} + \text{a} + \text{nām} & \rightarrow \text{ašvānām} & \text{"horse"} \\
\text{dev} + \text{a} + \text{nam} & \rightarrow \text{devānām} & \text{"god"}
\end{align*}
\]

1c. Unaccented: possessive adjective -vant; fem. -I:

\[
\begin{align*}
\text{h} & \quad \text{h} \\
\text{ašv} + \text{a} + \text{vant} + \text{i} + \text{nam} & \rightarrow \text{ašvavatinam} \\
\text{pad} + \text{vant} + \text{i} + \text{nam} & \rightarrow \text{padvatinam}
\end{align*}
\]

"having horses" "having feet"

We propose that Proto-Indo-European had in essence the same system, and that similar systems, though modified in various important respects, are retained in the other Indo-European languages that preserve the inherited accentual mobility. We will briefly trace the development in two Indo-European branches, Baltic and Slavic.

**Lithuanian**

As a starting point, let us return to the athematic paradigms set out in (4).

Metatony, the rule that accents the stem-final syllable before unaccented suffixes, is in practice disjunctive in relation to the initial accent assigned by the BAP—that is, BAP catches everything which has for any reason failed to undergo Metatony, as for example the pumams type (5). Suppose now that Metatony ceased to apply before case endings. What would happen? The result would be that a particular subset of stems in a particular subset of cases, namely inherently unaccented stems in strong cases, would acquire initial accent. But this was in fact the key innovation that led to the distinctive Balto-Slavic accent pattern, illustrated by the Lithuanian forms in (9):

\[
\begin{array}{c|c|c|c|c}
\text{A} & \text{B} \\
\text{Acc. Sg.} & \text{Gen. Sg.} \\
\text{strong case} & \text{weak case} \\
(1) \text{movable} & \text{dukteri} & \text{"daughter"} & \text{dukters} (\text{kas}) \\
(2) \text{fixed} & \text{móteri} & \text{"mother" (archaic)} & \text{móter(e)s}
\end{array}
\]

The same initial−desinential accent mobility was extended not only to inherently unaccented athematic nouns but to inherently unaccented nouns of all declensions, as shown in (10):

\[
\begin{align*}
\text{strong} & & \text{weak} \\
\text{movable (class 3)} & \text{Lith. gálv + a "head"} & \text{gálv + os} & \text{golov + ámi} \\
& \text{R. gólov + u} & & \\
\text{fixed (class 1)} & \text{Lith. várn + a "crow"} & \text{várn + os} & \text{vorón + ami}
\end{align*}
\]
At this point the two branches diverge onto very different paths. Baltic (as whose representative we take the archaic system of Lithuanian) underwent two major innovations. First, it developed a contrast between a falling ("acute") and rising ("circumflex") accent on long vowels and diphthongs (including tautosyllabic vowel + nasal and vowel + liquid sequences). The intonation difference reflects an older length difference. We shall return in the concluding section to a discussion of the history of these changes.

In Lithuanian, unlike Sanskrit and Slavic, the accent may strike not only vowels but also post-vocalic sonorants, provided that they belong to the same syllable as the immediately preceding vowels. Any such sonorant segment that may receive the accent will be referred to here as a mora. It is true of both underlying and surface forms that only moras may bear the accent. We assume here that the traditional rising ("circumflex") pitch corresponds to the case where the accent strikes the last sonorant in a tautosyllabic sequence beginning with a vowel, and we assume that the traditional falling ("acute") pitch corresponds to the case where the accent strikes any other sonorant in the sequence. We illustrate this in (11).

(11)  

<table>
<thead>
<tr>
<th>designation</th>
<th>standard spelling</th>
<th>abstract representation</th>
<th>historical source</th>
</tr>
</thead>
<tbody>
<tr>
<td>falling (&quot;acute&quot;)</td>
<td>V VR</td>
<td>h V V VR</td>
<td>*V *VR</td>
</tr>
<tr>
<td>rising (&quot;circumflex&quot;)</td>
<td>V ~V</td>
<td>h V ~VR</td>
<td>*V *VR</td>
</tr>
</tbody>
</table>

Since syllable structure may change in the course of a derivation, such changes may be accompanied by purely mechanical changes in the word accent.

(12)  

<table>
<thead>
<tr>
<th>Inf.</th>
<th>3. p. past</th>
</tr>
</thead>
<tbody>
<tr>
<td>/vir/</td>
<td>/vir + ti/ /vir + e/</td>
</tr>
<tr>
<td>/mir/</td>
<td>/mir + ti/ /mir + e/</td>
</tr>
</tbody>
</table>

When /mir/ is followed by a vocalic ending, the r is no longer a mora, and can therefore no longer be accent-bearing. The accent moves to the vowel on the left.

These changes in syllabic structure interact with a rule of Lithuanian which lengthens (geminates) accented e and a:

<table>
<thead>
<tr>
<th>/ar/</th>
<th>/ar + ti/ /ar + e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>/tar/</td>
<td>/tar + ti/ /tar + e/</td>
</tr>
</tbody>
</table>

"cook"  
"die"  
"plow"  
"speak"
These examples illustrate how misleading the orthographic representation of Lithuanian accent can be: the orthographic change of acute to circumflex in ārtis ~ āria corresponds to no change in the position of the accent but simply to the resyllabification of the r with the vocalic suffix in āria, which turns the accented a into the final mora of its syllable. On the other hand, the circumflex in tārtis designates the original accent on the r; while in āria the accent is shifted onto the vowel since the resyllabified r is no longer syllable final.

The second major innovation in Baltic was the deletion of an accent from the last mora of a stem if the first mora of the suffix was accented—resulting in the forward shift known as deSaussure's Law (DSL). The salient features of the accentual pattern of the Lithuanian declension are illustrated in (13).

(13) Accentual classes in the declension

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>n. sg.</td>
<td>h</td>
<td>n. sg.</td>
<td>h</td>
<td>n. sg.</td>
</tr>
</tbody>
</table>
|   varn + a "crow" | rank + a "hand" | galv + a "head" | barzd + a "bearc"
|   h | h   | h   | h   |
| g. sg. | h   | g. sg. | h   | g. sg. | h   |
|   varn + oos | rank + oos | galv + oos | barzd + oos |
|   h | h   | h   | h   |
| d. sg. | h   | d. sg. | h   | d. sg. | h   |
|   varn + ai | rank + ai | galv + ai | barzd + ai |
|   h | h   | h   | h   |
| l. sg. | h   | l. sg. | h   | l. sg. | h   |
|   varn + a | rank + a | galv + a | barzd + a |

The four columns represent the four accentual stem classes, and the four horizontal rows represent the four types of accentual suffix classes.

Paradigms I and II represent original inherently accented stems; paradigms III and IV represent original unaccented stems that get initial accent in strong cases. Within each, DSL splits off a subclass, II and IV, representing stems which were historically short, and are synchronically accented on the last mora.

It is clear that as a result of these changes some reanalysis has taken place in Baltic. After much experimentation we have tentatively concluded that the following system is most adequate: Morphemes are marked with a feature + high (h) on some mora, and a + high may further be either dominant +h or recessive -h. The location of the +h corresponds to the original length, and the dominant +h's correspond to the old inherently accented syllables. The BAP is then revised (for Lithuanian) as follows:

(14) The first dominant +h vowel in a word is accented; if there is none, the first +h is accented.
Accordingly, in the stems of classes I and II the accented vowels are dominant, whereas they are recessive in the stems of nouns of classes III and IV. (In this analysis every morpheme has one \(+h\) mora.) Case suffixes can also be dominant, e.g. \(-a\) (n. sg.) \(+h\)

\[ h \quad h \]

and \(-oos\) (g. sg.), or recessive, e.g. \(-a\) (i. sg.) and \(-al\) (dat. sg.). The crucial difference between the members of each pair of paradigms is whether or not the accent falls on the first mora: this fact determines the application of DSL. The underlying representations of the noun case forms illustrated in (13) are as shown in (15), with * indicating dominant h's.

\[(15)\]

\(\begin{array}{cccc}
\text{n. sg.} & \text{varn} + a & \text{rank} + a & \text{galv} + a & \text{barzd} + a \\
\text{g. sg.} & \text{varn} + oos & \text{rank} + oos & \text{galv} + oos & \text{barzd} + oos \\
\text{d. sg.} & \text{varn} + ai & \text{rank} + ai & \text{galv} + ai & \text{barzd} + ai \\
\text{i. sg.} & \text{varn} + a & \text{rank} + a & \text{galv} + a & \text{barzd} + a \\
\end{array}\)

It is obvious from an inspection of forms in (15) that the modified Basic Accentuation Principle by itself determines the surface word accent in all forms except the n. sg. and i. sg. of classes II and IV. These are just the case forms where \(+h\) resonants appear on two consecutive syllables. In such cases deSaussure's Law applies and deaccents the first of the two vowels:3

\[(16)\]

desoussure's Law \( V \rightarrow [-h] / C_0 [+h] \)

The accentual patterns in derived nouns are somewhat more varied than those just encountered in the declension because some derivational suffixes, but no case endings, trigger the Deaccentuation rule (3). Incidentally in Lithuanian it must be assumed that the Deaccentuation rule eliminates the diacritic feature * which marks the high pitch on the morpheme in question as dominant.

In (17) we have illustrated a suffix with inherent dominant accent which triggers Deaccentuation; such a suffix will form nouns of accentual class I, if its accent is on a prefinal mora: \(eln + ien + a\) \(\xrightarrow{\text{Deacc}}\) \(eln + ien + a\) \(\xrightarrow{\text{BAP}}\) \(eln + ien + a\)

\[(17)\]

Dominant accent; deaccenting

\(\begin{array}{ll}
eln + ien + a & \text{I "venison"} \quad \text{cf. } \text{eln} + ias \text{ I "stag"} \\
klek + ien + a & \text{I "rabbit meat"} \quad \text{cf. } klek + is \text{ II "rabbit"} \\
ooz + ien + a & \text{I "goat meat"} \quad \text{cf. } ooz + yys \text{ III "goat"} \\
vilk + ien + a & \text{I "wolf meat"} \quad \text{cf. } vilk + as \text{ IV "wolf"} \\
\end{array}\)
The suffix -ena illustrated in (18) forms nouns belonging to the accent class III regardless of base accent:

(18) No dominant accent; deaccenting

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>h</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>klaun + en + a</td>
<td>&quot;marten pelt&quot;</td>
<td>(cf. klaun + ə I &quot;marten&quot;)</td>
<td>klaun + en + al</td>
</tr>
<tr>
<td>vers + en + a</td>
<td>&quot;calf skin&quot;</td>
<td>(cf. vers + is II &quot;calf&quot;)</td>
<td>vers + en + al</td>
</tr>
<tr>
<td>oož + en + a</td>
<td>&quot;goat skin&quot;</td>
<td>(cf. oož + yys III &quot;goat&quot;)</td>
<td>oož + en + al</td>
</tr>
<tr>
<td>vilk + en + a</td>
<td>&quot;wolf skin&quot;</td>
<td>(cf. vilk + as IV &quot;wolf&quot;)</td>
<td>vilk + en + al</td>
</tr>
</tbody>
</table>

Nouns of this type exhibit desinential accent before dominant desinences, (as in the left column of (18) where nom. sg. forms are quoted) but preserve their original accent of the base form before recessive desinences (as in the right column of (18) where dat. sg. are quoted). The modified Deaccentuation rule derives all examples correctly as long as we assume that the suffix -en- is itself unaccented. This is shown in (19).

(19) Deacc. BAP

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th>h</th>
<th>h</th>
<th>h</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>klaun + en + a</td>
<td>klaun + en + al</td>
<td>vilk + en + a</td>
<td>vilk + en + al</td>
<td></td>
<td></td>
</tr>
<tr>
<td>klaun + en + a</td>
<td>klaun + en + al</td>
<td>vilk + en + a</td>
<td>vilk + en + al</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In most of the languages we are concerned with, derivational suffixes are typically deaccenting; as a result, the accentuation of the words that these suffixes form is independent of the base accent. Non-deaccenting derivational suffixes are relatively uncommon, though they tend to include some very productive ones. Non-deaccenting derivational suffixes are in a sense the analog to "stress-neutral" derivational suffixes in English-like stress systems, e.g. English -er, -ish. This situation is well illustrated by Lithuanian. There appears to be only one clear case, to our knowledge, of a non-deaccenting suffix, that is also dominant and accented; this is the very productive suffix -ininkas, corresponding to English -er. It forms class I nouns from nouns of class I and II, and class II nouns from nouns of classes III and IV. We illustrate this in (20):

(20) Dominant accent; non-deaccenting

<table>
<thead>
<tr>
<th></th>
<th>h</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>auks + inink + as</td>
<td>&quot;goldsmith&quot;</td>
<td>(cf. auks + as I &quot;gold&quot;)</td>
</tr>
</tbody>
</table>
This pattern follows directly from the assumption a) that stems of classes III and IV have recessive accented vowels, and b) that the suffix 

\[ \text{mink + as} \]

has a dominant accent on the last mora—i.e., on the \( h \).

The fourth possible combination whose existence our system allows is a suffix which has no dominant accent and does not trigger deaccentuation. Such a suffix would produce words with base accent in all case forms when added to dominant bases. When added to non-dominant bases such suffixes would form words with desinential accent if the desinence has dominant accent (weak cases), and with base stress if the desinence does not have dominant accent (strong cases). In other words, if the base does not have dominant accent, the derived noun would belong to class III; if the base has dominant accent, the derived noun would belong to class I. Modern Lithuanian appears to lack suffixes of this fourth type. But at least one suffix had exactly this predicted fourth pattern in the XVIth century dialect of Dauksa, whose accentology was studied by Skardzius (1935). At expected, this is a very frequent and productive suffix, the adjective-forming 

\[ \text{isk} \] -\[ \text{isk} \] -\[ \text{isk} \] -\[ \text{isk} \]. Skardzius (pp. 157-9) cites the following class I derivatives from bases with inherent dominant accent (class I or II):

<table>
<thead>
<tr>
<th>Base</th>
<th>Meaning</th>
<th>Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>dvālis kas</td>
<td>&quot;spiritual&quot;</td>
<td>(dvāsia II &quot;spirit&quot;)</td>
</tr>
<tr>
<td>gediškas</td>
<td>&quot;shameful&quot;</td>
<td>(gēda I &quot;shame&quot;)</td>
</tr>
<tr>
<td>kuniškas</td>
<td>&quot;corporal&quot;</td>
<td>(kuńas I &quot;flesh&quot;)</td>
</tr>
<tr>
<td>moteris kas</td>
<td>&quot;feminine&quot;</td>
<td>(moteris I &quot;woman&quot;)</td>
</tr>
<tr>
<td>pagonis</td>
<td>&quot;pagan&quot; (adj.)</td>
<td>(pagonis, pagonis I &quot;pagan&quot;)</td>
</tr>
<tr>
<td>vyriskas</td>
<td>&quot;masculine&quot;</td>
<td>(vyras I &quot;man&quot;)</td>
</tr>
<tr>
<td>žemiškas</td>
<td>&quot;earthly&quot;</td>
<td>(žėme II &quot;earth&quot;)</td>
</tr>
<tr>
<td>bróliskas</td>
<td>&quot;brotherly&quot;</td>
<td>(brólis I &quot;brother&quot;)</td>
</tr>
<tr>
<td>Žydiškas</td>
<td>&quot;Jewish&quot;</td>
<td>(Žydas II &quot;Jew&quot;)</td>
</tr>
</tbody>
</table>

These contrast with the following class III derivatives from bases with no inherent dominant accent (classes III and IV) which also are cited by Skardzius:

<table>
<thead>
<tr>
<th>Base</th>
<th>Meaning</th>
<th>Stem</th>
</tr>
</thead>
<tbody>
<tr>
<td>kūnig + išk + as</td>
<td>&quot;priestly&quot;</td>
<td>(kūningas III &quot;priest&quot;)</td>
</tr>
<tr>
<td>diėv + išk + as</td>
<td>&quot;divine&quot;</td>
<td>(diėvas IV &quot;god&quot;)</td>
</tr>
<tr>
<td>Lithuanian</td>
<td>English</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>ángeliškas</td>
<td>&quot;angelic&quot;</td>
<td></td>
</tr>
<tr>
<td>mergiškas</td>
<td>&quot;girlish&quot;</td>
<td></td>
</tr>
<tr>
<td>suniškas</td>
<td>&quot;filial&quot;</td>
<td></td>
</tr>
<tr>
<td>vėliniškas</td>
<td>&quot;childish&quot;</td>
<td></td>
</tr>
<tr>
<td>žmogiškas</td>
<td>&quot;ghostly&quot;</td>
<td></td>
</tr>
<tr>
<td>krikščioniškas</td>
<td>&quot;Christian&quot;</td>
<td></td>
</tr>
</tbody>
</table>

In some cases, the accent class of the base word has changed in modern Lithuanian; for example, Христионис, IV in Dauksa's dialect, is now Христионис, belonging to class I.

The suffix -iskas has in modern Lithuanian assumed a different accentual behavior.

**Russian**

We discover most of these features also in the accentual system of Russian. We begin by examining the accentual patterns in the declensional paradigms. Just like Sanskrit and Lithuanian, Russian has accented as well as unaccented morphemes and is subject to the BAP (I). Like Sanskrit nouns, Russian nouns may be subject to Metatony; however, unlike Sanskrit, Russian requires that the suffix triggering Metatony be itself accented and, moreover, whether Metatony will or will not take place before an accented suffix is a purely Idiosyncratic property of a given stem. (We shall comment below on the fact that Metatony in the declension is apparently a relatively recent development in Russian.) Like the other languages reviewed, Russian has "strong" and "weak" case endings. When affixed to accentless stems "weak" case endings appear always accented, e.g., the fem. gen. sg. -y/-i: skovorody "frying pan" and golovy "head"; while "strong" case endings may appear accented with some stems, and unaccented with other stems, e.g., the fem. acc. sg. -u: skovoroda "frying pan" but golovu "head." We shall assume that the "strong" case endings are underlyingly accentless and the "weak" case endings are underlyingly accented.

The "strong" (unaccented) endings are:

(21) In the singular, the acc. fem. -u - suffix

<table>
<thead>
<tr>
<th>Case Suffixes</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>All case suffixes of the non-fem. declension, except the -u suffix</td>
<td></td>
</tr>
<tr>
<td>All the case suffixes of fem. consonantal declension, except the loc. -i suffix</td>
<td></td>
</tr>
</tbody>
</table>

All other case endings are "weak", i.e. inherently accented. To account for the instances where "strong" case endings appear accented we postulate the Post-stem Accentuation Rule (PAR).
In the majority of nouns PAR applies to all case forms. The effects of this rule, however, are detectable only in forms with "strong" (inherently accentless) endings; since "weak" endings are inherently accented, PAR applies to them vacuously. As we shall see below, there is a substantial number of (mainly feminine) nouns in Russian where the PAR does not apply across the board, but only to case forms in one of the two numbers. Thus, whether or not PAR applies is an idiosyncratic property of each "strong" case form.

We digress briefly to comment on the fact that with the introduction of PAR into the grammar we have two distinct ways of accenting a suffix; either by supplying the accent in the underlying representation of the suffix, or by letting PAR apply to the suffix. It might be objected that this is a useless redundancy which should be eliminated. The obvious move would be to postulate that all suffixes are unaccented in their underlying representation, and that whenever a suffix appears accented in the output this is the result of the application of PAR. If we adopt this solution we would have three classes of instances to consider:

a) suffixes that are always accented; i.e., where PAR applies in all contexts. These are the so-called "weak" case endings. (cf. (21)).

b) suffixes that are never accented; i.e., that are everywhere marked as being exceptions to PAR. (Modern Russian accidentally does not provide an example of this type of suffix among its case endings. There are such suffixes among the derivational suffixes, e.g., yë in mal-yëmÌl "little fellows" - that the gap in the declension is accidental is shown further by the fact that such unaccented case endings existed in Old Russian and still exist in other Modern Slavic languages, such as Czech and Serbo-Croatian, i.e., the vocative sq. -ë.)

c) suffixes that are accented after some stems and accentless after other stems; i.e., that are marked as exceptions to PAR in certain context only. These are the "strong" case endings. (cf. (21)).

The problem is clearly with regard to the first class of suffixes. Since they are always accented on the surface, they would be underlyingly represented as unaccented only to be accented by PAR. If we allowed this solution then we would have no way to rule out such obviously absurd moves as the following. Since English has a rule which voices final obstruents in certain morphological environments (cf., device-devise; bath-bathe) we could then represent all word final obstruents as voiceless, and capture the distinction between voiced and voiceless obstruents by letting the
former always undergo the voicing rule. Straightforward considerations of descriptive
economy will eliminate a solution such as the one just sketched, and they will also
exclude the proposal to represent all suffixes as unaccented. It hardly needs saying
that the same considerations will also rule out the proposal to represent all suffixes
as underlyingly accented and to subject some of them to a deaccentuation rule, which,
incidentally, would have to be distinct from the rule deaccenting stems before certain
suffixes that functions in Russian much as in other Indo-European languages.

The Deaccentuation Rule (3) applies in Russian both in the inflection (see (24)
below) as well as in the derivation (see (35) below). It is thus an essential component
of the machinery determining the accentuation of Russian words.

In addition to the Deaccentuation Rule (3), Russian also makes use of a Metatony
Rule, which retracts the accent one syllable (in a few special cases, two syllables)
towards the beginning of the word. We give here the Metatony rule in its barest
essentials only:

(23) Metatony

\[ V \rightarrow [+h] / _{C_o} + _{C_o} [+h]. \]

It should be noted that Metatony accents a vowel only if it precedes an accented suffix.
Metatony must be ordered after PAR since, as we shall see below, suffixes accented
by PAR trigger Metatony.

The various components just reviewed which play a role in determining the
accentuation of the declension pattern interact in various ways to produce in a large
number of nouns a striking surface contrast in the accentuation of the singular vs.
plural. As we shall show below in a great many nouns - though by no means in all
nouns - if the singular forms have suffixal stress, the plural forms have stem stress;
and vice versa. It is of considerable interest, moreover, that during the last two
hundred years this surface accentual contrast has noticeably increased. 4

We begin our review of the different accentual patterns found in the declension
by examining the patterns exhibited by inherently accented stems. As will be recalled
in nouns with accented stems the accentuation of the suffix will normally not affect
the surface accent placement, for SAP insures that the accent will remain on the same
syllable of the stem regardless of the nature of the case suffix, as illustrated below.

<table>
<thead>
<tr>
<th></th>
<th>singular</th>
<th>plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>malín + y (gen.)</td>
<td>malín + am (dat.)</td>
</tr>
<tr>
<td>strong</td>
<td>malín + u (acc.)</td>
<td>malín + y (nom.)</td>
</tr>
</tbody>
</table>
There is, however, one striking exception to the preceding. In a large and growing class of masculine nouns, many of them clearly of foreign origin, the plural forms are subject to Deaccentuation (3). Since in these nouns all singular case endings are unaccented ("strong"), whereas all plural case endings are accented ("weak") (these nouns take the pl. nom. suffix -a) the result of Deaccentuation applying in the plural is a declensional pattern such as that illustrated in (24) where the singular with fixed stress on the stem contrasts with the plural with stress is fixed on the case ending.

(24)

<table>
<thead>
<tr>
<th></th>
<th>Singular</th>
<th>Plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>professor + a</td>
<td>not attested</td>
</tr>
<tr>
<td>strong</td>
<td>not attested</td>
<td>professor + ám</td>
</tr>
</tbody>
</table>

This is the first of several examples that will be discussed of the tendency of the language to extend its favorite accentual pattern to new classes of nouns.

Turning now to nouns with accentless stems we recall that surface stress placement will be largely governed by the accentual characteristics of the suffixes detailed in (21) above. It is immediately obvious from (21) that in the non-feminine declension there will be a stress contrast between the singular and the plural: in the singular, case endings are largely unaccented, in the plural the case endings are almost exclusively accented. (The same is true of the feminine consonantal declension, but this paradigm will be disregarded in our discussion below).

In the feminine -a declension the situation is quite different: here the majority of case suffixes are accented ("weak"). Thus, in the feminine -a declension there was originally no accentual contrast between the singular and the plural. But as we shall see this contrast between the accentuation of the singular and that of the plural was extended to nouns of this class too.

The original situation (and the one continuing the Indo-European "mobile" paradigm) is found in nouns like those illustrated in (25):

(25)

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>f. golov + y (gen.)</td>
<td>golov + ám (dat.) &quot;head&quot;</td>
</tr>
<tr>
<td></td>
<td>m. (na) zub + ū (loc.)</td>
<td>zub + ám (dat.) &quot;tooth&quot;</td>
</tr>
<tr>
<td></td>
<td>n. not attested</td>
<td>uš + ám (dat.) &quot;ear&quot;</td>
</tr>
<tr>
<td>strong</td>
<td>f. golov + u (acc.)</td>
<td>golov + y (nom.)</td>
</tr>
<tr>
<td></td>
<td>m. zúb + u (dat.)</td>
<td>zúb + y (nom.)</td>
</tr>
<tr>
<td></td>
<td>u. úx + u (dat.)</td>
<td>úš + i (nom.)</td>
</tr>
</tbody>
</table>
Here in the "weak" cases the endings are accented whereas in the "strong" cases the BAP assigns the accent to the first syllable of the word.

It should be noted that many masculine nouns and the overwhelming majority of neuter nouns have in the nom. pl. in place of the "strong" (unaccented) ı/y suffix, the "weak" (accented) -a suffix. Such words provide no examples of "strong" case endings in the plural. This fact will be of significance in our discussion below.

Alongside the "mobile" paradigm, feminine nouns provide examples also that mirror the old "oxytone" accentual pattern:

(26)

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>f.  kocerg + í (gen.)</td>
<td>kocerg + ám (dat.) &quot;stove poker&quot;</td>
</tr>
<tr>
<td></td>
<td>m. (na) post + ū (loc.)</td>
<td>post + ám (dat.) &quot;guard&quot;</td>
</tr>
<tr>
<td></td>
<td>n. not attested</td>
<td>božestv + ám (dat.) &quot;divinity&quot;</td>
</tr>
<tr>
<td>strong</td>
<td>f.  kocerg + ú (acc.)</td>
<td>kocerg + í (nom.)</td>
</tr>
<tr>
<td></td>
<td>m. post + ú (dat.)</td>
<td>post + ý (dat.)</td>
</tr>
<tr>
<td></td>
<td>n. božestv + í (dat.)</td>
<td>not attested</td>
</tr>
</tbody>
</table>

Here the accent falls on all case endings, "strong" as well as "weak". In order to account for such "oxytone" paradigms we shall assume that nouns in this class are subject to PAR (22). In the majority of nouns, PAR applies to all case forms. The effects of this rule, however, are detectable only in forms with "strong" (inherently accentless) suffixes; since "weak" endings are inherently accented, PAR applies to them vacuously.

There is a substantial number of (mainly feminine) nouns in Russian where the PAR does not apply across the board, but only to case forms in one of the two numbers. In (27) below we illustrate nouns where the PAR applies only in the singular.

(27)

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>f. skovorod + ý (gen.)</td>
<td>skovorod + ám (dat.) &quot;pan&quot;</td>
</tr>
<tr>
<td></td>
<td>m. not attested</td>
<td>kon + jám (dat.) &quot;steed&quot;</td>
</tr>
<tr>
<td></td>
<td>n. not attested</td>
<td>pleč + ám (dat.) &quot;shoulder&quot;</td>
</tr>
<tr>
<td>strong</td>
<td>f. skovorod + ú (acc.)</td>
<td>skovorod + ý (nom.)</td>
</tr>
<tr>
<td></td>
<td>m. kon + jú (dat.)</td>
<td>kon + l (nom.)</td>
</tr>
<tr>
<td></td>
<td>n. pleč + ú (dat.)</td>
<td>pleč + l (nom.)</td>
</tr>
</tbody>
</table>

[accent supplied by PAR]

In (28) we illustrate nouns where PAR applies in the (nom.) plural only. There are no feminine nouns in this category, and according to Zaliznjak (1967), only three
neuter nouns: ochki "glasses", mazly "daubers", treply "jabberers". On the other hand, there are about 80 masculine nouns in this class.

(28)

\[
\begin{array}{llll}
\text{sg.} & \text{pl.} \\
\text{weak} & m. & \text{sad} + \text{u} (\text{loc.}) & \text{sad} + \text{am} (\text{dat.}) "garden" \\
\text{strong} & m. & \text{sád} + \text{u} (\text{dat.}) & \text{sad} + \text{y} (\text{nom.}) \\
\end{array}
\]

[accent supplied by PAR]

The splitting of PAR into two subcases, one applying in the singular only and the other restricted to the plural, is yet another development of Russian that manifests the tendency of language to differentiate the accentuation of singular forms from that of the plural forms. The most striking manifestation of this tendency is in the increasingly important role that the Metatony rule (23) has been playing during the last 150 years in determining the accentuation of nouns. In the declension (with a handful of totally ad hoc exceptions) Metatony is restricted to plural forms. The effect of the Metatony rule is to accent the vowel preceding an accented suffix; in terms of its surface effects, Metatony results in a retraction of the accent by one syllable towards the beginning of the word. By restricting Metatony to the plural the language assures that there will be an accentual contrast between singular and plural forms in nouns that either are subject to PAR or that like the feminine -a declension nouns have inherently accented endings in the singular. We illustrate these contrasts in (29).

(29)

\[
\begin{array}{llll}
\text{sg.} & \text{pl.} \\
\text{weak} & f. & \text{kolbas} + \text{y} (\text{gen.}) & \text{kolbás} + \text{am} (\text{dat.}) "sausage" \\
& m. & \text{žen} + \text{y} (\text{dat.}) & \text{žén} + \text{am} (\text{dat.}) "wife" \\
& n. & \text{not attested} & \text{not attested} \\
\text{strong} & f. & \text{kolbas} + \text{u} (\text{acc.}) & \text{kolbás} + \text{y} (\text{nom.}) \\
& m. & \text{žen} + \text{u} (\text{acc.}) & \text{žén} + \text{y} (\text{nom.}) \\
& n. & \text{kazak} + \text{u} (\text{dat.}) & \text{kazák} + \text{i} (\text{nom.}) \\
\end{array}
\]

[accent supplied by PAR and Metatony]

The contrast in accentuation between the singulars and the plurals is almost as striking when the Metatony rule applies to nouns that either do not undergo PAR at all, or undergo it only in the singular or in the plural alone. As the manifestation of the contrast also differs in the masculines and neuters as against the feminines, we shall discuss these two groups separately.

As was pointed out above, in the nonfeminines all singular case endings are "strong" (unaccented), with the exception of the loc. -u ending, which, however, is quite restricted
in its distribution. As a consequence, if PAR does not apply in the singular, we obtain words with uniform initial accent in the singular. In the plural, on the other hand, regardless of gender, all suffixes, except for the nom. -i/y, are "weak" (accented). Hence we get a contrast between singular forms that have initial accent, and plural forms that have suffixal accent (cf. (25), (27), (28)). If we now let plural forms undergo Metatony, the contrast is changed: we get initial accent in the singular contrasting with presuffixal accent in the plural. This contrast, however, is effectively neutralized in nouns with monosyllabic stems; the accentual patterns of such nouns are indistinguishable from nouns with inherently accented stems. Since the number of polysyllabic unaccented stems is relatively small, it is hardly surprising that there is only a handful of nouns that exhibit this accentual pattern, cf.

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak m.</td>
<td>not attested</td>
<td>obod + jam (dat.) &quot;rim&quot;</td>
</tr>
<tr>
<td>n.</td>
<td>not attested</td>
<td>ozër + am (dat.) &quot;lake&quot;</td>
</tr>
<tr>
<td>strong m.</td>
<td>obod + u</td>
<td>not attested</td>
</tr>
<tr>
<td>n.</td>
<td>ozer + u</td>
<td>not attested</td>
</tr>
</tbody>
</table>

Incidentally all of the nonfeminine nouns in this small class take the "weak" -a suffix in the nom. pl. Thus, the question of whether or not these nouns are subject to PAR in the plural does not arise.

The situation is different in the feminine a-declension. Here there are only two "strong" case endings: the -u- ending of the accusative singular, and the -i/y ending of the nominative plural, and these nouns do not take the "weak" -a ending in the nom. pl. Moreover, as was noted above, there is a restriction on the application of PAR in this class of nouns: nouns subject to PAR only in the plural do not exist. Thus, corresponding to the four different treatments with regard to PAR encountered in the nonfeminines, the feminines have only three: a) PAR applies in both singular and plural, b) PAR applies in the singular alone, c) PAR applies neither in the singular or plural. We have already discussed the result of applying Metatony to the nouns in category a) (cf. (28)). It remains for us to investigate the consequences of applying Metatony to nouns in categories b) and c).

When Metatony is applied to nouns subject to PAR in the singular only we get a uniformly accented singular paradigm where the accent always falls on the case ending contrasting with a plural paradigm in which accent falls on the initial syllable in the nominative and on the presuffixal syllable elsewhere:

<table>
<thead>
<tr>
<th></th>
<th>sg.</th>
<th>pl.</th>
</tr>
</thead>
<tbody>
<tr>
<td>weak</td>
<td>sirot + y (gen.)</td>
<td>sirot + am (dat.) &quot;orphan&quot;</td>
</tr>
<tr>
<td>strong</td>
<td>sirot + u (acc.)</td>
<td>sirot + y (nom.)</td>
</tr>
</tbody>
</table>
This pattern is only minimally distinct from that of nouns which undergo PAR in the plural as well as in the singular (cf. (29)). Indeed, for monosyllabic stems the two patterns are indistinguishable. For in monosyllabic stems, initial and presuffixal stress coincide. We could have derived all the right forms of žena in (29) by assuming that it was a noun of the type sirotě (31), with no PAR in the plural, with initial stress (cf. sirotě) assigned in the nominative plural by BAP and presuffixal stress (cf. sirotam) elsewhere in the plural by Metatony. Our reason for considering monosyllabic stems like žena as being representatives of paradigm (29) rather than (31) is the extremely marginal status of the latter in present-day Russian. The noun sirotě is actually the only clear example of it, and it is a variant pattern at that. Although recorded in all major modern sources, the nom. pl. sirotě is regarded by Russian orthoepists as inferior to sirot' (see Kiparsky (1962), p. 223). (Another possible example is the plurale tantum xlopoty "troubles").

There is a strong tendency for plural retraction to be implemented by a single rule, either BAP or Metatony, but not both. Formally, this amounts to a redundancy rule

\[
(32) \quad [+ \text{Metatony}] \Rightarrow [+ \text{PAR pl.}]
\]

to which only sirotč (and perhaps xlopoty) are fragile exceptions.

When Metatony is applied to feminine nouns that are not subject to PAR in the singular, we get accent movement in the singular as well as in the plural. In the singular there will be initial accent in the accusative and suffixal accent elsewhere; the plural will have initial accent in the nominative as it is not subject to PAR, and presufflaxial accent in all other cases. In monosyllabic stems, we will again get the same result whether the nominative plural is subject to PAR or not, since initial and presuffixal accent coincide. As it happens, the pattern is attested only with a handful of monosyllabic stems:

\[
(33) \quad \begin{array}{ccc}
\text{sg.} & \text{pl.} \\
\text{weak} & \text{vod} + \dot{y} \quad \text{(gen.)} & \text{vod} + \dot{am} \quad \text{(dat.)} & \text{"water"} \\
\text{strong} & \text{vod} + \dot{u} \quad \text{(acc.)} & \text{vod} + \dot{y} \quad \text{(nom.)} \\
\end{array}
\]

Given the demonstrably productive redundancy rule (32), we will assume that the nouns in this class are all marked as undergoing PAR in the plural. On this analysis, nom. pl. vod + y is assigned by Metatony, like dat. pl. vod + am and not by BAP, like acc. sg. vod + u. 5

As we have just seen, in nouns with unaccented stems the surface accentuation is governed by three idiosyncratic features which determine whether or not a given form is subject to PAR in the singular, subject to PAR in the plural, subject to Metatony (in the plural only). These three binary features define eight accentual patterns, of which two are excluded by the redundancy rule (32). As there are few exceptions to
the redundancy rule (32), pattern (34\textsuperscript{\circ}), as discussed above, is attested by one or two nouns in the speech of many speakers.

(34)

<table>
<thead>
<tr>
<th></th>
<th>PAR sg.</th>
<th>PAR pl.</th>
<th>Metatony</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>golová, zúb, úxo (25)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>(\beta)</td>
<td>-</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>(\gamma)</td>
<td>sád, mazlo (28)</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>(\delta)</td>
<td>voda (33), obod, ózero (30)</td>
<td>-</td>
<td>+</td>
</tr>
<tr>
<td>(\epsilon)</td>
<td>skovorodá, kón, pleč (27)</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(\zeta)</td>
<td>(sirotá (31))</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(\eta)</td>
<td>kočergá, post, božestvó (26)</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(\theta)</td>
<td>kolbasa, kazák, kolesó (29)</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Comments on the Evolution of the Accentual Patterns in Modern Russian Nouns

Historically the modern accentuation derives from three patterns: the so-called "barytone" pattern with accent fixed on the stem preserved in modern Russian in the accentuation of nouns with accented stems; the so-called "mobile" pattern preserved in modern Russian in the accentuation of nouns such as golová, zúb, úxo (cf. (25), (34\textsuperscript{\circ}) and the so-called "oxytone" pattern preserved in the accentuation of nouns such as kočergá, póst, božestvó ((26), (34\textsuperscript{\eta})). Since we are interested only in nouns with unaccented stems we shall disregard here the "barytone" class. It is self-evident that the distinction between "mobile" and "oxytone" paradigms is reflected by whether or not PAR applies to a noun. Moreover, originally PAR applied across the board to all forms, and the distinction between nouns that are subject to PAR in only one of the two numbers is historically a more recent phenomenon. The same is true of Metatony: originally it played almost no role in the inflectional morphology of the language and its present important rule is of fairly recent date. The sense of evolution of the last two or three centuries, which is carefully documented in Kiparsky (1962), is towards the elimination of accentual "mobility" within a given number and towards the establishment of a contrast of the accent placement in the singular and in the plural.

The development can be seen most clearly in the feminines. Kiparsky (1962) describes it as follows:

"Von den oben aufgezählten 109 Wortern, die heute dem Typ A (our (34\textsuperscript{\theta}) and (34\textsuperscript{\upsilon})) folgen, scheint kein einziges diesen Typ vor Anfang des 19. Jhs. erhalten zu haben... Die grössten Verluste erlitt dabei der II. Typ (our (34\textsuperscript{\eta})), der an den Typ A (our (34\textsuperscript{\theta}) or (34\textsuperscript{\upsilon})), meist durch Vermittlung des Typs B (our (34\textsuperscript{\epsilon})), fast 50 Worter abgegeben hat..."
Starke Einbussen hat auch Typ III (our (34α)) auf Kosten
des A (our (34ξ) or (34θ), here former rather than latter)
erfahren, wobei die Entwicklung entweder III>B>A (our (34α)> (34ξ)> (34θ))... oder III>W>A (our (34α)> (34ξ)> (34θ))...
ging." (p. 229).

The following seven trajectories in the accentual paradigms of inherently un-
accented feminine stems are documented by Kiparsky (p. 229),

1. $\eta > \epsilon$ (e.g. sveča "candle")
2. $\eta > \Theta$ (e.g. kolbasá "sausage")
3. $\eta > \epsilon > \Theta$ (e.g. skolupá "rind")
4. $\alpha > \epsilon$ (e.g. borozdá "furrow")
5. $\alpha > \Theta$ (e.g. vodá "water")
6. $\alpha > \epsilon > \Theta$ (e.g. rosá "dew")
7. $\alpha > \epsilon > \Theta$ (e.g. vesná "spring")

Note that types $\epsilon$ and $\Theta$ both have two historical sources, $\eta$ and $\alpha$, and that $\Theta$
furthermore develops from either source by two different paths. On the other hand,
type $\Theta$ is always derived directly from $\alpha$. How are these observations to be
interpreted?

Taking as a starting point the two original unaccented stem patterns $\alpha$ (golová) and
$\eta$ (kočergá), suppose that the nouns in these two accent classes could undergo one or
both of two accentual changes: (A) the change to the "favored" marking for PAR
([+ PAR sg.] and [-PAR pl.]) and (B) becoming subject to Metatony. Moreover,
assume that(32) is implemented wherever it becomes applicable. The possible
developments can then be represented by means of the tree diagrams below, where
" $\alpha$ sg," and " $\beta$ pl." stand for " $\alpha$ PAR sg," and " $\beta$ PAR pl." respectively.
It will be seen that all possible trajectories within this schema are attested for some class of nouns in Kiparsky's observations. Note in particular that \( \eta \), having turned into \( \theta \) by (B) cannot further undergo (A), since this would be incompatible with (32).

Also, all trajectories observed by Kiparsky fit into the proposed schema. Thus the schema discloses a surprising systematicity in the apparently chaotic accentual history of the Russian declension. The proliferation of new accentual classes is thus an orderly development of the potentialities inherent in the Slavic pattern of accentual mobility.

We noted that there are no nouns with the feature combination \([-\text{PAR} \text{ sg.}, +\text{PAR} \text{ pl.}, -\text{Metatony}]\). From the viewpoint of the present analysis, this gap is synchronically arbitrary. Any synchronically arbitrary gap requires a historical explanation. Note then that the derivation of the modern accent patterns in the feminines just given provides such an explanation. Given the original paradigms and the direction of change, there is no way in which paradigms with the missing feature combination could have arisen.

We turn now to the neuters, where the situation is somewhat different. Since in the neuters there are normally no "strong" case endings in the plural and no "weak" case endings in the singular, the nouns belonging to the original "mobile" paradigm, i.e. to our class \((34 \alpha)\) will exhibit a contrast in the accent place for singulars and for plurals: in the singular we shall have stem (or initial) accent; in the plurals we shall have suffixal accent. If such nouns were subject to Metatony the contrast could only be preserved in polysyllabic stems. Since polysyllabic stems are rather few it should occasion little surprise that this development has affected only a small number of nouns and that among the neuters the original "mobile" class has been preserved in fact to a very large extent.
In the modern language there are only a few neuter nouns that belong to the "oxytone" accentual paradigm (34 7), and several of these have alternants that are subject to Metatony. The majority of neuter nouns that now are regularly subject to Metatony can be shown to have had the "oxytone" pattern at an earlier time. Kiparsky (1962) states that in 21 out of 52 nouns now subject to Metatony he was able to establish their earlier "oxytone" accentuation (p. 252).

In sum, in view of the fact that the language tends towards accentual paradigms where suffixal stress in one number contrasts with stem stress in the other number we should expect that nouns belonging to the original "mobile" paradigm (34 6) will remain unaffected by any changes whereas nouns belonging to the original "oxytone" paradigm (34 7) will undergo change and become subject to Metatony ((34 7) > (34 0)). And this is in fact what Kiparsky found. He notes that most of the 20 nouns that follow the "mobile" paradigm (34 6) in the modern language had the same accentuation in the distant past. 6

As regards the development of pattern (34 0) Kiparsky (1962) observes:

Von den oben untersuchten 52 Fällen, in denen die heutige Schriftsprache ganz oder fast ausschliesslich den IV. Typ (our (34 0)) gebraucht, haben wir in 28 Fällen einen relativ späten Übergang zu diesem Typ feststellen können." (p. 252).

This brings us to the last class of nouns, the masculines. The case suffixes of masculine nouns are almost all "strong" in the singular, like those of the neuters, whereas like those of the feminines the plural cases endings of the masculines consist of one "strong" ending (the nom. pl.) and four "weak" endings. Because of the predominance of monosyllabic stems it is to be expected that masculine nouns with unaccented stems will in general tend not to undergo Metatony as this would obliterate the contrast in accent placement between the singular and the plural. The only class where this would not be true is the "oxytones", but unlike the neuter nouns, masculine nouns have in this instance resisted the development of contrasting accentuations in the two numbers of the paradigm.

The situation was different with respect to the nouns that originally belonged to the "mobile" paradigm. Here the accentual contrast was, of course, almost in place to begin with since almost all singular masculine suffixes are inherently "strong" (unaccented) and plural suffixes are inherently "weak" (accented). The major exception is constituted by the nom. pl. -i/-y suffix which unlike other plural suffixes is "strong" (unaccented). This minor deviation was "liquidated" in quite a number of nouns by letting these nouns undergo PAR in the plural (i.e. (34 6) > (34 0)). Kiparsky (1962) documents this development in the case of several dozen nouns.
Among the nouns that at present follow this pattern there is a considerable number of fairly recent loan words which originally had inherent stem accentuation. This shows once again the strength of the tendency towards contrasting accentuations in the two numbers.

We now turn to the accentual patterns found in the derivational morphology of Russian nouns. As summarized by Red'kin (1971):

"The accentuation of derived nouns of class I depends on the accentuation of the base. If the base word has nonfinal accent in all of its case forms then in the derived noun, the accent will fall on the same syllable of the base. If, however, in even one case form of the base word we find desinential accentuation, then in the derived word the accent will fall either on the base final vowel (i.e., the syllable preceding the derivational suffix) (subclass 1), or on the suffix itself (subclass 2), or on the desinence (of the derived word - MH/PK) (subclass 3)... (p. 48).

"The accentuation of derived nouns of class II does not depend on the accentuation of the base. Among the nouns of class II there are three subclasses. In the case of derived nouns of subclass 1, the accent falls... on the syllable preceding the derivational suffix, in subclass 2 on the suffix itself, and in subclass 3, on the desinence." (p. 52).

In terms of the preceding discussion the difference between classes I and II is that the derivational suffixes in I are not deaccenting, so that stems always preserve the inherent accent before them. As a result accented stems exhibit different accentual behavior in the derivation than stems without inherent accent. In class II, on the other hand, the derivational suffixes are deaccenting so that accented and unaccented stems are treated alike. Consider now classes I,1; I,2; I,3; with unaccented base stems (in the second column of (35)). Their accentuation follows at once if we assume that their respective derivational suffixes are preaccenting in I,1, accented in I, 2, and unaccented in I, 3. The derivation of these forms is shown in the bottom part of (35).

<table>
<thead>
<tr>
<th>(35)</th>
<th>accented base</th>
<th>unaccented base</th>
<th>accented base</th>
<th>unaccented base</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>koróv+ij</td>
<td>tabáč+ij</td>
<td>soldá:ð+ik+u</td>
<td>slov+á:ð+ik+u</td>
</tr>
<tr>
<td></td>
<td>(koróv+u &quot;cow&quot;)</td>
<td>(tabak+ú &quot;tobacco&quot;)</td>
<td>(soldá:ð+u &quot;soldier&quot;)</td>
<td>(slov+á:ð+u &quot;diction&quot;)</td>
</tr>
<tr>
<td>2.</td>
<td>čeloveč+íšč+u</td>
<td>tabáč+íšč+u</td>
<td>Pavl+úš+u</td>
<td>Petr+úš+u</td>
</tr>
<tr>
<td></td>
<td>(čelovek+u &quot;human&quot;)</td>
<td>(tabak+ú &quot;tobacco&quot;)</td>
<td>(Pávl+u &quot;Paul&quot;)</td>
<td>(Petr+ú &quot;Peter&quot;)</td>
</tr>
</tbody>
</table>
III.

In class II the suffixes have the corresponding properties except that they are in addition deaccenting. Hence the derivations from accented bases proceed in exactly the same fashion as those from unaccented bases except, of course, that the deaccentuation rule applies vacuously to unaccented stems.

Our new treatment improves in two important respects on that of Halle (1973). First, the morphologically arbitrary constituent structure differences postulated there are now eliminated in favor of the phonological feature [+ accented] on suffixes. Secondly, a simple system underlying the derivational accent classes is now revealed. Setting aside the feature [+ Metatony], the classification of derivational suffixes in the two analyses compares as follows:

<table>
<thead>
<tr>
<th>Halle 1973</th>
<th>Present analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inherently accented</td>
<td>Constituent structure forming</td>
</tr>
<tr>
<td>1. -ast, -ay, -lt, -at</td>
<td>+</td>
</tr>
<tr>
<td>2. -ist</td>
<td>-</td>
</tr>
<tr>
<td>3. -ač, -ak</td>
<td>-</td>
</tr>
<tr>
<td>4. -ov, -ik</td>
<td>-</td>
</tr>
</tbody>
</table>
Concluding Remarks

We have already remarked on the essential stability of the main features of the Indo-European accentual system in all the languages surveyed. We have found a common core consisting of the three rules

\[(86)\]

a) Deaccentuation rule
b) Metatony rule
c) Basic Accentuation Principle

Moreover, the languages all had both accented and accentless stems as well as suffixes. What differentiated the languages was the role that Metatony played in the declensional paradigm. Whereas in Sanskrit Metatony applied before accentless case endings, in Lithuanian and Slavic Metatony ceased to apply in the declension altogether.

This development would be motivated if it were the case that the theme vowels at some point in the history of Baltic and Slavic ceased to be independent morphemes, and became reanalyzed as part of the case ending. That assumption seems in any case on purely morphological grounds inescapable for the modern Slavic languages, and at least plausible for the modern Baltic languages.

This line of reasoning has a corollary of great interest. We see that the two principal innovations of Baltic and Slavic, the development of marginal mobility and its extension to thematic declensions, involve no addition to the accent system; they are as it were contained in embryo in the inherited Indo-European system. In particular, the second innovation, i.e. the extension of marginal mobility to the thematic declension, would seem to follow from the morphological coalescence of theme vowels and case suffixes. This means that the dating of these innovations to a common Balto-Slavic period is no longer as inevitable as it has appeared. Moreover, since Germanic quite clearly underwent the same morphological reanalysis, it is no longer such a mystery that Verner's Law doublets in Germanic attest to precisely the same initial/final accent mobility in thematic declensions at some prehistoric state of that branch as well. The attempts to project this type of mobility in thematic nouns back into the protolanguage are therefore quite ill-founded.

In any case, we can assume that Baltic and Slavic originally had just two types of paradigm, corresponding to the Indo-European fixed and movable type. In addition to losing Metatony as a rule affecting the accentuation of the declensional paradigm, there have been wholesale changes in the inherent accentuation of whole classes of stems. This topic has been carefully studied by V. I. Ilič-Svityč (1963), who has found that Baltic nouns with short stem vowels basically preserve their Indo-European accentual class—i.e., those that belong to the Lithuanian class II can be shown to be cognate with inherently accented stems in other Indo-European languages, whereas
those that belong to the Lithuanian class IV have cognates with inherently unaccented stems in the related languages. The situation is rather different with regard to Baltic nouns with long stem vowels. Here there appears to have been a wholesale shift from the unaccented to the accented class. Illič-Svityč suggests that flirt was basically correct when he speculated that the shift affected words with a nonapophonic long vowel; i.e., with a vowel "whose length arose as a result of the dropping of laryngeals after syllables... In those instances where the root contained an original short vowel... or a long apophonic element... i.e., where no new length developed, a lack of correspondence in accentuation did not exist." (p. 81). In addition, Illič-Svityč notes that there were also shifts in the opposite direction: some inherently accented stems became unaccented; these shifts, however, were not systematic.

The situation in Slavic is different. Illič-Svityč established that all original short-vowel stems acquired columnar stress on the ending, i.e. shifted to the type kočerɡu - kočerɡam of (26). In terms of our analysis, all stems with short vowels lost their inherent accent and became subject to a rule which places an accent on all case endings, formulated as the Post-stem Accentuation rule in (25). At a more recent period, Russian reintroduced Metatony in the declension of certain classes of nouns, but this time morphologically conditioned, by any accented plural suffix. This results in the types of accentual pattern discussed above (see (25) - (31)).

If we now compare the Lithuanian and modern Russian accent systems, we see essential similarity in the derivational systems but quite fundamental differences in inflection. Historically, the four accent classes of Lithuanian are completely stable: accent changes in this language have been, as already noted, a matter of shifts in the underlying accentuation of certain stems and endings, leaving the four classes of stems and the four classes of endings unchanged. Russian, on the other hand, has experienced a proliferation of minor subtypes among inherently unaccented, movable stems.

How is the different trend of Russian and Lithuanian to be explained? The reason would appear to be the fact that the Lithuanian system remained entirely phonological -- the accentuation of any inflectional form is determined in mechanical fashion by rules (3, 14, 16) from its underlying representation; whereas the Russian system has been morphologized, in particular being tied to number in the noun declension. The entire range of subtypes which have been spawned by analogy can in fact be understood as the result of the tendency to recruit the accent alternations to signal the singular/plural distinction.

The common core of the Indo-European accentual system consisting of the three rules, a) Deaccentuation, b) Metatony, c) BAP and associated lexical markings is fairly powerful, but it is important to note that not just any imaginable accent paradigm could be accommodated in it. Of particular interest is the fact that
there are imaginable kinds of accent paradigms which would be extraordinarily complex and unnatural in our framework no matter how many new rules we added, while they might be quite simple to describe in some other approaches to morphological stress systems. An example of such an alternative approach, no less plausible on a priori grounds than ours, perhaps, is that proposed by Garde (1968). Garde's idea is that morphemes are ordered into levels of accent "strength", such that the actual stress of a word is determined by its strongest morpheme. For example (denoting increasing strength by increasing subscripted integers) let $A_3$ by the strength of inherently accented stems, such as Skt. bhrātar-, āśva. Garde's theory would allow, and indeed predict, the existence of derivational suffixes $B_5$ and $C_4$, such that derivatives formed by them are accented as in (37):

(37) 
(a) $A_3 + C_4$
(b) $A_3 + B_5 + C_4$

On our theory this would be an anomaly, since (a) indicates that C triggers deaccentuation, while (b) indicates that it does not. If such cases are in fact nonexistent or rare, as they appear to be, that would be evidence favoring our treatment over Garde's.

By way of further support we may add that accentual systems with exactly the properties we postulate for IE are by no means unprecedented; one such system has been described by Jane and Kenneth Hill in IJAL 1968. We quote from their study:

"A root may have inherent stress. Stress in such roots is invariant [fixed stress] except as discussed [below]. Stress may be determined by its presence in certain affixes, if the root(s) to which they are attached have no inherent stress... [movable stress]. If neither the root nor the affixes in a word have inherent stress, stress is placed on the first vowel of the word... [BAP]. Root stress overrides affix stress, and only the first of a series of root stresses within a word is retained phonetically... [cf. the compounds]. Some suffixes place the stress on the last vowel of the root." [Metatony].

The Hills, of course, were describing not Indo-European, but the Uto-Aztecan languages, Cupeño, Cahulla, and Luiseno. Their stress system differs from that of Indo-European mainly in having no deaccentuation rule. However, the remarkable similarities between the Hills' system and the one presented here clearly show that we are not dealing with an isolated and unprecedented system. In fact, in our reading we seem lately to come across one instance after another of this system in widely separated language groups. To mention but one, we found a system quite
similar to that of Indo-European, including even the deaccentuation rule, in Asurini as described by Carl Harrison in the recent SIL volume on the Tupi languages of South America. While such parallels, of course, do not establish the correctness of the proposed analysis, they do provide what Calvinists would call "comfortable assurance" that the enterprise may not be totally misconceived, and at this point in our work that is about all that we would claim for what we have presented here.

NOTES

1. We do not mean that this is all there is to say about the accentuation of compounds. There are, of course, many subsidiary regularities requiring their own rules, as well as the sprinkling of exceptions to be expected in any such lexical domain. One quite general rule is that a compound in which both members are inherently unaccented gets accented on its last syllable. Hence, there are no accentually mobile compounds. In classical Sanskrit, as described by Pāṇini, compounds are accented on quite different principles.

2. We take the Sanskrit rather than the Balto-Slavic pattern to be the original one for the reasons given in Kiparsky (1973).

3. There is a problem with the dative dual and dative plural forms:

   várnoms   galvóms   ránkoms   barzdóms

Given the facts of galvóms we should represent it as

   h  h
    galv + ooms

And the correct output would be obtained. However, this at once leads to problems with ránkoms. If underlyingly it is represented as

   h  h
    rank + ooms

Saussure's law should then apply, giving the incorrect

   h  h
    rank + ooms.

To get the correct output here we would need

   h  h
    rank + ooms.

However then we would get (as in the gen. sg.) a rising tone in the case suffix, i.e.

   * galv + oms.
The solution is reasonably simple. In designating the tones in (11) above we have considered only cases of two mora sequences. And we have therefore said that if the accent goes on the first mora in the sequence we get falling tone; if it goes on the last mora we get rising tone. As we have seen, this leads to incorrect result when we extend our formulation to three mora sequences. To correct this we need, however, only say that we get falling tone, when accent goes on the nonfinal mora; and as before we get rising tone when accent falls on the final mora.

Additional support for this analysis comes from certain dialectal facts. In the Zemaitė dialect (as in Latvian) there is syncope in the nom. sg. masc. suffix: as → s. This syncope is accompanied by a change in tone:

- **St. Lith.** džaras underlyingly džas
- **Zemaitė** dvaras underlyingly *h dvaras + s

4. This accentual differentiation of singular vs. plural forms has been previously noted; see, especially, R. Jakobson (1957) and E. Stankiewicz (1968).

5. N. Durnovo (1932) reported that in a XIV century Pskov manuscript the nom. pl. vody "waters" is transcribed with a "kamora" over the o. Since this is the traditional way of representing the "neo-acute" vowel, the transcription indicates that this noun was subject to Metatony already in the XIV century, much earlier than the overwhelming number of examples adduced in Kiparsky (1962).

6. Recall that because of the fact that neuter case endings are normally "strong" (unaccented) in the singular and "weak" (accented) in the plural, the "mobile" paradigm shows stem accent in the singular and suffixal accent in the plural. The example ūxo cited in (23) is exceptional in that it takes the "strong" i-suffix in the nom. pl.

**REFERENCES**


Kurjtowicz, J. (1968) Indogermanische Grammatik II (Heidelberg, Carl Winter).


The Yuman language family consists of ten closely related languages still spoken in the Southwest. This paper will assume the following classification which is reasonably well agreed upon among Yumanists (Kroeber, 1943; Joel, 1964; Wares, 1968):

- Pai group: Walapai, Yavapai, Havasupai (all in Arizona); Paipai (Baja California)
- River group: Mojave, Yuma (both on the Colorado River); Maricopa (Arizona)
- California-Delta group: Cocopa (Colorado River–Sonora); Diegueño (Southern and Baja California)
- Kiliwa (Baja California)

All descriptions of Yuman languages agree that these are stress languages, and that stress is assigned morphologically to the root of the word. The canonical shape of a Yuman root may be reconstructed as a monosyllabic morpheme bearing stress (Langdon 1975) consisting of at least a vowel (short or long), most commonly preceded, followed, or surrounded by single consonants. Examples from Diegueño, which may be considered typical in this respect, illustrate the various possible types.


Roots may function as words but more commonly words are formed by, among other processes, surrounding the root by prefixes and suffixes which produce unstressed syllables. Uninflected words tend to have stress on the last syllable in direct consequence of the great number of derivational prefixes and the very small number of derivational suffixes. Even when derivational suffixes are present, since they typically consist of a single consonant, no post-stress syllable results. Thus

(2) Diegueño: ts –xə–mə –k’ wán–p ‘is tangled up’

has three pre-stress syllables and no post-stress syllable. Nevertheless, suffixes producing unstressed syllables are common, though they are mostly syntactic in function. In fact, complex suffixal sequences are possible, particularly in verbal constructions with criticized stressed higher predicates, as in

Note that regardless of the complexity of the word, the stress remains fixed on the root. The only thing that can happen to a stressed syllable is that the stress reduces in compound formation to secondary stress on the first member of the compound, thus maintaining the preference for primary stress on the last syllable. Secondary stress in compounds may also further reduce so as to become indistinguishable from other unstressed material.

(1)  Diegueño:  xáta:pá: 'coyote' (xáta 'dog' + pá: 'person')
matátay'mountain' (mát 'ground' + táy 'big')

The major stress-related problem in the analysis of Yuman languages has to do with the unstressed parts of utterances, which exhibit much reduction and insertion of vowels in ways which do not lend themselves to elegant formulation by means of phonological rules. This has been described in considerable detail by Yumanists (e.g. Crawford 1966, Halpern 1946-47, Joel 1966, Langdon 1976, Mixco 1971, Redden 1966, Shaterian ms, Winter 1966) and will not be further discussed here. This paper will instead focus on the behavior of stressed syllables in several Yuman languages and explore in some detail the peculiar fact that, unique among Yuman languages, Mojave appears to exhibit stress shifts.

What is labelled stress in Yuman is a complex combination of features of loudness, pitch, and length resulting in greater perceived prominence of the so-called stressed syllable. On the other hand, vowel length is underlyingly contrastive for all Yuman languages and is clearly to be reconstructed for Proto-Yuman, though some language-specific complications arise which will be discussed below. Also to be reconstructed for Proto-Yuman are three contrastive vowel qualities which, coupled with length, give a total of six contrastive vowels, both in stressed and unstressed position. The California-Delta subgroup and Kiliwa have preserved a 3-vowel system but the other languages have developed 5-vowel systems through the operation of a set of rules I have described elsewhere (Langdon 1976), and exhibit to various degrees a tendency to minimize the contrastive importance of vowel length. Nowhere has length become fully predictable, however, because it serves not only lexical differentiation purposes but also morphological functions in stem alternations denoting plurality and derivation. The Pai subgroup and Kiliwa no longer have vocalic length contrasts in unstressed position, while these remain in the River and California-Delta languages.

Of the 3-vowel languages, Cocopa has preserved the Proto-Yuman system best, both for quality and quantity. This is therefore a good starting point for a look at pitch and intonation contours. Crawford (1966) notes that Cocopa is stress-timed, and it is likely that this observation is valid for the other Yuman languages as well. He also describes distinctive intonation contours that characterize various sentence types. The stress-timed feature is responsible for the great variation in unstressed parts of utterances while the intonation contours may select for greatest prominence.
one of the stressed syllables in the sentence by giving it higher pitch, although a very common pattern keeps the pitch of successive stressed syllables even. When the higher pitch option is taken, the selection of the syllable which receives higher pitch is based on syntactic considerations, the main pattern being for the main verb to receive the highest pitch. Final contours then interact with the position of highest pitch to produce the characteristic sentence melody, the most common one being a falling contour for declarative sentences or those questions formed by suffixation. An interesting variant occurs only with negative sentences which take a slightly rising contour on the negative morpheme la:x, which is sentence final.

(5) Cocopa:

a. declarative

\[ \text{n'a:p n'wá-y p'á:-c šayu} \] (me my=house-in it=stands=same=subject is=there) 'It stands there in my house.'

b. negative

\[ \text{apá-c u'á:m} \] (man-subject he=did) 'The man did it.'

c. interrogative

\[ \text{mapú-c kafi: m-a? is m-ší-a} \] (you-subject coffee you-have you-drink-question) 'Did you drink your coffee?'

d. alternative declarative

\[ \text{n'a:-c wá-y man spac} \] (I-sub house-in leave emerge) 'I went out of the house.'

Diegueño is very similar to Cocopa, except that it has neutralized the short/long vowel contrast in roots ending in a vowel, as the compared items in (6) indicate, so that Cocopa is crucial for determining the underlying length of such forms.

(6) Diegueño

\[ \text{'I drink'} \]

\[ \text{'your nose'} \]

\[ \text{'bird'} \]

Comparing with Cocopa:

\[ \text{'I drink'} \]

\[ \text{mi:xu} \]

\[ \text{'bird'} \]

\[ \text{'your nose'} \]

Timing and melody are comparable to those of Cocopa. In addition, the pitch configurations of certain canonical word forms described for the Mesa Grande dialect in Langdon (1976) can be summarized as follows. The stressed syllable of a word is typically higher in pitch than the rest, although it has a less common variant with lower pitch than the rest. Unstressed syllables tend to have low pitch, except the directly pre-stress syllable if it contains a long vowel or better yet a short vowel followed by glottal stop. The examples in (7) show some typical word pitch contours, the final drop in pitch representing the typical utterance-final contour. Note
particularly the pattern with glottal stop in the unstressed syllable followed by a stressed syllable with short vowel (i.e. the word for ‘the pretty one’) where the first vowel has a higher pitch than the second.

(7) Diegueño:  
- xá má: 'he sleeps'  
- k’a má: 'the high one'
- səxə pən: 'he kneels'  
- k’a səsənu 'the pretty one'
- a:k wəlvək 'if he licks'

For Kiliwa, Mixco (1971) claims not only the 3 vowel qualities, each occurring both short and long, but also contrastive pitches, three for long vowels (high, low, and high-falling), two for short vowels (high and low). He supports this analysis by the classical method of presenting minimal pairs, a sampling of which is given in (8a) for short vowels and in (8b) for long vowels.

(8) Kiliwa:  

(a) short vowels
- p hə 'belly'
- yəl 'bark (of tree)'
- səl 'hand'
- p hə 'nose'
- xəp 'metate'
- p hə 'wash face'
- yəl 'to crush'
- səl 'to fry'
- p hə 'stinkbug'
- xəp 'pond'
- səl 'to roast'
- səl 'to penetrate'

(b) long vowels
- kumf 'man'
- kəm 'shaman'
- ḡa: 'to grind'
- ḡa: 'to go'
- mə 'to fear'

He notes further that lexical differentiation is the only function of the pitch contrasts, and also that vowel length contrasts are often neutralized in connected discourse, in which case a high-falling pitch often associated with long vowels is indistinguishable from high pitch. Since in addition there is only one known set of lexical items demonstrating the contrastiveness of three pitches (i.e. 'man', 'shaman', 'fear' above) it
is probable that no more than two contrastive pitches are truly functional in Kiliwa. An examination of the minimal pairs for short vowels also reveals a tendency for the high pitch to be more often associated with nouns and the low pitch with verbs (e.g. 'bark (of tree)' vs 'to crush', 'belly' vs 'to wash face', 'hand' vs 'to fry'). An explanation for this might be sought in the verb final nature of Yuman languages and the common tendency for pitch to drop toward the end of an utterance. Since not all Yuman languages have the distinction made in Kiliwa, this suggestion must at this point be considered as purely speculative. A phonological observation is that all items in minimal pairs with short vowels end either in glottal stop (which is apparently predictable in this environment, so that these stems are essentially vowel-final), or in 1, suggesting the possibility of an earlier neutralization of length in these environments. This is reminiscent of a process attested in Diegueño (see 6 above). As to the three-way pitch contrast for long vowels, comparative evidence can be brought to bear on the one set of contrastive forms to show that two very different historical sources underlie the Kiliwa forms as shown in (10) where tentative Proto-Yuman reconstructions are indicated.

(10) Kiliwa

\[ \text{(ko)m}\ddot{\text{f}}:\text{y} \rightarrow \text{'man'} \]
\[ (k)m\ddot{\text{f}}:\text{y} \rightarrow \text{'shaman'} \]
\[ \text{m}\ddot{\text{f}}:\text{y} \rightarrow \text{'to fear'} \]

Proto-Yuman

\[ *\text{xm}\ddot{\text{f}}: \rightarrow \text{'male'} \]
\[ *\text{m}s\ddot{\text{y}}:\text{a}\ddot{\text{y}} \rightarrow \text{'to fear'} \]

A certain amount of fluctuation in pitch is also noted in different forms of what is clearly the same root.

(11) Kiliwa

\[
\begin{align*}
\text{k}^\text{w}\ddot{\text{n}}\ddot{\text{f}}\ddot{\text{y}} & \rightarrow \text{'the black one'} \\
\text{x}\ddot{\text{a}}\ddot{\text{n}}^\text{w} & \rightarrow \text{'black water'} \\
?\ddot{\text{n}}\ddot{\text{f}}\ddot{\text{y}} & \rightarrow \text{'I have'} \\
\ddot{\text{n}}\ddot{\text{f}}\ddot{\text{y}} & \rightarrow \text{'he has'} \\
\text{c}\ddot{\text{n}}^\text{w}\ddot{\text{f}}\ddot{\text{i}}\dot{\text{l}}\ddot{\text{u}}p & \rightarrow \text{'rolls it'} \\
\text{p}\ddot{\text{c}}\ddot{\text{n}}^\text{w}\ddot{\text{i}}\dot{\text{l}}\ddot{\text{u}}p & \rightarrow \text{'rolls about'}
\end{align*}
\]

Other possible reasons for the development of contrastive pitch may be inferred from some synchronic phonological rules of Kiliwa, particularly the following:

1. an unstressed vowel followed by glottal stop has high pitch (note the similarity to the Diegueño facts in 7), and
2. otherwise, unstressed vowels have the inverse pitch of the stressed syllable in the same word (for the purposes of this rule, high and high-falling pitch behave the same). Examples are

(12)
1. paʔ miː 'you said'
2. takaːn 'mouse' salkaː 'gopher' nǐwːy 'a ceremony'

A historical interpretation of these facts might be that at an earlier stage the pitch level of an unstressed syllable may have been determined by conditions independent of the pitch of the stressed syllable and may in turn have had some effect on the pitch of the stressed syllable itself. Other factors have undoubtedly also contributed to the emergence of pitch contrasts. Among these, we may note that Kiliwa has undergone more drastic sound changes than other Yuman languages. Note for example a still operating synchronic rule of Kiliwa changing stressed u to i in some rather widespread environments, such as closed syllables.

We now turn to Walapai, where we have two descriptions (Redden 1966, Winter 1966) which choose diametrically opposed solutions to what is clearly a serious problem in the analysis of vowel length. Redden (1966) who was also first to note the stress-timed character of Yuman languages in his work on Walapai, considers length non-distinctive and predictable from the operation of the stress-timing dynamics, though he fails to formulate the appropriate rules. For a suggestion on what is involved, I reproduce here two renditions of the same Walapai sentence from Redden (1966, p. 13) using different options of the stress-timing pattern.

(13) Walapai yektm nʔa man-m ha cipak vatvokun' (this=morning when arise-different=subject water emerge)9 'When I got up this morning, we had water again.'

\[ yéektm nʔa má:mm aháa + cipáak vátvökun' hh \]
\[
\begin{array}{ccccc}
1 & 2 & 3 & 4 & 5 \\
\end{array}
\]

vs. \[ yéektm nʔa má:mm hà cipáak vátvökun' hh \]
\[
\begin{array}{ccccc}
1a & 2a & 3a & 4a \\
\end{array}
\]

On primary stress, Redden (1966, p. 15) notes: "A primary-stressed vowel is often higher in pitch than immediately adjacent vowels, but this is by no means always the case. What is heard as stress is a combination of loudness, pitch and length. One, or even two, of these features may not be present, and stress will still be heard." Also, final statement contour "...is marked by sharp decrease in stress, pitch, and voicing." Note that the decrease in voicing is marked by hh in the figure above.
In the other description of Walapai, Winter (1966) identifies a phoneme called "vowel raiser" which clearly corresponds to vowel length in other descriptions of Yuman languages. Although it enters into a number of phonological rules and differentiates lexical forms, its functional load in the language is somewhat limited.

For Paipai, Joel (1966), struggling with what is clearly a version of the same situation, states that there are three types of stressed vowels, short, long, and intermediate or indifferent. In addition, I have had access to a long Paipai wordlist (Kaufmann and Shaterian, 1974) transcribed in fine and careful phonetic notation. This list agrees with Joel in noting three degrees of length, but also indicates three pitch levels, high, falling, and rising. When these data are examined, a strong tendency emerges for each vowel length to be more commonly associated with a particular pitch. Thus, short vowels tend to have high pitch, mid vowels falling, and long vowels rising. Once this correlation has been established it becomes instructive to look at successive recordings of the same items in various environments. Let us first illustrate the clear contrasts, by a few examples of minimal pairs of short vowel with high pitch vs long vowel with rising:

(14) Paipai

<table>
<thead>
<tr>
<th>Short Vowel</th>
<th>Long Vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>pu:k</td>
<td>xəmά:</td>
</tr>
<tr>
<td>púk</td>
<td>xəmά</td>
</tr>
<tr>
<td>xəmά:n</td>
<td>muvʃį</td>
</tr>
<tr>
<td>xəmά:n:n</td>
<td>muvʃį</td>
</tr>
</tbody>
</table>

To my knowledge no minimal sets involving the three degrees of length are available.

In (15) below are exhibited variant recordings of what is believed to be the same underlying root. Where possible, a true contrastive item is also shown.

(15) Paipai

a. Alternations of stressed short vowels: \( \bar{V} \sim \bar{V} \sim \bar{V} \)

<table>
<thead>
<tr>
<th>Short Vowel</th>
<th>Resultant Vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>xʷát</td>
<td>'red'</td>
</tr>
<tr>
<td>xʷát:</td>
<td>'red'</td>
</tr>
<tr>
<td>cxʷát:</td>
<td>'blood'</td>
</tr>
</tbody>
</table>

b. Alternations of Intermediate length stressed vowels: \( \bar{V} \sim \bar{V} \sim \bar{V} \)

<table>
<thead>
<tr>
<th>Intermediate Vowel</th>
<th>Resultant Vowel</th>
</tr>
</thead>
<tbody>
<tr>
<td>ʃpó:k</td>
<td>'knows it'</td>
</tr>
<tr>
<td>ʃpó:tem</td>
<td>'doesn't know'</td>
</tr>
<tr>
<td>ʃtupé-k</td>
<td></td>
</tr>
</tbody>
</table>
c. Alternations of stressed long vowels: \( \tilde{v} \): 'sells it' \( \tilde{v} \): 'buzzard' \( \tilde{v} \): 'eats it' 
\( \tilde{V} \): 'is selling it' \( \tilde{V} \): 'buzzard' \( \tilde{V} \): 'I am eating' 
\( \tilde{W} \): 'is eating' \( \tilde{W} \): 'in order to eat' \( \tilde{W} \): 'food' (short vowel)

CONTRASTING WITH

It should be noted that the alternations of sets a) and b) are essentially the same, suggesting that there is no basic contrast between short and intermediate length vowels, nor between high and falling pitch, and that the true contrasts are between sets a) and b) on the one hand, i.e. short or intermediate length vowels with high or high-falling pitch, and set c) on the other, i.e. long vowels with rising pitch. It thus seems that the best clue to contrast is pitch rather than length, so that rising pitch is always a clue to the presence of vowel length. The only ambiguity is possible when a vowel is recorded with intermediate length and high pitch which the examples in (13) demonstrate can be a variant of both types. Thus, given only m\( \tilde{s} \)tupé 'you believe' and cm\( \tilde{a} \)cu 'in order to eat' without access to other forms of the same paradigm, it would be impossible to decide whether the stressed vowel is long or short. I would like to suggest that the difficulty encountered in analyzing the vowels of Paipai results from this very ambiguity.

For Yavapai, Shaterian (ms) finds it necessary to set up in stressed syllables not only the 5 vowels common to the Pai group, but also 3 degrees of length and 2 contrastive pitches. Minimum sets for length are shown in 14a and for pitch in 14b.

(14) Yavapai

a. length contrasts

\( \tilde{v} \) 'worm' \( \tilde{v} \) 'water' \( \tilde{v} \) 'name'
\( \tilde{v} \) 'be steep' \( \tilde{v} \) 'be bitter' \( \tilde{v} \) 'sugar'
\( \tilde{v} \) 'thread' \( \tilde{v} \) 'cottonwood' \( \tilde{v} \) 'be sweet'

b. pitch contrasts

\( \tilde{v} \) 'I am sitting' \( \tilde{v} \) 'be ungrateful' \( \tilde{v} \) 'horn'
\( \tilde{v} \) 'dwelling, house' \( \tilde{v} \) 'to scoop out' \( \tilde{v} \) 'tree'

Shaterian acknowledges serious difficulties with this scheme since this formidable apparatus does essentially very little work in the language.

"The distinctive pitch accents have a very low functional load.

Pitch, stress, and length are interrelated phenomena. Two pitches
and three vocalic lengths can cooccur only with primary stress; with secondary stress the two contrasting pitches are neutralized and only two vocalic lengths are present. Furthermore, there is considerable free variation with respect to pitch and length occurring with primary stress; a specific lexical item may carry the same pitch in elicitation form consistently only to appear with the opposite pitch after affixation occurs." (p. 27)

To illustrate the extent of variation possible for a single item, I reproduce below (item 15) all the variants noted by Shaterian (ms) for the fifteen words of 14.

(15) Yavapai

'worm'  ?i?\breve{a}la ~ ?i?\breve{a}la

'be steep'  ?i?\breve{f}:li, mat?i?\breve{a}la 'cliff (steep place)'  ?wi\textsuperscript{1}k?\breve{f}:la 'mountain (steep rock)'

'thread'  ?i?\breve{f}:la (said to be from Spanish hilo)

'water'  ?ah\breve{a} ~ ?ah\breve{a}

'be bitter'  ?ah\breve{a} ~ ?ah\breve{a}

'cottonwood'  ?ah\breve{a}\breve{a} ~ ?ah\breve{a}: ~ ?ah\breve{a}.

'be sweet'  múl ~ múl ~ m\breve{o}\breve{l}, pa:m\breve{u}-lva 'prominent person'

'be sweet'  miy\breve{u}-li ~ miy\breve{u}:li

'sugar'  miy\breve{u}-la ~ miy\breve{u}la

'I am sitting'  ?w\breve{a}

'dwelling, house'  ?w\breve{a}, ?w\breve{a}ca 'house-subject', n\breve{y}w\breve{a}\breve{a} 'home'

'be ungrateful'  h\breve{n}\breve{u}, h\breve{n}\breve{u}-km 'is ungrateful', ?h\breve{n}ukm 'I am ungrateful'

'to scoop out'  h\breve{n}\breve{u} ~ h\breve{n}

'horn'  k\breve{w}\breve{a} ~ k\breve{w}\breve{a}, k\breve{w}-vc l\breve{a}\breve{w}li 'horns', but k\breve{w}\breve{a}:vi 'horned' (long vowel)

'tree'  k\breve{w}\breve{a}

Among the minimal sets for lengths, it is noteworthy that each consists of either two nouns and one verb or two verbs and one noun, suggesting only a basic two-way contrast for any category. It should also be noted that the pairs exemplifying the pitch contrasts all involve short root final vowels, suggesting once again (as in Diegueño and Kiliwa) an earlier length neutralization in that position, with the pitch preserving the contrast. While much remains to be explained before we can fully understand how the Yavapai situation developed, the above discussion makes clear that, in addition to phonological considerations, the whole question of length and pitch must take into account such syntactic facts as part of speech and therefore the place a word can occupy in a sentence and the types of affixes it can take.
We finally come to Mojave, where a very different situation is reported. Even early phonetic observations about the language (Kroeber, 1911) have made reference to the fact that, although the typical stress pattern for words in Mojave is to have stress on the root vowel as in other Yuman languages, a number of words are perceived with stress elsewhere and that some minimal pairs of stress placement result.

\[(16) \text{Mojave}^{13} \text{\begin{tabular}{l}
\'aha 'water' \\
\'ah:ve 'rattlesnake' \\
\'i:do 'my eye' \\
\'ah?á: 'cottonwood' \\
a:vé 'mouse' \\
i:do 'willow'
\end{tabular}}\]

The data to be examined for Mojave are from two tapes recorded on a Nagra I Vd tape recorder by Pamela Munro and Linda Munsen containing word lists constructed with particular relevance to the stress problem. Each word on the list was elicited from two native speakers, Thomas Stevens (TS), and Nellie Brown (NB). Two separate phonetic transcriptions were made of these tapes, one by myself, and one by Donald Crook. In the areas of the place of stress and the general pitch contours, our subjective recordings coincide remarkably well, and we disagree only slightly in our perception of vowel length. We both agree with other investigators that some Mojave words are always perceived with stress on a syllable other than the root, that some words have stress move from one syllable to another upon repetition, and that some others behave as good Yuman words should, with stressed syllables on the root. In addition, we both heard vowel length in stressed and unstressed position, as have all other investigators.

Some distinct patterns emerge from an examination of these transcriptions. These are exhibited in (17) below and demonstrate that the place of perceived stress can be predicted from the various combinations of short and long vowels in the pre-root and root syllables (other syllables are never perceived as stressed). In the formulas summarizing the various patterns, X stands for any consonant or even, in medial and final position, consonant cluster. In addition, in final position, X can also be zero.

\[(17) \text{Mojave}^{14} \text{\begin{tabular}{l}
a. \text{\begin{tabular}{l}
\( X \ V \ X \ V \ X \end{tabular}}'aha 'water', cúksa (TS)~çúška (NB) 'head', ká?ak 'he kicks it', númme 'mountain lion', \( \check{\text{á}} \text{se} ' (TS)~\check{\text{á}} \text{se} (NB) 'buzzard', \( \check{\text{á}} \text{h} \text{w} 'enemy', hán' ò 'lake', \( \check{\text{á}} \text{m} \text{m} ' (i:)yavó'mme 'beard'
\end{tabular}}
\text{\begin{tabular}{l}
b. \text{\begin{tabular}{l}
\( X \ V \ X \ X \ X \end{tabular}}'á:hma 'quail', hì:ya 'mouth', \( \text{ú:ce} 'coals', \( \text{i:} \text{e} 'my hair', \( \text{á:k} 'cloud', \( \text{á} : \text{ve} 'rattlesnake', \( \text{é} : \text{t} \text{o} 'one', \( \text{h} \text{i} : \text{t} \text{o} 'stomach', \( \text{h} \text{i} : \text{d} \text{o} 'eye', \( \text{a} \text{v} \text{é:n} 'y 'road'
\end{tabular}}
\text{\begin{tabular}{l}
c. \text{\begin{tabular}{l}
\( X \ V \ X \ X \ X \end{tabular}}'aspa: (TS)~'aspa: (NB) 'eagle', vahá: 'guts', \( \text{án} \text{yà:} 'sun', \( \text{áv} \text{í:} 'rock', \( \text{á} \text{yà:} 'mesquite', \( \text{h} \text{i} \text{d} \text{ú:} 'is', \( \text{á} \text{c} \text{i:} 'fish'
\end{tabular}}
\end{tabular}}\]
c2. X V X V: C(voiced): hupá:m 'lays down', himí:m 'he cries'
c3. X V X V: C(voiceless) ~ X V X V: C(voiceless): hupá:ka ~ húpa:ka 'snow', hívë:sk ~ hívë:sk 'he runs', nákú:t(k) ~ náku:tk 'man's father', ?átá:yk ~ ?átá:yk 'a lot', ámma:t 'land' (no alternant was recorded for this word which perhaps does not belong to this group and will be discussed below)
d. X V: X V: X


The generalizations suggested by these facts may be stated as follows. When both the root and pre-root syllable have short vowels, or when the pre-root vowel is long and the root vowel short, stress is perceived on the pre-root vowel (types a and b). When the pre-root vowel is short and the root vowel long, stress is perceived on the root (types cl, 2, and 3), with a possible alternant with stress on the pre-root vowel if the root vowel is followed by a voiceless consonant (type c3). When both vowels are long, stress is perceived on either one, or both are perceived as equally stressed (type d).

These statements account for the vast majority of the words transcribed from the tape and make clear that, when the root vowel has been identified and the vowel length is as indicated in the examples, the place of perceived stress is fully predictable. We can therefore assume that Mojave, like other Yuman languages, has underlying stress on the root and derives the surface stress patterns by the application of stress retraction rules embodying the generalizations just stated.

The facts, however, are slightly more complicated. In assigning vowel length to the forms displayed in (17), a certain amount of normalization had to be resorted to. Since for each form there were 8 transcriptions (both speakers said the same word twice and each token was transcribed by two people), if all agreed with respect to vowel length, there was of course no problem. If the distribution was skewed, the most common form was assumed to represent the norm, but if the distribution was even, outside evidence was resorted to to resolve the difficulty, such as data from other sources than the tape (e.g. Munro 1974 and private communication) as well as comparative evidence from other Yuman languages which are known to be conservative with respect to vowel length. Upon carefully examining the transcriptions from the tape, it was noted that an important clue to the perception of stress was the perception of pitch, so that what we recorded as the highest pitch level coincided with the place of perceived stress. A similar observation is found in Kroeber (1911):

Of more importance than stress in the accentuation is pitch. While the writer's ear is poorly trained for the observation of this quality, he
finally received the impression that that syllable of a word which appeared to be most strongly stressed was spoken in a higher pitch than the others. Measurements of wave lengths made in a number of tracings that are of large enough size to render this procedure feasible corroborated this subjective judgment. (p. 64) [the reference here is to kymograph tracings Kroeber obtained in the field]... it seems probable that increased stress and higher tone coincide in words accented on the ultimate [i.e. the root], but that when the accent falls elsewhere it consists only of a rise in pitch. (p. 65)

These are extremely insightful observations, and our subjective observations of the modern Mojave data in general confirm Kroeber's assessment. A summary of the pitch patterns exhibited by our transcriptions is given below for each of the stress types of (17):

\[\text{(18) Mojave} \]

\begin{itemize}
  \item a. \(X \vee X \vee X\)
  \item b. \(X \vee: X \vee X\)
  \item c1. \(X \vee X \vee:\)
  \item c2. \(X \vee X \vee: C(\text{voiced})\)
  \item c3. \(X \vee X \vee: C(\text{vl}) \sim X \vee X: C(\text{vl})\)
  \item d1. \(X \vee: X \vee:\)
  \item d2. \(X \vee: X \vee: C\)
\end{itemize}

These patterns demonstrate that the place of perceived stress coincides with the place of perceived highest pitch. In particular, they make clear the difficulty in establishing the stressed syllable for the d) patterns which are perceived with even pitch on both syllables. In general, it seems that high pitch is most typically associated with long vowels, except for the optional variant of pattern c3) which might be accounted for by the fact that the drop in pitch associated with the final voiceless consonant is allowed to begin earlier in the word. The only genuinely surprising pattern is a), where no motivated reason for stress shift is apparent, both vowels being short. It might be conjectured, however, that this pattern, which an examination of the forms in 17a shows to affect a number of words beginning in \(?v\), may have originated in words of this type, given the already noted affinity of syllables containing glottal stop for high pitch. Another tendency illustrated by the examples in 17a is for root-initial consonants to lengthen. Root-initial is precisely the position which I have argued elsewhere (Langdon 1975) to be the phonologically strong position in Yuman. The phonetic result of this feature is that the pre-root syllable thus acquires more
length as the onset of the long consonant is interpretable as part of the preceding 
syllable. This may then be interpreted as an extension of the correlation between 
length and high pitch noted above.

Keeping the above discussion in mind, it is now possible to look systematically 
at fuller data from the tapes, i.e. at alternate transcriptions of the same word. In 
19 are listed those words from 17 for which several variants were recorded.

\[(19)\] Mojave

a. \(XVXV\) ?áha (TS-NB) \(\sim\) ?áha: (TS) 'water'
   ká?:ak (TS) \(\sim\) ká?:ak (TS) 'he kicks it' (note also ka?:km (NB))
   ?ámmo (TS-NB) \(\sim\) ?ámmo: (NB) 'sheep'
   yavómmé (TS) \(\sim\) i:yavómmé (NB) \(\sim\) yavómmé (TS) 'beard'

b. \(X:i\) \(XV\) ?á:hma (NB) \(\sim\) ?á:hma: (NB) \(\sim\) ?á:hma (NB-TS) 'quail'
   ?í:i?:e (TS-NB) \(\sim\) ?í:i?:e (TS) 'my hair'
   ?í:k:i:e (TS-NB) \(\sim\) ?í:k:i:e (NB) 'cloud'

c. \(XVXV\) ?aspá: (TS) \(\sim\) ?aspá: (NB) \(\sim\) ?aspá: (TS) \(\sim\) ?aspá: (NB) 'eagle'
   vahá: (TS-NB) \(\sim\) vahá: (TS) 'guts'
   ?aví: (TS-NB) \(\sim\) ?aví: (TS) 'rock'
   ?ací: (NB) \(\sim\) ?a:ci: (TS) 'fish'

c2. \(XVXV\) C(vd) himí:m (TS-NB) \(\sim\) hi:mí:m (TS) 'he cries'

c3. \(XVXV\) C (vl) nakú:tk (TS) \(\sim\) nakú:tk (NB) 'man's father'

d1. \(XVXV\) ?ā:ve: (NB-TS) \(\sim\) ?ā:ve: (TS) \(\sim\) ?ā:ve: (NB) 'mouse'
   ?í:pa: (NB) \(\sim\) ?í:pa: (NB) \(\sim\) ?í:pa: (TS) 'man'
   ?í:i:do: (TS) \(\sim\) ?í:i:do: (TS) 'willow tree'
   hi:hu: (TS-NB) \(\sim\) hi:hu: (NB) \(\sim\) hi:hu: (TS) 'nose'

d2. \(XVXV\) C hi:mi:θa (NB) \(\sim\) hi:mi:θa (NB) \(\sim\) hi:mi:θa (TS) \(\sim\)
   hi:mi:θ (TS) 'fur'
   ck a:θá:o (NB) \(\sim\) ck a:θá:o (TS) 'lungs'
   dokó:pi:t (NB) \(\sim\) dokó:pi:t (TS) 'owl'

The startling observation to be made here is that vowel length, which had been crucial 
for the determination of the place of perceived stress, exhibits much fluctuation, particularly 
in the root syllable when the vowel is root final (once more, the parallel to the Diegueño 
rule must be pointed out). In fact, this can be accounted for under the analysis proposed 
above as follows. Long underlying vowels are pronounced with high pitch which in turn 
is perceived as stress. Since they are thus kept distinct from short vowels, their 
actual duration is less crucial to their identification. On the other hand, short vowels, 
pronounced with lower pitch and perceived as not bearing stress may be non-distinguishably 
lengthened. While this situation creates a certain amount of surface ambiguity which 
may well lead to a reanalysis of stress placement in Mojave, the facts as presented 
still allow the unambiguous identification of the underlying stressed syllable as above.
This solution also accounts for the additional fact that, in connected discourse, all forms appear with the underlying root stress. An example from the tapes illustrates this:

(20) Mojave

\[ ?\text{a}:\text{yk} \] (TS) 'a lot'
\[ ?\text{a}:\text{yt}:\text{hank} \] (TS) 'a whole lot'

Finally, I will note three apparent exceptions to the generalizations arrived at above. They all involve a short pre-root vowel followed by a long root vowel, perceived with stress on the first, i.e. short vowel, although the prediction is that stress should be on the root.

(21) Mojave

\[ ?\text{a}:\text{ti}: \] (TS-NB) 'wood'
\[ ?\text{k}:\text{ri}: \] (TS-NB) 'basket'
\[ ?\text{am}:\text{ma}:\text{t} \] (TS-NB) 'ground'

The long vowel in 'wood' is confirmed by comparative evidence; perhaps the two instances of glottal stop conspire to let stress be perceived on the first syllable. The word for 'basket' was perceived consistently with a long root vowel in the tape-recorded instances. However, Pamela Munro (private communication) maintains she always hears it short. Given the rules outlined above, her perception must be the correct underlying form, and the alternants with long vowel can be construed to be no more than an instance of the variation possible in pattern a. The same explanation is probably correct for the word for 'ground' which comparative evidence shows to have a short root vowel, and therefore to belong to pattern a, rather than being a variant of pattern c3, which would be the other possible alternative.

I hope to have shown that in Yuman languages, stress, length, and pitch are independent variables which interact with each other in complex ways. Stress is basically morphological, it is a property of the root; length is contrastive in the manner of other segmental phonemes; and pitch is dependent on particular configurations of sequences of syllables in the word and in connected discourse, but has become contrastive in some languages. A full understanding of these complex interactions in each of the languages as well as the unraveling of the historical developments underlying the present situation will require much detailed research. The remarks made in this paper are presented as an indication of the directions such research might profitably take.

Epilogue

After the original version of this paper was read at the Stressfest, independent evidence bearing on the Mojave problem became available from the instrumental investigations conducted by Birgitte Bendixen at the Phonetics Laboratory at UCSD using as input the two tapes which served as basis for the subjective transcriptions.
analyzed in the body of the paper. While the full results of this work would be too long to incorporate in this paper, it was felt that it would be of interest to give at least a sample which would afford a comparison of the subjective transcription with the reality of the signal. No substantive changes have been made in the transcriptions and analysis presented above.

The appended seven figures give a sample word for each of the pitch-stress patterns discussed. Each word is illustrated by one version for each speaker, Thomas Stevens (TS), and Nellie Brown (NB). The figures were drawn by Birgitte Bendixen. In each figure the top trace represents the fundamental frequency ($F_0$) of the voice (in Hz) and the bottom trace is the overall voice intensity (in dB).

An examination of the figures allows the following observations. When highest pitch and amplitude coincide on the same vowel, that vowel is invariably perceived as stressed, whether it is short or long (Fig. 1 'water' TS-NB, Fig. 2 'quail' TS, Fig. 3 'eagle' TS, Fig. 5 'man's father' TS, Fig. 7 'lungs' TS-NB). Highest pitch + length can be perceived as stress when no clue is provided by amplitude (Fig. 2 'quail' NB, Fig. 3 'eagle' NB). Highest amplitude + length can be perceived as stress when no clue is provided by pitch (Fig. 4 'he lay down' NB). Highest amplitude alone can be perceived as stress (Fig. 5 'man's father' NB, where this happens in spite of the fact that the second vowel is perceived as long). Length alone may be perceived as stress (Fig. 4 'he lay down' TS. Note also how in Fig. 7 'lungs' NB the second long vowel is perceived as just as stressed as the first one, even though pitch and amplitude are clearly lower.) Finally, and even though no example is found in the illustrations, our data show that pitch alone can be perceived as stress. This is the case in one version of 'we 'enemy' by both TS and NB and, Interestingly, in some versions of the three exceptions noted in (21).
Figure 1. Pattern a. 'water'

Figure 2. Pattern b. 'quail'
Figure 3. Pattern cl. 'eagle'

Figure 4. Pattern c2. 'he lay down'
Figure 5. Pattern c3. 'man's father'

Figure 6. Pattern d1. 'man'

Figure 7. Pattern d2. 'hunan'
FOOTNOTES

1 Financial support under NSF grant SOC 74-18043 is gratefully acknowledged. I would also like to thank Susan Norwood and Don Crook for help in culling from the Yuman archives at UCSD the relevant data for this paper and for many useful discussions. I have also greatly profited from comments by Alice Grundt.


3 For the purposes of this paper, I have standardized the orthography of Yuman languages in non-ambiguous ways as follows: since the topic of the paper is very specifically stress, I have marked stress on all forms, even though this is rarely needed. I have also marked long vowels with : throughout to avoid confusion with the intermediate length indicated by * in some languages.

4 Some investigators (e.g. Redden 1966 and Halpern 1946–47) describe suffixes which have primary stress so that some "words" in Yuman have two primary stresses. I have ignored these in this paper on the assumption that these forms function as two words.

5 These syntactic conditions are not explicitly stated by Crawford, but I believe they constitute a valid restatement of his description.

6 The pitch contours indicated on these sentences are also a restatement of Crawford's notation, which uses numbers for pitch level and various types of punctuation for final contours.

7 Throughout the remainder of this paper, pitch level and contour will be marked by lines both below and above the forms. Their interpretation is straightforward. It should be pointed out, however, that no inference about the actual difference in pitch between low and high should be made from this notation, which is adopted purely for convenience.

8 The alternation γ / l is an instance of the very common type of sound symbolism exhibited in Yuman languages. All data on Killwa presented in this paper are from Mixco (1971) or from copies of unpublished fieldnotes of his which he kindly deposited in the Yuman archives at UCSD.

9 The morphological analysis of this sentence, such as it is, is my own, not Redden's.
10 I am grateful to Alan Shaterian for giving permission to use forms from this list, though he warns that it has not been fully rechecked. I therefore do so at my own risk.

11 These labels reflect my interpretation of the symbols used by Kaufman and Shaterian, namely: / high pitch, \ rising pitch, ^ falling pitch.

12 Commenting on an earlier version of this paper, Joel confirms this observation: "It seemed to me that the high-level was associated with short vowels, and the mid-rising with long vowels. In fact, I came to rely on pitch to tell me whether the vowel was short or long." (Joel, personal communication). The Palpal examples in 14 are from the same communication by Joel.

13 Mojave forms not transcribed from the tapes identified below have been checked against field notes of Pamela Munro.

14 The only exception is the word for 'willow' for which the tapes only have versions by TS, each heard with a final short vowel. Other recordings (e.g. Wares 1968) as well as comparative evidence all point to a long underlying root vowel which is the justification for placing this form under pattern dl.

15 We are grateful to Timothy Smith for guidance in the technical aspects of this study.

REFERENCES

Shaterian, A. V. ms. Yavapai Phonology.
ACCENT IN JAPANESE

James D. McCawley
University of Chicago

1. Accent in the modern standard language

1.1 General considerations

Standard Japanese is a particularly clear example of a pitch-accent system. Phrases differ from each other accentually only as regards place where pitch falls; that is, leaving out the effects of intonational contours which may be superimposed on a sentence, the pitches on a phrase are predictable from its segmental composition plus information as to where (if anywhere) the pitch falls. For example, among phrases consisting of four short syllables (illustrated here by phrases consisting of a 3-syllable noun followed by a 1-syllable case marker) there are only the following four distinct melodies:

(1)  
- makura ga  'pillow' (Nom.)
- kokoro ga  'heart'
- atama ga  'head'
- sakana ga  'fish'

These can be characterized as exhibiting (respectively) fall in pitch after the first syllable, after the second, after the third, and nowhere, and can be represented as follows, using ' to mark place where pitch falls:

(2)  
- ma'kura ga
- koko'ro ga
- atama' ga
- sakana ga

The pitch on a phrase can be predicted from the accent mark by the following ordered rules:

(3)  
1. Make everything high pitched.
2. Make everything after a ' low pitched.
3. Make the first syllable low pitched if the second is high pitched.

Consideration of long syllables necessitates some revision of (3). Japanese exhibits both short syllables (consisting of consonant plus short vowel, or of just short vowel) and long syllables (consisting of (C)VV or of (C)VC, where long vowels are represented as geminate). The distinction between long syllable and short syllable is a distinction between syllables that 'count as two units' and syllables
that 'count as one unit' for certain purposes, as in the scansion of Japanese poetry, which typically involves lines alternately of 5 and 7 units such as the 5-7-5 pattern of haiku. The units just referred to are called moras. The only reasonable definition of 'mora' that has been proposed is: 'something of which a long syllable consists of two and a short syllable consists of one'. Since a long syllable consists of what would make up a short syllable, plus additional material, one can take the initial (C)V- of a long syllable to be its first mora and the remaining -V or -C to be its second mora. For example, gakko 'school' divides into the two syllables gak-koo and the four moras ga-k-ko-o.\(^2\)

Among speakers of standard Japanese, there is some variation as to whether an unaccented CVV or CVn first syllable is left high pitched or its first mora made low pitched, i.e. both sensei and sensei are common pronunciations of sensei 'teacher', the latter being more common among younger speakers. Pitch may fall after the first mora but not after the second mora of a long syllable. Thus, while the a-forms in (4) are real phrases, the b-forms are phonologically inadmissible:

(4)  
\[
\begin{align*}
\text{a. } & \text{ se}^\prime \text{i} \text{k} \text{a} \text{ ga} & \text{'floral arrangement'} \\
& \text{ sense}^\prime \text{i} \text{ ga} & \text{'teacher'} \\
& \text{ sl}^\prime \text{nnen ga} & \text{'New Year'} \\
\text{b. } & *\text{se}^\prime \text{i} \text{k} \text{a} \\
& *\text{sense}^\prime \text{i} \text{ ga} \\
& *\text{sl}^\prime \text{nnen ga}
\end{align*}
\]

Each syllable thus affords only one possible place for the pitch to fall: at the end of its first mora. One may thus take the accent to be borne by the syllable: specifying the accent of a phrase amounts to specifying the syllable, if any, where (=after the first mora of which) pitch falls. (3) thus must be revised as follows:

(5)  
\[
\begin{align*}
\text{i. } & \text{ Make everything high pitched.} \\
\text{ii. } & \text{ Make everything after the first mora of the 'ed syllable low pitched.} \\
\text{iii. } & \text{ (for some speakers) Make the first mora low pitched if the second is high pitched; (for other speakers) make the first mora low pitched if the second is high pitched and not of the form -V or -n.}
\end{align*}
\]

Regardless of whether sakana ga is treated as unaccented or as having an accent on ga, the above rules will assign the same pitches to it. It is in fact essential that the rules work in this way, since the distinction between underlying final accent on a short syllable and underlying accentless is neutralized. Since the accents manifested in (l) are obviously contributed by the four nouns, the nouns require the underlying forms

(6)  
\[
\begin{align*}
& /\text{ma}^\prime \text{k} \text{ura}/ \\
& /\text{koko}^\prime \text{ro}/ \\
& /\text{atama}^\prime/ \\
& /\text{sakana}/
\end{align*}
\]
When pronounced in isolation, these four nouns take exactly the same pitches as are obtained by subjecting the underlying forms (6) to the rules of (5):

(7) makura kokoro atama sakana

and the accentual distinction between atama and sakana is obliterated.

1.2 Accent on simple words.

I turn now to the question of the lexical representation of accent. The discussion in the last paragraph suggests that for any positive integer n, nouns of n syllables divide into n + 1 underlying accentual types: 1-st syllable accent, 2nd syllable accent, ... n-th syllable accent, and unaccented. This is in fact the case. The following examples illustrate the possibilities for nouns of one through four syllables:

(8) 1-syllable:  

   e ga 'picture' /e/  
   e Ga 'handle' /e/  

2-syllable  

   kaki ga 'oyster' /ka'ki/  
   kaki ga 'fence' /ka'ki/  
   kaki ga 'persimmon' /kaki/  

3-syllable: (see (1) and (6))

4-syllable:  

   kamakiri ga 'preying mantis' /ka'makiri/  
   kudamono ga 'fruit' /kuda'mono/  
   kagaribi ga 'campfire' /kagaribī/  
   kamisori ga 'razor' /kamisori/  
   kamaboko ga 'fish pudding' /kamaboko/  

So far the only NP's which I have considered are those consisting of N + ga. Substitution of other particles or combinations of particles for ga makes clear that particles also exhibit underlying accent differences:

(9)  

<table>
<thead>
<tr>
<th>Noun</th>
<th>'from...'</th>
<th>'to...'</th>
<th>'only...'</th>
</tr>
</thead>
<tbody>
<tr>
<td>the pillow</td>
<td>ma'kura kara</td>
<td>ma'kura made</td>
<td>ma'kura sika</td>
</tr>
<tr>
<td>the heart</td>
<td>koko'ro kara</td>
<td>koko'ro made</td>
<td>koko'ro sika</td>
</tr>
<tr>
<td>the head</td>
<td>atama' kara</td>
<td>atama' made</td>
<td>atama' sika</td>
</tr>
<tr>
<td>the fish</td>
<td>sakana kara</td>
<td>sakana ma'de</td>
<td>sakana' sika</td>
</tr>
</tbody>
</table>

When added to an accented noun, kara, made, and sika are pronounced on a low pitch, with the pitch falling in the place determined by the noun. However, when they are added to an unaccented noun, they display distinct accentual behaviors: made has an
accent on its first syllable, sika is preceded by an accent, and kara exhibits the high-ending shape which is ambiguous between unaccented and final-accented. Since kara can in fact be followed by the topic-marker wa, it is possible to test whether sakana kara is unaccented or final-accented, and it turns out in fact to be final-accented: sakana kara' wa. The accent on the phrases in (9) can then be accounted for by setting up the underlying forms

(10) /kara'/, /ma'de/, /'sika/ (i.e. sika carries with it an accent that goes on the preceding syllable)

and having a rule that deletes all but the first accent of a phrase that contains more than one accent. Selected phrases in (9) will then have the following derivations:

(11) ma'kura ma'de sakana ma'de koko'ro 'sika sakana 'sika
    → ma'kura made -- koko'ro sika --

Among 2-syllable particles there thus appear to be at least three underlying accentual types: first-syllable accent, second-syllable accent, and preaccented. This is in fact all that there is, subject to some qualifications given at the end of this section. In the case of 1-syllable particles, there are only two underlying accentual types, first-syllable accent and preaccented:

(12) /ni/ (dative) ka'ki ni wa 'to the oyster'
    kaki ni wa 'to the fence'
    kaki ni' wa 'to the persimmon'

/'ra/ (pluralizer)3 kore wa 'this': kore'ra 'these'
    bo'ku wa 'I': bo'kura 'we'

There are certain one-syllable particles which can not be shown definitely to be final-accented rather than unaccented, since, for syntactic reasons, they cannot be followed in surface structure by any other particle. This is the case with wa (topic marker), ga (nominative), and mo 'even, also'. However, there are no particles which have to be analyzed as unaccented, which means that wa, ga, and mo must be treated as final-accented, since to treat them otherwise would require setting up an accentual type for particles which is otherwise unnecessary. Thus, to the extent that one can generalize about accentual possibilities within such a limited domain as particles, it appears that n-syllable particles divide into n + 1 accentual types: first syllable accent,... n-th syllable accent, and preaccented. In §1.3 I will take up the question of whether the extra type for nouns ('unaccented') can be identified with the extra type for particles ('preaccented').

Verbs and adjectives, by contrast, have only two accentual possibilities regardless of length. Some verb and adjective stems contribute an accent to the word and some do not, and the place of the accent is predictable in the case of those which do. The accentual possibilities for vowel-stem verbs, consonant-stem verbs, and adjectives are as follows:
<table>
<thead>
<tr>
<th>(13)</th>
<th>Vowel-stem verbs</th>
<th>Consonant-stem verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>accented</td>
<td>unaccented</td>
</tr>
<tr>
<td></td>
<td>'eat'</td>
<td>'open'</td>
</tr>
<tr>
<td>Present</td>
<td>tabe'-ru</td>
<td>ake-ru</td>
</tr>
<tr>
<td>Past</td>
<td>ta'be-ta</td>
<td>ake-ta</td>
</tr>
<tr>
<td>Cond.</td>
<td>ta'be-tara</td>
<td>ake-ta'ra</td>
</tr>
<tr>
<td>Provis.</td>
<td>tabe'-reba</td>
<td>ake-re'ba</td>
</tr>
</tbody>
</table>

Adjectives

<table>
<thead>
<tr>
<th></th>
<th>accented</th>
<th>unaccented</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'high'</td>
<td>'dark'</td>
</tr>
<tr>
<td>Present</td>
<td>tak'a'-i</td>
<td>kura'-i</td>
</tr>
<tr>
<td>Adv.</td>
<td>ta'ka-ku</td>
<td>kura-ku</td>
</tr>
<tr>
<td>Past</td>
<td>ta'ka-kat-ta</td>
<td>kura'-kat-ta</td>
</tr>
<tr>
<td>Cond.</td>
<td>ta'ka-kat-tara</td>
<td>kura'-kat-tara</td>
</tr>
<tr>
<td>Provis.</td>
<td>ta'ka-ker-eba</td>
<td>kura'-ker-eba</td>
</tr>
</tbody>
</table>

Except for the present tense (/ru/ for verbs, /i/ for adjectives) and for the absence in the adjective of the imperative and some other forms and the absence in the verb of an adverbial form, the same endings appear with adjectives as will verbs (past tense /ta/, conditional /ta'ra/, provisional /re'-ba/), though in adjectives some material (the remains of the adverbial ending /ku/ plus an auxiliary verb) intervenes between stem and ending.

The conditional and provisional endings and the material between adjective stems and endings exhibit accents after unaccented stems but not after accented stems. This fact is accounted for by the accent-deletion rule discussed above, provided those endings are assigned accented underlying forms (/ta'ra/, /re'ba/, /'kat/, /'ker/) and other endings are assigned unaccented underlying forms. For example, ta'kakata'ra will arise through application of accent deletion to /ta'ka-ka-ta'ra/. In addition, there is an alternation between stem-final accent and stem-penultimate accent in adjectives and vowel-stem verbs. That the alternation involves stem-penultimate accent and not, say, stem-initial accent, is made clear by a consideration of longer words:

(14) sirabe'-ru, sir'a-be-ta 'investigate'
tsadasl'-i, tada'si-ku 'correct'

There is a small amount of evidence for taking the stem-penultimate accent to be more basic than the stem-final accent. Specifically, besides the 'inflected adjectives' discussed so far, Japanese also has 'uninflected adjectives', which are always used in combination with a copula. A certain number of adjective stems can be used both as inflected adjectives and as uninflected adjectives, and when used as uninflected adjectives they exhibit stem-penultimate accent:
ooki'-i, o'oki-ku 'large'; o'oki na kaban 'large suitcase'
tiisa'-i, ti'lsa-ku 'small'; ti'lsa na kaban 'small suitcase'
okasi'-i, oka'si-ku 'strange'; oka'si na hanasi' 'strange story'

The dependent copula na has no effect on the accent of the preceding item; in particular, it does not affect final accents such as that of beta' 'unskillful': beta' na e' 'poor(ly drawn) picture'. Thus, if both the uninflected and the inflected forms of these adjectives have the same basic position of accent, the basic position could not be stem-final (since there is nothing in o'oki na kaban etc., to cause the accent to shift off of the final syllable of the adjective), whence it must be stem-penultimate and the present tense morpheme /i/ must attract accent towards it: ta'ka -i → taka'-i. If verbs and adjectives are as parallel as they appear to be, this means that stem-penultimate position must be basic in vowel-stem verbs also and that the present tense ending /ru/ and the provisional ending /re'ba/ attract accent towards them: ta'be-re'ba → tabe'-re' ba (→tabe'-reba).

A number of verbs and adjectives do not conform to the above accentual paradigms. One case of regular deviation from the paradigms is that in which the regular accent would fall on a voiceless syllable. Short high vowels flanked on both sides by voiceless consonants are normally devoiced in Japanese. In the following examples, capitalization indicated voicelessness of the vowel and " indicates where the accent would be according to the rules given so far, when that differ from the actual accent:

(16) 'near' 'deep' 'low' 'obstinate' 'foul-smelling'
    tlka'i    hU'ka'i    hku'i    si'tUko'i    kUsa'i
    tI'ka'ku    hU'ka'ku    hI'ku'ku    si'tU'ko'ku    kU'sa'ku
'thick' 'wise' 'rescue' 'get moist' 'blow'
    hUto'i    kasIko'i    tasUke'ru    sI'ke'ru    hU'ku'
    hU'to'ku    kasI'ko'ku    tasU'ke'ta    sI'ke'ta    hu'ita
'adhere' 'conceal'
    tu'ita    kaku'sita

When related forms exist in which the syllable in question has a voiced vowel, the accent fits the usual paradigm:

(17) tyuu-l-buka'i, tyuu-l-bu'kaku 'cautious' (lit. 'caution-deep')
    kl-zu'ku, kl-zu'ita 'notice' (lit. 'spirit-adhere')

There is thus evidently a rule shifting accent to the right when it falls on a voiceless syllable. However, this rule has exceptions, both systematic and idiosyncratic. Accent is never shifted off of an adjective stem onto an ending nor off of a noun onto a particle:

(18) 'was okay' 'from the town'
yoroS'i'katta  matl' kara
*yoroska'tta  *matl ka'ra
A number of adjectives having stem-final accent throughout their entire paradigms are listed in McCawley (1968:155-6); many (but not all) of these are subject to minor regularities such as that adjective stems in ...na- have uniform stem-final accent.

Before going on to the accent of compound forms, I will note briefly a few additional points about the accent of simple nouns, verbs, and adjectives.

1. There are a number of morphemes whose accent predominates over that of the words that they are attached to, e.g.

   (19) a. 'to the extent of ...' (used in comparative sentences)
   makura gu'rai 'as ... as a pillow'
   sakana gu'rai 'as ... as a fish'

   b. 'only'
   makura dake', makura ba'kari 'only a pillow'
   sakana dake', sakana ba'kari 'only a fish'

   c. 'let's'/'probably is'
   tabeyo'o 'let's eat'
   akeyo'o 'let's open'
   tanomo'o 'let's ask for'
   susumo'o 'let's advance'

   d. 'probably does not'
   taberu ma'i
   akeru ma'i
   tanomu ma'i
   susumu ma'i

   e. (politeness marker)
   tabema'su/tabema'sita 'eats/ate'
   akema'su/akema'sita 'opens/opened'
   tanomima'su/tanomima'sita 'asks/asked for'
   susumima'su/susumima'sita 'advances/advanced'

2. Final accent in a noun is lost before the genitive marker no, except in the case of one-mora nouns or where the genitive morpheme is attached to a noun-phrase that consists of more than just the noun:

   (20) atama no ooklsa 'the size of the head'
   o' no ooklsa 'the size of the tail'
   uma no sippo' 'a horse's tail'
   ano uma' no sippo' 'the tail of that horse'
3. A couple of morphemes attract the accent of the preceding word. The distance which the accent moves can exceed the 1-mora attraction noted above in the present and provisional (tabe-ru → tabe'-ru):

(21) a. 'appears'
   nakisoo desu 'he looks like he's going to cry'
   kakiso'o desu 'he looks like he's going to write'
   omotasoo desu 'it looks heavy'
   okasio'o desu 'it looks strange'
   (present tense forms: naku, ka'ku, omota, okasi)

b. 'while, although'
   akenagara 'while opening'
   tabena'gara 'while eating'
   syatyo nagara 'although being a company president' (syatyo)
   sensei na'gara 'although being a teacher' (sensei)
   nikuya na'gara 'although being a butcher' (nikuya)

4. The negative form of verbs is not subject to the usual accent attraction, though otherwise it is conjugated like other adjectives (it is accented if formed from an accented verb, unaccented if formed from an unaccented verb):

(22) yoma'nai/yoma'nakatta 'does/did not read'
    tabe'nai/tabe'nakatta 'does/did not eat'
    tanoma'nai/tanoma'nakatta 'does/did not ask for'

This fact provides additional (though rather weak) evidence that the stem-penultimate position of accent is basic: under that hypothesis, negative forms are merely exceptions to an existing rule, whereas if the stem-final form were basic they would require an additional rule of accent placement.

1.3 Compound and derived words.

Derived verbs are of the same 'accentedness' as the verb from which they are derived, though of course the accent will appear on a different syllable, as a result of the principles of accent placement in verbs. The same is true of verbs derived from adjectives and of adjectives derived from verbs, though in saying this I am assuming a somewhat arbitrary distinction between compounding and derivation:

(23) a. causative
   akesaseru, akesaseta 'makes/made open
   tabsasase'ru, tabsasaset 'makes/made eat'
   ugokaseru, ugokaset 'makes/made move
   tanomase'ru, tanomaset 'makes/made ask for'
b. 'passive'\(^9\)
- akerareru, akerareta
- taberare'ru, tabera'reta
- ugokareru, ugokareta
- tanomare'ru, tanomare'ta

c. passive of causative
- akesaserare, akesaserareta
- tabesaserare'ru, tabesasera'reta
- ugokaserareru, ugokaserareta
- tanomaserare'ru, tanomaserare'ta

d. (various non-productive and semi-productive morphemes)
- ak-u 'open', ak-e-ru 'open'
- kawa'k-u 'dry', kawak-a's-u 'dry'
- mo(y)-e-ru 'burn', moy-as-u 'burn'
- kaku-re'-ru 'be concealed'; kaku'-s-u 'conceal'

e. verbs derived from adjectives
- kata-i 'hard', kata-mar-u 'harden'
- haya'-i 'fast, early', haya-mar-r-u 'be hasty, rash'
- maru-l 'round', maru-me-ru 'make round'
- huto'-i 'thick', huto'-r-u 'grow fat'

f. adjectives derived from verbs
- ka(w)-u 'buy', ka(w)-i-ta-i 'want to buy'
- yom-u 'read', yom-i-ta-i 'want to read'

Compound verbs and adjectives follow quite different rules than do derived verbs and adjectives. In older speakers, a verb compounded from two verbs has the opposite of its first member's accentuation:

\[(24)\]
- naki-ya'm-u 'stop crying' (unaccented + unaccented)
- naki-da's-u 'burst into tears' (unaccented + accented)
- kaki-yam-u 'stop writing' (accented + unaccented)
- kaki-das-u 'start writing' (accented + accented)

However, for many younger speakers, compound verbs are generally pronounced accented regardless of the underlying accent of their components. Compounds of verb + adjective or adjective + verb are generally accented:

\[(25)\]
- kai-niku'-i 'difficult to buy' (kau)
- yomi-niku'-i 'difficult to read' (yomu)
- hakai-yasu'-i 'easy to put on' (haku)
- kai-yasu'-i 'easy to write' (ka'ku)
- too-sugi'-ru 'too distant' (tool)
- taka-sugi'-ru 'too expensive/tall' (taka'i)
- kai-ta-ga'r-u 'behaves like he wants to buy' (kai-ta-i)
- yomi-ta-ga'r-u 'behaves like he wants to read' (yomi-ta-i)
- samu-ga'r-u 'behaves like he is cold' (samu'-i)

The accentual phonology of compound nouns is far more intricate than that of compound verbs or adjectives. There appears to be little regularity in the accentuation
of one very important class of compounds, namely compounds of two Sino-Japanese morphemes, though there are certain morphemes for which some rule can be set up (e.g. compounds beginning with i- 'than' all have initial accent: ɪ'zyoo 'at least', ɪ'nai 'within', ɪ'gai 'except', ɪ'rai 'from that time onwards', etc.): I will henceforth ignore this class of compounds. The accent of compound nouns is largely determined by the accent of compound nouns is largely determined by the accent and length of the final member: only in extra-short compounds and before a small number of final elements does the accent of the initial member play any role. ¹⁰

It is useful in what follows to draw a distinction between 'long' and 'short' final members of compounds. A 'long' final element is an item which either is at least 3 moras long or is a Sino-Japanese compound. When a compound has a long final member, it has an accent on that element. If the long final member is unaccented or is accented on its final syllable when used alone, it has an accent on its first syllable when it is the final member of a compound; otherwise the accent is on the syllable which bears the accent when the final member is used independently:

(26) no'ogyoo + kumial → moogyoo-ku'mial 'agricultural union'
roodoo + kumial → roodoo-ku'mial 'labor union'
tyu'uka + ryo'orl → tyuuka-ryo'orl 'Chinese cooking'
huransu + ryo'orl → huransu-ryo'orl 'French cooking'
genzi + monoga'tari → genzi-monoga'tari 'the Tale of Genji'
iso'ppu + monoga'tari → isoppu-monoga'tari 'Aesop's fables'
hanauri' + musume' → hanauri-mu'sume 'girl who sells flower'
inaka + musume' → inaka-mu'sume 'country girl'

This generalization can be expressed in the form of the ordered rules:

(27) i. In a compound noun X # Y in which Y is long, the accent of Y predominates (i.e. the accent of X is eliminated).

ii. In a compound noun X # Y in which Y is long and either accented on its final syllable or unaccented, put accent on the first syllable of Y.

The reference to syllables in (27ii) is intentional: it is accent on the last syllable rather than on the last mora that conditions the application of (27ii), since final elements with an accented long final syllable behave like musume rather than like monoga'tari:

(28) asobi + hanbu'n → asobi-ha'nbun 'half for pleasure'
zyooda'n + hanbu'n → zyooodan-ha'nbun 'half as a joke'
hutai + zyooke'n → hutai-zyo'oken 'incidental condition'
ka'izyo + zyooke'n → kaizyo-zyo'oken 'cancellation condition'
Let us now turn to compound nouns with a long first member and a short final member. With a few exceptions (such as -zin; see fn. 10), the final element of these compounds determines the accent of the compound, without regard to the underlying accent of the first member. Leaving aside the exceptional items just alluded to, the final members can be divided into the following classes: (a) those which yield an unaccented compound, (b) those which yield an accent on the immediately preceding syllable of the compound, and (c) those which take an accent on their first syllable:

\[(29)\]  

\[a.\] sansui-ga 'landscape' (sa'nsui)  
   pasutera-ga 'pastel' (pa'sutera)  
   syoozoo-ga 'portrait' (syoozoo)  
   huukei-ga 'landscape' (hu'ukel)  
   ratai-ga 'nude' (ratal)  
   huusen-dama 'balloon' (huusen)  
   zyuzu-dama 'rosary bead' (zyuzu')  
   garasu-dama 'glass bead' (garasu)  
   nanki'n-dama 'glass bead' (nanki'n)

\[b.\] kensetu'-syoo 'Ministry of Construction' (kensetu)  
   gaimu'-syoo 'Foreign ministry' (ga'imu)  
   ookura'-syoo 'National Exchequer'ill  
   noori'n-syoo 'Ministry of Agriculture and Forestry'  
   monbu'-syoo 'Ministry of Education'

\[c.\] momen-i'to 'cotton thread' (momen)  
   tumugi-i'to 'pongee thread' (tumugi')  
   situke-i'to 'basting thread' (situke)

Some of these final members can occur as independent words: others only occur as bound morphemes. Chew (1964) noted an important correlation between the accentual behavior of these three kinds of elements in compounds and their behavior as independent nouns: most items of type (a) are final-accented when used independently, most items of type (b) are unaccented, and all items of type (c) have first-syllable accent when used independently. Moreover, the exceptions to this correlation are not randomly distributed: they are mainly final-accented or initial-accented nouns which nonetheless are preceded by an accent when used as the final member of a compound. In McCawley (1968), I proposed integrating this correlation into a general treatment of compounds in the following way. Recall that both for nouns and for particles, the number of accentual types among items of \(n\) syllables was \(n + 1\): 1st syllable accent, 2nd syllable accent, . . . ,
n-th syllable accent, and an additional type, which for nouns was 'accented' and for particles was 'preaccented'. Suppose that the 'additional types' were both given the same underlying accentual representation, namely preaccented (i.e. the underlying form of 'ball' would be /'tama/). Then both type (b) and type (c) would conform to a generalized version of rule (27): the underlying accent of the final member would predominate. It is thus possible to replace (27) by the following more general system of rules:

(30) i. In a noun compound $X \# Y$, the accent of $Y$ predominates.
   ii. If $Y$ is long and final-accented or unaccented, put accent on the first syllable of $Y$.
   iii. If $Y$ is short and final-accented, deaccent the whole compound (i.e. put accent before $X$).

Most of the exceptions to Chew's correspondence can be integrated into this scheme by representing them with two accents, e.g.

(31) syota'i-nusi 'householder' (nu'si) /'nu'si/
    hayar'i-uta 'popular song' (uta') /'uta'/

If unaccented words are to be represented as underlingly pre-accented, the pitch assignment rules will have to ignore a word-initial '. Thus, they will assign the same pitch to /'nu'si, 'uta'/ as they would have assigned to /nu'si, uta'/. (301) will make both accents predominate in compounds such as those of (31), but only the first of those accents will have a phonetic realization, since accent-deletion eliminates all accents in a phrase other than the first one. (The rule, of course, will have to be stated in such a way that it ignores accents at the beginning of a word: it will thus affect nusi and uta only when they are at the end of a larger word).

1.4 Sentence and phrase accent

A surface constituent of more than one word, exclusive of 'particles', often allows the following options as to how it is pronounced: (a) as two or more separate phrases, each with its own accent (if any), (b) as a single group of phrases, with the accented syllable of the first phrase in the group higher in pitch than the accented syllables of the subsequent phrases, and (c) as a single phrase, with the accent of the first accented constituent of the phrase predominating:

(32) a. ma'do ga @ kowa'rete iru 'the window is broken'
    b. ma'do ga % kowa"rete iru
    c. ma'do ga kowarete iru

My description of option (c) presupposes a rule that I have not yet given: accent on the last mora of a phrase is lost, even if the phrase boundary closing that phrase is subsequently lost. Thus, 'He lives in Tokyo' is pronounced tookyoo ni sumima'su and not *tookyoo ni' sumimasu (cf. tookyoo ni' wa). A syntactically complex expression allows options (b) and (c) to be combined, in accordance with its constituent structure:
(33) do'o % itta'ra % i'i desu ka? 'How would it be best to go?'
do'o ittara % i'i desu ka?
do'o ittara i'i desu ka?
*do'o % itta'ra i'i desu ka?

Since doo ittara 'if (one) went how' is a surface constituent, the phrase boundary between itta'ra and i'i cannot be eliminated unless that between do'o and itta'ra is also: the elimination of phrase boundary in option (c) consists in eliminating all phrase boundaries from a surface constituent, and in (33) the only constituent containing the boundary between itta'ra and i'i is the whole sentence.

Both the lowered pitch of option (b) and the accent elimination of option (c) involve the predomination of the first accent in a surface constituent. The rule that the first accent in a constituent predominates is the Japanese analogue to the English 'nuclear stress rule' (Chomsky and Halle 1968), according to which the last accent in a constituent predominates. Not only are the 'nuclear stress rules' of English and Japanese mirror images, but so are the rules for compound nouns: in English, the first member of a compound noun generally predominates, whereas in Japanese the final member generally predominates. Interestingly, both languages share a systematic exception to the compound rule: 'dvandva compounds' (i.e. compounds paraphraseable by an expression in which the two members are conjoined) are accented like phrases rather than like compounds:

(34) secretary-treasurer
     yo'ru-hiru 'night and day'

2. Accent in Japanese dialects

The general outlines of the accentual system of standard Japanese also apply to the accentual systems of the dialects of Hokkaidō, Eastern Honshū (except for a large area running from Chiba prefecture to Miyagi prefecture, in which all accentual contrasts have been lost), Western Honshū, the western tip of Shikoku, and eastern Kyūshū. In all of those dialects, only the place of fall in pitch is distinctive, and the place where pitch falls more often than not agrees with the standard language (but see §3 for some systematic differences in place of fall in pitch), though the various dialects differ quite a bit as regards how the pitch on the various moras is related to the place where pitch falls. For example, in Aomori, the whole phrase is low-pitched except for (the first mora of) the accented syllable; in Tarō (Iwate pref.) and in Nakamura (Kōchi pref.), the syllables preceding the accented syllable have the opposite of the standard pitch: they are low except that if two or more moras precede the accented syllable, the first mora is high:

(1) Standard Aomori Tarō
    atama ga adama ga adama ga

In two large areas, significantly different accentual systems are found: the area
encompassing central Honshū and most of Shikoku (this is the area traditionally called 'Kansai'), and the area encompassing western and southern Kyūshū, the Amami Islands, and the Ryūkyū Islands.

2.1 Kansai dialects

In the dialects of the Kansai area (including Kyōto, Osaka, Wakayama, and most of Shikoku), not only the location of fall in pitch (if any) is distinctive, but also the pitch level (high or low) on which the phrase starts. In the dialect of Befu, Hyōgo pref. (data from Hirayama 1960), one of the more conservative Kansai districts as regards accent, the following accentual types occur in 1-, 2-, and 3-syllable nouns:

(2) ka, ka ga 'fly' kaki, kaki ga 'persimmon' sakana, sakana ga 'fish'
    ha, ha ga 'leaf' hasi, hasi ga 'bridge' azuki, azuki ga 'red bean'
    ta, ta ga 'field' hasi, hasi ga 'chopsticks' awabi, awabi ga 'abalone'
    mado, mado ga 'window' usagi, usagi ga 'rabbit'
    kabuto, kabuto ga 'helmet'

In each case there are more distinct accentual types than could be distinguished in terms of just place of fall in pitch. Since there are pairs of words which do not differ as regards whether the word starts on a high or on a low pitch (e.g. 'persimmon' vs. 'chopsticks'; 'red bean' vs. 'helmet'), the distinction between low-initial and high-initial is an obvious choice for an additional characteristic to take as one of the distinctive differences among phrases in this type of dialect. It is reasonable to use the symbol for 'fall in pitch' also to stand for 'low initial' when written at the beginning of a phrase, since there in fact would be a fall in pitch if something ending on a high occurred before the word in question, e.g.

(3) ano usagi 'that rabbit' cf. ano awabi 'that abalone'
    ano kabuto 'that helmet'

I will thus transcribe (2) as follows:

(4) ka kaki sakana
    ha' ha'si azu'ki
    'ta 'hasi a'wabi
    'mado 'usagi
    'kabuto

The relationship between accent marks and pitches is given by the following ordered rules:

(5) i. Make everything high pitched.

ii. Make low-pitched all moras after a ' except (a) the last mora of the phrase, if the phrase has no ' other than one preceding the phrase, and (b) a mora bearing a '.
Combinations of noun plus particle are subject to the same rule of accent deletion as in the standard language, with the qualification that the preposed accent signifying low-initial causes no loss of underlying accent in the particle:

(6)  

<table>
<thead>
<tr>
<th>Noun</th>
<th>Particle</th>
<th>Underlying Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>usi</td>
<td>'yori</td>
<td>/'yori</td>
</tr>
<tr>
<td>u'ma</td>
<td>'yorl</td>
<td>/'yorl</td>
</tr>
<tr>
<td>'hasi</td>
<td>'demo</td>
<td>/'demo</td>
</tr>
<tr>
<td>'saru'</td>
<td>'kara</td>
<td>/'kara</td>
</tr>
</tbody>
</table>

The accent deletion rule thus acts as follows:

(7)  

\[
\text{u'ma + 'yorl + 'saru + 'hasi + 'demo + usi + 'yori} \\
\rightarrow \text{u'ma + yori + 'saru + demo} \\
\]

The presumable underlying forms for 'than' and 'even' are /'yorl, demo/. In the case of 'from', the data given in (6) do not make clear whether its underlying form should be /'kara/ or /'kara'/. In the speech of the one Kansai speaker from whom I have obtained the relevant data (a native of Kyōto: Kyōto agrees with all that I have said so far about Kansai dialects, except that in Kyōto all nouns that at one time were 00'0 have shifted to 0'00), the following are the pronunciations:

(8)  

<table>
<thead>
<tr>
<th>Noun</th>
<th>Particle</th>
<th>Underlying Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>usi</td>
<td>'kara-wa</td>
<td>/'kara</td>
</tr>
<tr>
<td>u'ma</td>
<td>'kara-wa</td>
<td>/'kara</td>
</tr>
<tr>
<td>'hasi</td>
<td>'kara-wa</td>
<td>/'kara</td>
</tr>
<tr>
<td>'saru'</td>
<td>'kara-wa</td>
<td>/'kara</td>
</tr>
</tbody>
</table>

Since there is no fall in pitch after kara in the forms with 'cow' and 'chopsticks', and since wa in these dialects does not introduce or remove accent, the underlying form of kara must be unaccented. The dative marker ni is likewise unaccented for this speaker (cf. (12) in §1.2 for data showing that the standard language has /ni'/ as underlying form):

(9)  

<table>
<thead>
<tr>
<th>Noun</th>
<th>Particle</th>
<th>Underlying Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>usi</td>
<td>ni</td>
<td>/ni</td>
</tr>
<tr>
<td>u'ma</td>
<td>ni</td>
<td>/ni</td>
</tr>
<tr>
<td>'hasi</td>
<td>ni</td>
<td>/ni</td>
</tr>
<tr>
<td>'saru'</td>
<td>ni</td>
<td>/ni</td>
</tr>
</tbody>
</table>

Hattori (1960) cites facts which also confirm an unaccented underlying form /kara/ but in addition suggest that the underlying form for 'than' should be /'yorl/ rather than /'yorl/. Specifically, he cites (pp. 429, 436) the following as the pronunciation of the above forms when emphasis is placed on the particle:

(10)  

<table>
<thead>
<tr>
<th>Noun</th>
<th>Particle</th>
<th>Underlying Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>usi</td>
<td>yori</td>
<td>/yori</td>
</tr>
<tr>
<td>u'ma</td>
<td>yori</td>
<td>/yori</td>
</tr>
<tr>
<td>gama</td>
<td>yori</td>
<td>/yori</td>
</tr>
<tr>
<td>'saru'</td>
<td>yori</td>
<td>/yori</td>
</tr>
<tr>
<td>usi</td>
<td>demo</td>
<td>/demo</td>
</tr>
<tr>
<td>u'ma</td>
<td>demo</td>
<td>/demo</td>
</tr>
<tr>
<td>gama</td>
<td>demo</td>
<td>/demo</td>
</tr>
<tr>
<td>'saru'</td>
<td>demo</td>
<td>/demo</td>
</tr>
<tr>
<td>usi</td>
<td>kara</td>
<td>/kara</td>
</tr>
<tr>
<td>u'ma</td>
<td>kara</td>
<td>/kara</td>
</tr>
<tr>
<td>gama</td>
<td>kara</td>
<td>/kara</td>
</tr>
<tr>
<td>'saru'</td>
<td>kara</td>
<td>/kara</td>
</tr>
</tbody>
</table>
The behavior of demo indicates that when a particle is emphasized, its underlying accent is preserved. However, yori exhibits not only an underlying preposed accent (manifested by the high pitch on the second syllable of gama) but also a falling pitch on its final syllable, like the falling pitch on ‘leaf’ and on the final syllable of ‘window’ and ‘monkey’ when they are used in isolation. The second accent on yori would be eliminated except when, by virtue of its being made a separate word by emphasis, its second accent is not in the environment for accent deletion. By contrast, kara exhibits no falling pitch on its final syllable. My Kyōto informant has a slightly different pronunciation of the forms with kara, though the differences only point to a difference in the rules for the realization of accent marks in terms of concrete pitches: usi kara, uma kara, gama kara, saru kara.

The identification of low-initial with fall in pitch is supported by some facts about accent in adjectives in the Kyōto dialect. In Kyōto, the distinction between what in standard Japanese are accented and unaccented adjectives has been lost\(^{1}\). Monosyllabic adjectives are low-initial in the present tense, and longer adjectives are high initial:

\[(11)\]
\[
\begin{array}{l}
\text{Present} \\
\text{Adverbial}
\end{array}
\]
\[
\begin{array}{ll}
\text{ee} & \text{‘good’ ( < ‘yo + i)} \\
\text{a'kai} & \text{‘red’} \\
suru'doi & \text{‘sharp’}
\end{array}
\]

The identification of low-initial with fall in pitch makes it possible to state a single generalization covering both the monosyllabic case and polysyllabic case: accent goes one mora before the present tense ending. There is also an alternation between low-initial and high-initial in monosyllabic and disyllabic adjectives:

\[(12)\]
\[
\begin{array}{ll}
\text{Present} & \text{Adverbial} \\
\text{ee} & \text{yo'o ( = ‘yo + ‘u)} \\
a'kai & \text{ako'o ( = ‘aka + ‘u)} \\
\text{(cf. kana’sli, kana’syuu ‘sad’)}
\end{array}
\]

Evidently, one of the two accents in ‘ako'o is contributed by the adjective stem and one by the ending. A consideration of compounds provides strong evidence that the low-initial allomorph of disyllabic adjectives is more basic than the high-initial allomorph: in Kansai dialects, a compound starts high or low, depending on whether its first member starts high or low\(^{17}\), and in Kyōto, almost all compounds beginning with a disyllabic adjective stem are low-initial:

\[(13)\]
\[
\begin{array}{llll}
\text{‘aka-hata} & \text{‘red flag’} & \text{‘taka-i’biki} & \text{‘loud snore’} \\
\text{‘aka-si’ngoo} & \text{‘red traffic light’} & \text{‘taka-go’e} & \text{‘shrill voice’} \\
\text{‘aka-to’nbo} & \text{‘red dragonfly’} & \text{‘taka–’hiku} & \text{‘high and low’}
\end{array}
\]

This means that the present tense form must involve attraction of accent from before the root (where it would be realized as low-initial) to the first syllable of the root (where it is realized as high pitch):

\[(14)\]
\[
\text{‘aka + i } \to \text{ a’ka-i}
\]
The adverbial form of 'good' has undergone deletion of one of two underlying accents: 'vo + 'u → yo-'u. Of course, some rule or other would have to apply to it, since modern Kansai dialects do not allow surface forms which are low-initial and have first syllable accent.

As in the standard language, verbs exhibit only a two-way accentual opposition. In Kyōto, the opposition is between high-initial verbs and low-initial verbs. For monosyllabic and disyllabic verbs, these generally correspond to the unaccented and accented verbs, respectively, of the standard language; however, longer verbs are high-initial except for derivatives of low-initial words (e.g. 'taka-sugir-u 'be too expensive,' 'yomi-ta-gar-u 'behaves like he wants to read', 'yom-are-ru 'is read'). In Kōchi, the opposition is as in the following examples:

(15) Unaccented               Accented
    uru  'sell'     soru  'shave'
    neru  'sleep'  deru  'leave'
    agaru  'rise'  u'goku  'move
    akeru  'open'  l'kiru  'live'
    utagau  'doubt'  yorokobu  'rejoice'
    kasaneru  'pile up'  kazou'reru  'count'

These facts are further evidence in favor of identifying low-initial with fall in pitch; if such an identification is made, then the placement of accent in accented verbs allows a uniform generalization, namely that the accent is put before the stem-final (C)V(C).

My subsequent remarks about Kansai verbs will be restricted to the Kyōto dialect.

The past tense in Kyōto generally involves a fall in pitch, though its location depends on the segmental and accentual form of the verb stem:

(16) Unaccented                 Accented
    Present  Past                  Present  Past
    'go'      iku  it'ta           'bite'      kamu  'kanda
    'produce' umu  u'nda           'cut'       kiru  'kitta
    'sell'    uru  ut'ta           'keep as pet'   kau  'koota
    'carry'   ou  o'ota           'write'      kaku  'kafta
    'put on'  kiru  ki'ta         'come'      kuru  ki'ta
    'do'      suru  si'ta         'emerge'    deru  de'ta
    'sleep'   neru  na'ta          'see'       miru  mi'ta
    'rise'    agaru  a'gatta       'walk'      aruku  'aru'ita
    'float'   ukabu  u'kanda       'conceal'    kakusu  'kaku'sita
    'loathe'  kirau  ki'roota     'enter'      hairu  'hai'tta
    'open'    akeru  a'keta       'waken'     okiru  'oki'ta
    'borrow'  kariru  ka'rito     'hang'       kakeru  'kake'ta
    'ridicule' azakeru  azaketa'  'conceal'    kakureru  'kaku'reta
    'inquire' ukagau  ukago'ota    'oneself'   kakureru  'kaku'reta
    'give'    ataeru  ata'eta
    'pile up' kasaneru  kasa'neta
The forms which exhibit no fall in pitch in the past are those of accented CVC verb stems. Assuming an underlying form /'ta/ for the ending, these would be derived from /kam + 'ta, etc. by means of a rule deleting the second accent from #'CVC + 'CV (I know of no more general statement of the rule which fits the facts of the dialect). The high-initial past tense of accented CV verbs evidently arises via the regressive accent deletion noted above in connection with the adverbial forms of monosyllabic adjectives. The same phenomenon is also found in the imperative and conditional:

(17) Past Imperative Conditional

'mi + 'ta → mi'ta  'mi + 'ro → mi'ro  'mi + 'tara → ml'tara
'de + 'ta → de'ta  'de + 'ro → de'ro  'de + 'tara → de'tara

Cf. 'yo + 'u → yo'o ('good' + adverbial ending)

Other forms with an underlying '0'0 eliminate it by shifting the second accent to the right:

(18) 'mat + 'reba → 'mate'ba (conditional of 'matu 'wait')
'mat + 'e → 'mate' (imperative of 'matu)
'aka + 'ke + 'reba → 'aka'kere'ba (conditional of 'akai 'red')

The regressive accent deletion of (17) and the accent shift of (18) are both involved in negative forms:

(19) 1'kahen 'does not go'  kama'hen 'does not bite'
'se'enhen 'does not do'  de'ehen 'does not leave'
'aga'kahren 'does not rise'  'aru'kahren 'does not walk'
'a'kehren 'does not open' oki'hin 'does not wake'
u'kagawahen 'does not inquire'

In unaccented verbs, there is an accent one mora before -hen (two moras before it if there are at least four moras before it). In accented verbs, however, the accent is immediately before -hen when what precedes is '00, and the low initial is lost from roots of the form '0. If the accent accompanying -hen basically goes one mora before it, the rules involved in (17)-(18) will cover these facts:

(20) 1'kam + 'a + hen → kama'hen
'de + 'e + hen → de'ehen
(alternate form: 'de + 'ya + hen → do'yahen)

The Kyōto pronunciations of the examples of compound nouns given for standard Japanese in §1.3 are as follows:
(21) a. (cf. (26) of 1.3)

'a noogyoo 'kumiai $\rightarrow$ 'noogyoo-'kumiai 'agricultural union'
roodoo 'kumiai $\rightarrow$ roodoo-'kumiai 'labor union'
hura'nsu + ryo'ori $\rightarrow$ huransu-ryo'ori 'French cooking'
tyuuka-ryo'ori 'Chinese cooking'
genzi + monoga'tari $\rightarrow$ genzi-monoga'tari 'the Tale of Genji'
'isoppu + monoga'tari $\rightarrow$ 'isoppu-monoga'tari 'Aesop's Fables'
hana-u'ri + mu'sume $\rightarrow$ hanauri-mu'sume 'girl who sells flowers'
inaka + mu'sume $\rightarrow$ inaka-mu'sume 'country girl'

b. (cf. (28) of 1.3)

'asobi + ha'nbun $\rightarrow$ 'asobi-ha'nbun 'half for pleasure'
zyo'dan + ha'nbun $\rightarrow$ zyoodan-ha'nbun 'half as a joke'

'hutai + zyooken $\rightarrow$ 'hutai-zyooken 'Incidental condition'
ka'izyo + zyooken $\rightarrow$ kaizyo-zyooken 'cancellation condition'

c. (cf. (29) of 1.3)
sansui-ga 'landscape' (sa'nsui)
pasutera-ga 'pastel' (pa'sutera)
syoozoo-ga 'portrait' (syoozoo)
huukei-ga 'landscape' (hu'ukei)
'ratai-ga 'nude' ('ratai)

huusen-dama 'balloon' (huusen)
zyuzu-dama 'rosary bead' (zyu'zu)
garasu-dama 'glass bead' (ga'rasu)
nankin-dama 'glass bead' (na'nkin)
kensetu'-syoo 'Ministry of Construction' (kensetu)
gaiimu'-syoo 'Foreign Ministry' (ga'imu)
ookura'-syoo 'National Exchequer'
'moorin'-syoo 'Ministry of Agriculture and Forestry'
'monbu'-syoo 'Ministry of Education'

'kabuto'-musi 'beetle' ('kabu(to)
'abura'-musi 'cockroach' ('abura)
nankin'-musi 'bedbug' (na'nkin)
tentoo'-musi 'bedbug'
'sanada'-musi 'tapeworm'
In all but one of the examples in (21), the compound agrees with its first member in being high-initial or low-initial. The one exception, 'gaimu'-syoo 'Foreign Ministry', is low-initial, even though its first member, ga'imu 'foreign affairs', is high-initial. Ga'imu in fact requires an underlying form 'ga'imu, since words beginning with the Sino-Japanese morpheme gai 'outside' are low-initial except when accented on the first syllable. Similarly, gi'mu 'obligation' requires an underlying form 'gi'mu, since compounds beginning with it are always low-initial (e.g. 'gi'mu-kyo'oiku 'compulsory education'). The regressive accent-deletion of (17) will delete the preposed accent of 'gi'mu and 'ga'imu when they are used as independent words.

As in the standard language, a compound with an unaccented final member gets an accent on the first mora of the final member. A final-accented long final member is not possible in Kyōto Japanese, since only 'short' words can be final-accented; accent on any other syllable is preserved in the compound. The first pair of examples in (21a) suggest that low-initialness of the final member is eliminated, and an examination of further examples shows that to be the case:

(22) kootoo + 'gakkoo →'kootoo-gak'koo 'high school'
goma + 'abura →goma-a'burə 'sesame oil'
hukuro + 'nezumi →hukuro-ne'zumi 'opossum' (lit. 'bag mouse')
   + 'kuzi'ra →makkoo-ku'zira 'sperm whale'
yoroi + 'kabu'to →yoroi-ka'buto 'armor helmet'
te'tu + 'kabu'to →te'tu-ka'nuto 'iron helmet'

Note, however, that the accent on the second syllable of 'kuzi'ra and 'kabu'to is also missing from the compound. My conjecture as to why this should be the case is that Kyōto lacks entirely nouns of the accentual type 00'0, and that once the low-initialness of 'kuzi'ra, etc. is eliminated, the resulting final member has an inadmissible accentual form and must be changed somehow, and shifting its accent one mora to the left is the minimal change which yields an admissible output compound. Alternatively, it might be the case that (22), etc., involve shift of the preposed accent of the final member one mora to the right and that the usual accent deletion rule then eliminates the other
accent. I tentatively conclude that, as in standard Japanese, the accent of a long final member of a compound predominates, with the qualification that low-initialness of the final member is eliminated, and low-initialness of the first member is not affected by the predomination of the second member.

As in the standard language, some short final members of compounds make the compound unaccented (except for the low-initialness that it inherits from its first member) and other final members yield compounds in which the final member is preceded by an accent. There are likewise short final members that take an accent on their first syllable, though none happen to have figured in (21):

(23) nankin-ma'me 'peanut' (na'nkin)
ingen-ma'me 'green bean'

I am not prepared to state any correlation between the accent of 1- and 2-mora nouns and their effect as final member of a compound other than to observe that unaccented high-initial words are preceded by an accent when used as the final element of a compound, e.g.

(24) yuubin + hako \(\rightarrow\) yuubin'-bako 'mailbox'
syussatu + kuti \(\rightarrow\) syussatu'-guti 'gate where tickets are punched'
koori + mizu \(\rightarrow\) koori'-mizu 'ice water'
hurumai + sake \(\rightarrow\) hurumai'-zake 'treat of wine'

Whatever the correlation is, it will not be possible to set up an analysis as neat as that which I proposed for the standard language, since it is not possible to take the accent contributed by an unaccented final member as part of the underlying form of the final member: such an analysis was possible in the standard language only because there was no contrast between 'unaccented' and 'pre-accented' in the standard language.

One striking feature of the compounds in which the first member ends in a long syllable is that the accent imposed by final members such as -syoo and -musi goes on the second mora of the long syllable. While simple words in Kyōto never have a contrast between accented first mora and accented second mora, such a contrast is possible in compounds, since the accent inserted into a long second member goes onto its first mora, whereas the accent put in by a 'pre-accenting' final member goes on the immediately preceding mora:

(25) si'nkee + syoo \(\rightarrow\) sinkee'syoo 'neurosis'
sin + se'ehu \(\rightarrow\) sinse'ehu 'new government'

2.2 Western Kyūshū and the Ryūkyū Islands

In western Kyūshū and parts of the Amami and Ryūkyū Islands, not only verbs
and adjectives but also nouns exhibit only a two-way accentual opposition, regardless of length. The details as to which moras or syllables are pronounced high and which ones low varies considerably from one locality to another\textsuperscript{32}. The simplest variant of this type of accentual system is found in Kuma (Fujitsu-gun, Saga pref.). Using the terms 'falling' and 'level' for the two accentual types, a phrase beginning with a falling morpheme has its first mora high pitched and all subsequent moras low pitched; a phrase beginning with a level morpheme is on a low level pitch:

\begin{enumerate}
\item a. Falling nouns
\begin{itemize}
\item \textit{hi}, \textit{hi} ga 'day'
\item \textit{hana}, \textit{hana} ga 'nose'
\item \textit{kuruma}, \textit{kuruma} ga 'vehicle'
\item \textit{kamaboko}, \textit{kamaboko} ga 'fish pudding'
\end{itemize}
\item b. Level nouns
\begin{itemize}
\item \textit{hi}, \textit{hi} ga 'fire'
\item \textit{hana}, \textit{hana} ga 'flower'
\item \textit{abura}, \textit{abura} ga 'oil'
\item \textit{yomikata}, \textit{yomikata} ga 'pronunciation'
\end{itemize}
\end{enumerate}

Other localities in Western Kyūshū have accentual systems that presumably developed from this system via spread or shift of the high pitch to the right, e.g.

\begin{enumerate}
\item Higashi-Kawanobori (Saga) Nagasaki Kagoshima
\begin{itemize}
\item \textit{hi}, \textit{hi} ga
\item \textit{hana}, \textit{hana} ga
\item \textit{kuruma}, \textit{kuruma} ga
\item \textit{kamaboko}, \textit{kamaboko} ga
\end{itemize}
\item Nagasaki
\begin{itemize}
\item \textit{hi}, \textit{hi} ga
\item \textit{hana}, \textit{hana} ga
\item \textit{kuruma}, \textit{kuruma} ga
\item \textit{kamaboko}, \textit{kamaboko} ga
\end{itemize}
\item Kagoshima
\begin{itemize}
\item \textit{hi}, \textit{hi} ga
\item \textit{hana}, \textit{hana} ga
\item \textit{kuruma}, \textit{kuruma} ga
\item \textit{kamaboko}, \textit{kamaboko} ga
\end{itemize}
\end{enumerate}

In Nagasaki, the high pitch has shifted to the second mora (syllable?) of the phrase, provided that the phrase has at least three moras (syllables?). In Kagoshima, the accent has shifted to the next-to-last syllable of the phrase. In Koshikijima (Kagoshima pref.), it has shifted to the next-to-last mora:

\begin{enumerate}
\item Kagoshima (syllable counting) Koshikijima (mora counting)
\begin{itemize}
\item \textit{meirei} 'order'
\item \textit{hootai} 'bandage'
\item \textit{kyuusyu}u 'Kyushu'
\item \textit{kankel} 'relationship'
\end{itemize}
\item Koshikijima (mora counting)
\begin{itemize}
\item \textit{meirei}
\item \textit{hootai}
\item \textit{kyuusyu}u
\item \textit{kankel}
\end{itemize}
\end{enumerate}

The references to 'next-to-last syllable/mora' must be qualified to the extent of noting that certain morphemes do not count as part of the phrase for the purposes of accent shift. For example, when the above falling nouns are combined with the copula \textit{da} in Kagoshima, the result is:
There thus is a surface distinction between phrases with high-pitched penultimate syllable and phrases with high-pitched ante-penultimate syllable.

Except for one locality, the only variation in the pronunciation of phrases beginning with a level morpheme is whether the phrase is pronounced on a low level pitch or has a rise at the end. In Kagoshima, there is a rise, and it does not affect items such as the copula which do not 'count' accentually; thus a phrase ending . . . LHL can be either a falling phrase that ends with material that 'counts' accentually (e.g. *hana ga* 'nose (Nom.)') or a rising phrase that ends with a one-syllable morpheme that 'does not count' (e.g. *hana da* 'it's a flower'). The exceptional dialect is Makurazaki (Kagoshima pref.), which has undergone a pitch inversion, as is clear from a comparison with Kagoshima forms (data from Tojō, vol. 4, pp. 277-8), and thus has a fall in pitch in 'level' phrases:

(5)  

<table>
<thead>
<tr>
<th>Falling words</th>
<th>Level words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kagoshima</td>
<td>Makurazaki</td>
</tr>
<tr>
<td>ひ, ひ が 'day'</td>
<td>ひ, ひ か</td>
</tr>
<tr>
<td>はな, はな が 'nose'</td>
<td>はな, はな か</td>
</tr>
<tr>
<td>あそぶ 'play'</td>
<td>あそぶ</td>
</tr>
</tbody>
</table>

The only Makurazaki forms which are not exact tonal inversions of the corresponding Kagoshima forms are the isolation forms of monosyllables. The falling pitch of Kagoshima 'day' and Makurazaki 'fire' amounts to compression on a single syllable of the melody that is manifested when か follows.

In western Kyūshū it is the first member of a compound that determines the accent on the compound, e.g. (Kagoshima forms, taken from Hirayama 1960):

(6)  

| は 'leaf' + さくら 'cherry tree' | はざくら 'post-blossom foliage' |
| "  'stone' + か 'fence' | "  'stone fence' |
| はやま 'mountain' + さくら 'cherry tree' | はやざくら 'wild cherry tree' |
| いろ 'color' + しろ 'white' | いぞろ 'fair-skinned person' |

Most of the dialects of the Amami Islands and the Ryūkyū Islands also exhibit
only a two-way accent contrast, and there is the same kind of variation from one loca-
ty to another as to the details of the contrast, e.g. in Miyako-Ōura, just as in Nagas-
aki, the second mora of 'falling' phrases is high, except when there are only two mo-
as, in which case the first is high; in Miyako-Karimata, just as in Koshikijima, the
enultimate mora of a falling phrase is high-pitched; in Miyako-Shimoji, the second
mora of a falling phrase is high-pitched even if there are only two moras23.

However, there are dialects in this area which make more than a two-way ac-
centual distinction, for example, the dialects of Kametsu (Amami Tokunoshima) and
ronaguni (the group of islands near Taiwan), in which a 3-way distinction among 2-syl-
lablable nouns is maintained.

1. History

The principal correspondences for 2-syllable nouns between Kansai Japanese and
Tokyo Japanese (which for present purposes can be identified with the standard language)
are illustrated by the four words

<table>
<thead>
<tr>
<th></th>
<th>Kansai</th>
<th>Tokyo</th>
</tr>
</thead>
<tbody>
<tr>
<td>'cow'</td>
<td>usi</td>
<td>usi</td>
</tr>
<tr>
<td>'bridge'</td>
<td>ha'si</td>
<td>hasi'</td>
</tr>
<tr>
<td>'chopsticks'</td>
<td>'hasi</td>
<td>ha'si</td>
</tr>
<tr>
<td>'window'</td>
<td>'mado'</td>
<td>ma'do</td>
</tr>
</tbody>
</table>

Since there is no segmental characteristic from which the difference between the 'chop-
sticks' type and the 'window' type is predictable, the parent accentual system must have
had an accentual distinction between the two which was lost in the development of Tokyo
Japanese. The data in (1) are in fact consistent with the hypothesis that the parent system
is identical to the modern Kansai system and that the Tokyo system developed by a shift
of accent one mora to the right.

A consideration of additional dialect evidence makes it necessary to set up an ad-
ditional accentual type in the parent language. The 'bridge' type divides into two distinct
sets of reflexes in western Kyūshū, eastern Kyūshū (e.g. Ōita), and in northern Honshū
e.g. Akita):

<table>
<thead>
<tr>
<th></th>
<th>Kansai</th>
<th>Tokyo</th>
<th>Akita/Oita</th>
<th>W. Ky.</th>
<th>Reconstruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. usi 'cow'</td>
<td>00</td>
<td>00</td>
<td>00</td>
<td>F</td>
<td>00</td>
</tr>
<tr>
<td>2. hasi 'bridge'</td>
<td>0'0</td>
<td>00'</td>
<td>00</td>
<td>F</td>
<td>0'0</td>
</tr>
<tr>
<td>3. hana 'flower'</td>
<td>0'0</td>
<td>00'</td>
<td>00'</td>
<td>L</td>
<td>'0'0</td>
</tr>
<tr>
<td>4. hasi 'chopsticks'</td>
<td>'00</td>
<td>0'0</td>
<td>0'0</td>
<td>L</td>
<td>'00</td>
</tr>
<tr>
<td>5. mado 'window'</td>
<td>'00'</td>
<td>0'0</td>
<td>0'0</td>
<td>L</td>
<td>'00'</td>
</tr>
</tbody>
</table>
In sets 1, 2, 4, and 5, Kansai high-initial corresponds to W. Ky. F (= 'falling') and Kansai low-initial corresponds to W. Ky. L (= 'level'). Set 3, which is not distinct from set 2 in Kansai and Tōkyō, deviates from this correspondence by virtue of having L instead of F in W. Ky. The obvious difference to posit between set 2 and set 3 in the proto system is thus that of high-initial (set 2) vs. low-initial (set 3). Northern Honshū and eastern Kyūshū largely agree with Tōkyō in accent but have unaccented rather than final-accented reflexes for set 2. The agreement between northern Honshū and eastern Kyūshū in distinguishing set 2 (about 20 nouns) from set 3 (about 40 nouns) is striking, since the two areas are separated by a 1000 km stretch of Honshū in which they are not distinguished.

It is not obvious how to interpret the accentual formula reconstructed for set 3. How would a form like /'ha'na/' be pronounced? A low pitched syllable followed by a still lower one? A rising pitched syllable followed by a low? Or perhaps the first ' would make the first syllable low and the other one would make subsequent syllables low, thus making the word level low? The last possibility seems the best bet, since it makes the reconstruction conform quite well to the pitches recorded in the Ruijumyōgishō (a thesaurus compiled in Kyōto about 1100 A.D.):

(3)  

<table>
<thead>
<tr>
<th></th>
<th>Reconstruction</th>
<th>Ruijumyōgishō</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>00</td>
<td>HH(H)</td>
</tr>
<tr>
<td>2.</td>
<td>'0'0</td>
<td>HL(L)</td>
</tr>
<tr>
<td>3.</td>
<td>'0'0</td>
<td>LL(H)</td>
</tr>
<tr>
<td>4.</td>
<td>'00'</td>
<td>LH(H)</td>
</tr>
<tr>
<td>5.</td>
<td>'00'</td>
<td>LH(L)</td>
</tr>
</tbody>
</table>

(the pitches given in parentheses are those of a following wa or ga). If this reconstruction is correct, then the following historical developments have occurred:

i. (W. Kyūshū) Syllables past the first become L (thus, only the pitch on the first syllable remains distinctive).

ii. (E. Kyūshū [Ōita] and N. Honshū [Akita]) Accent is lost in high-initial words.

iii. (All Japan)26 '0'0 → 0'0. 

iv. (E. Honshū [incl. Tōkyō, Akita], W. Honshū [Hiroshima], E. Kyūshū [Ōita]). Accent shifts one mora to the right.

The accent shifts in Eastern and Western Japan are presumably independent developments. Accent shift in fact appears to be a rather common phenomenon in the history of Japanese dialects. The data cited in Hirayama 1966b suggest that the same accent shift took place independently in most of Amami-Tokunoshima, where 2-syllable nouns exhibit the same pitches as in Akita except that unaccented phrases (sets 1 and 2) have the shape LLL rather than the level low of Akita. Okuda (1975) argues that the same accent shift has occurred again in Izumo, which had undergone it once already with the
The correspondences for 3-syllable nouns are much less neat than those for 2-syllable nouns. In the following table, I have listed the 7 distinct pitch forms recorded in the Ruijumyōgishō, with the most common manifestations of those types in the modern dialects. I have used parentheses to mark the many cases where, because there are so few representatives of a particular type and their modern reflexes are so diverse, it is not completely clear what modern reflex is the 'regular correspondence'. Kansai dialects are represented here by Befu (Hyogo pref.), which is more conservative than either Kyoto or Osaka.

<table>
<thead>
<tr>
<th>No.</th>
<th>Ruijumyōgishō</th>
<th>Kansai</th>
<th>Tōkyō</th>
<th>Akita</th>
<th>Ōita</th>
<th>W. Kyūshū</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>sakana 'fish'</td>
<td>HHH(H)</td>
<td>000</td>
<td>000</td>
<td>000</td>
<td>F</td>
</tr>
<tr>
<td>2.</td>
<td>tokage 'lizard'</td>
<td>HHL(L)</td>
<td>('00'0)</td>
<td>(000)</td>
<td>(000)</td>
<td>F</td>
</tr>
<tr>
<td>3.</td>
<td>awabi 'abalone'</td>
<td>HLL(L)</td>
<td>0'00</td>
<td>(0'00</td>
<td>(00'0)</td>
<td>(000')</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(00'0)27</td>
</tr>
<tr>
<td>4.</td>
<td>atama 'head'</td>
<td>LLL(H)</td>
<td>00'0</td>
<td>000'</td>
<td>000'</td>
<td>L</td>
</tr>
<tr>
<td>5.</td>
<td>namida 'tear'</td>
<td>LLH(L)</td>
<td>0'00</td>
<td>(0'00</td>
<td>(00'0)</td>
<td>(000')</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>&gt;00'0</td>
</tr>
<tr>
<td>6.</td>
<td>nezumi 'mouse'</td>
<td>LHH(H)</td>
<td>'000'</td>
<td>000</td>
<td>00'0</td>
<td>00'0</td>
</tr>
<tr>
<td>7.</td>
<td>kabuto 'helmet'</td>
<td>LHL(L)</td>
<td>'00'0</td>
<td>0'00/000</td>
<td>(00'0)</td>
<td>(0'00)</td>
</tr>
</tbody>
</table>

The correspondences between western Kyūshū and Ruijumyōgishō work out perfectly, but nothing else does. Moreover, while representations in terms of accent marks suggest themselves for five of the seven accentual types of the Ruijumyōgishō, there is no obvious assignment of accentual formulas to types 4 and 5. One of them would have to be represented as '0'00 and the other as '000', but which would be which? The fact that both type 4 and 2-syllable type 3 are all low followed by a high particle suggests assigning to 4 the formula '0'00. However, that assignment gives no clue as to why the reflexes of type 4 are what they are in Kansai and Tōkyō (if 'head' and 'flower' were parallel, Kansai should have あたま rather than あたま), though neither would the reverse assignment of the two formulas.

Hayata (1973) argues that proto-Japanese had more accentual distinctions than were recorded in the Ruijumyōgishō. He points out that by reconstructing accentual shapes 00' and 000', which are unattested in the Ruijumyōgishō as in most modern Kansai dialects, one can account for some otherwise exceptional correspondences. Specifically, since such accentual types, if subjected to the accent shift (rule iv, above), would yield unaccented words in Tōkyō and Hiroshima, they would give rise to correspondences such as those in (5a) and account for the distinction between them and those in (5b):
Hayata proposes that the Kansai reflexes of the words in (5a) underwent an accent retraction, which applied to all final accents except those where there is not 'enough room' for it to take place, as in na' 'name' and mado' 'window'.

There is an additional class of words to which Hayata's proposal is applicable. In Tōkyō, 3-syllable nouns of type 7 divide evenly as regards whether they are unaccented or initial accented:

(6)

<table>
<thead>
<tr>
<th>Kansai</th>
<th>Tōkyō</th>
<th>W. Ky.</th>
<th>Proto Jap.</th>
</tr>
</thead>
<tbody>
<tr>
<td>'person'</td>
<td>hito</td>
<td>F</td>
<td>00'</td>
</tr>
<tr>
<td>'lizard'</td>
<td>tokage</td>
<td>F</td>
<td>000'</td>
</tr>
<tr>
<td>'bridge'</td>
<td>hasi'</td>
<td>F</td>
<td>0'0</td>
</tr>
<tr>
<td>'red bean'</td>
<td>azuki'</td>
<td>F</td>
<td>00'0</td>
</tr>
</tbody>
</table>

Suppose that the words in the first column were assigned a proto form '00'0 and those in the second column a proto form '000'. The hypothesized Kansai accent retraction would obliterate the difference between the two types in Kansai dialects. In both Tōkyō and Hiroshima, type 6 nouns (with presumable proto form '000) are unaccented. While it is not clear what steps are involved in getting from proto '000 to unaccented reflexes in Tōkyō and Hiroshima, the same steps would presumably yield unaccented reflexes for the hypothetical '000', since accent shift would obliterate the difference between it and '000.

Hayata's approach leads to a hypothetical proto-language in which there were 6 accentual possibilities for 2-syllable nouns and 9 for 3-syllable nouns. In addition, Hattori (1951) proposed reconstructing as LLH(H) those words of type 5 (such as abura 'oil') which are unaccented in Eastern and Western Honshu, to distinguish them from those (such as nami'da 'tear', na'mida in Tōkyō) which are accented:

(7)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>HH(H)</td>
<td>1.</td>
</tr>
<tr>
<td>2a.</td>
<td>HH(L)</td>
<td>2a.</td>
</tr>
<tr>
<td>2b.</td>
<td>HL(L)</td>
<td>2b.</td>
</tr>
<tr>
<td>3.</td>
<td>LL(H)</td>
<td>3.</td>
</tr>
<tr>
<td>4.</td>
<td>LH(H)</td>
<td>4.</td>
</tr>
<tr>
<td>5.</td>
<td>LH(L)</td>
<td>5a.</td>
</tr>
<tr>
<td>5b.</td>
<td>LLH(H)</td>
<td></td>
</tr>
</tbody>
</table>
6. LHH(H)
7a. LHH(L)
7b. LHL(L)

Note, however, that if these reconstructions are correct, it is impossible to represent accentual distinctions in proto-Japanese in terms of high- vs. low-initial plus place of fall in pitch: there are 10 distinct accentual types for 3-syllable nouns, whereas there are only 8 (= 2 x 4) possible combinations of high vs. low-initial plus accent on first, second, third, or no syllable.

An alternative mode of representing accentual distinctions in proto-Japanese is suggested by Okuda (1975), namely in terms of place of rise (written \( \uparrow \)) and place of fall (written \( \downarrow \)), where both are in principle distinctive. Suppose, to modify Okuda's notation slightly, high-initial words are written with a preceding 'rise', so that every word will have a rise and somewhere after the rise there may (but need not) be a fall. The above accentual types would then be represented as

\[
\begin{array}{ll}
1. & \uparrow 00 \\
2a. & \uparrow 00 \\
2b. & \uparrow 0 0 \\
3. & 0 0 \uparrow \\
4. & 0 \uparrow 0 \\
5a. & 0 0 \uparrow 0 \\
5b. & 0 0 \uparrow 0 \\
6. & 0 \uparrow 0 \\
7a. & 0 r 0 0 \\
7b. & 0 r 0 0 \\
\end{array}
\]

This would make complete use of the combinatory possibilities for 'rise' and 'fall': if every word must have a rise and if a fall can only follow a rise, then there are 6 possibilities for 2-syllable nouns and 10 possibilities for 3-syllable nouns.

4. Typology
4.1 Syllables and moras

Both the syllable and the mora play a role in the structure of standard Japanese. A distinction must be drawn between long syllables, which have the shape (C)VV or (C)V(C)VC (with the final consonant being either the first part of a geminate or a nasal whose exact articulation is determined by the surrounding segments), and short syllables, which have the shape (C)V. Universally, the distinction between 'long syllable' and 'short syllable' amounts to the distinction between 'syllable that counts as two units' and 'syllable that counts as one unit'. There are two distinct ways in which a syllable might count as two units: (a) with respect to rules that are sensitive to the 'length' of an item or to the 'distance' between two points in an utterance,
and (b) with respect to whether it provides two possible locations or only one for accentual phenomena.

The following phenomena in standard Japanese are sensitive to 'length' or 'distance'. (i) Compound nouns with a long final member are subject to different rules than compounds with a short final member; something three or more moras long counts as 'long' and something less than three moras long counts as 'short', except that compounds count as long even if they are only two moras long. (ii) Recent loan words are accentuated either where the stress in the source language was or on the third mora from the end. The following examples show that it is a mora count rather than a syllable count that determines the location of accent when it diverges from the stress of the source language:

(1) to'rio 'trio' eroti'kku 'erotic' kuude'taa 'coup d'etat'
bura'ziru 'Brazil' reko'odo 'record'
ra'peru 'lapel' paamane'nto 'permanent' (wave)
sokura'tesu 'Socrates'

In the items in the first column, the last two syllables are short, and the accent goes on the preceding syllable. In the items in the second column, the final syllable is short and the preceding syllable long, and it is on that syllable that the accent goes. In the items in the third column, the final syllable is long and the accent goes on the preceding short syllable. A fourth possibility remains, namely that in which the third mora from the end is the second mora of a long syllable. Here the accent is four moras from the end:

(2) esukare'etaa 'escalator'
wasi'nton 'Washington'
rechure'kkusu 'reflex'

However, a single generalization still covers both the examples in (1) and those in (2): the accent is on the syllable containing the third-from-last mora. (iii) Japanese poetry is written in meters that are sensitive to the number of moras in a line but not to how those moras group together into syllables. The traditional meters involve lines alternately of 5 and of 7 moras.

Standard Japanese is thus a 'mora-counting language' rather than a 'syllable-counting language'. However, it is syllables rather than moras that bear the accent, and standard Japanese can thus be described as a 'mora-counting syllable language'. (i) There is no opposition between accented first mora and accented second mora:
each syllable provides only one possible place where pitch may drop. By contrast, Lithuanian allows a contrast between stressed first mora and stressed second mora (e.g. ėnis 'duck' vs. ėnis 'breast') in long syllables, though there is no such contrast in short syllables. In Lithuanian, the number of possible places that stress might fall in a syllable thus equals the number of moras in the syllable. (ii) Accent placement rules in standard Japanese give outputs in which pitch drops after the first mora of a long syllable, even if it is the second mora rather than the first that meets the conditions for application of the rule. For example, the accent that appears before 'preaccenting' final members of compounds goes on the first mora of the preceding syllable, e.g. tento'o-musi 'ladybug' (cf. kabuto'-mubi 'beetle'). Similarly with the accent that precedes certain particles:

(3) oosaka 'Ōsaka', oosaka' sika 'only Ōsaka'
tookyoo 'Tōkyō', tokyo'o sika 'only Tōkyō'
(c.f. na'goya, na'goya sika)

And, as noted in the preceding paragraph, the accent inserted on the third mora from the end of a loan word appears on the fourth mora from the end if the third mora from the end is the second mora of a long syllable. (iii) Rules which are sensitive to whether something is 'final accented' treat items with an accented long final syllable and items with an accented short final syllable the same way, even though the accent in the former is not on the final mora. This can be illustrated by the rule for compound nouns with a final-accented long final member (§1. 3, exx. (26), (28) and by the rule that removes final accent before the genitive marker no:

(4) uma' ga, uma no 'horse' (nom., gen.)
taiwa'n ga, taiwan no 'Taiwan'
(c.f. ne'ko ga, ne'ko no 'cat')

Long syllables have developed within the last 1500 years in Japanese, by loss of consonants or vowels, and in borrowing from Chinese. Before large-scale borrowing from Chinese began, in the 6th and 7th centuries A. D., all Japanese syllables had the form (C)V, and syllables without an initial consonant occurred only word-initially. The accentual system of proto-Japanese thus made no provision for long-syllables, and there is no reason to expect that all Japanese dialects would treat long syllables the same way. I have noted above that Japanese dialects in fact differ as regards whether the syllable or the mora is the unit of length or is the unit to which accents are assigned (see §2.2, ex. (3)). Kyōto Japanese provides a counterexample to my earlier claim (McCawley 1968) that a language is a 'syllable language', a 'syllable counting language', etc. throughout the entire derivation (i.e. that if one rule in a language measures length in syllables, then all rules will take the syllable as the unit of distance; if one rule in a language treats the syllable rather than the mora as the bearer of accent, then all rules of the language will),
Kyōto Japanese changes from a syllable language to a mora language in the course of a derivation: in underlying forms of lexical items, there is no contrast between long syllable with first-mora accent and long syllable with second-mora accent, but such a contrast is created by the rules for noun compounds (see §3.1, ex. (25)).

4.2. Pitch accent vs. tone system vs. ?

A phonological characteristic (be it high pitch, vowel length, or whatever) is used accentually in a language when what is distinctive is its location; underlying forms contain at most a specification of the location of this characteristic (rather than specifying independently for each syllable whether it has the characteristic or not), and rules inserting that characteristic in one location have the side effect of eliminating or 'weakening' any other occurrences of it that might be present, even at a distance of several syllables.

Strictly speaking, accent is a phonological notion but not a phonetic notion. 'Accent rules' operate in terms of abstract accent marks, and eventually rules must apply which replace accent marks by the phonetic characteristics in which they are manifested. The number of degrees of stress which a language distinguishes depends on the number of phonetic characteristics which figure in the realization of 'accent': only when more than one characteristic is involved is it possible for the language to distinguish 'reduced stress' from both 'full stress' and 'unstressed'. A pitch accent system is an accentual system in which the 'accent' is manifested in terms of the pitch on the accented syllable or mora, relative to the pitches on the other syllables or moras. The conversion of accent marks into pitches need not be the last step of the derivation, however, and on this ground I have argued (McCawley 1970a) against the popular conception of 'pitch accent system' and 'tone system' as opposite terms of a dichotomy. I hold rather that there is a continuum with regard to how 'deeply tonal' a language is, and there are languages (such as Ganda and Tonga) which are pitch accent systems as far as their underlying forms are concerned but whose phonological rules operate almost entirely in terms of pitches on specific syllables rather than the location of 'accent'.

There in fact appears to be at least one rule of Japanese phonology that applies to representations in which pitches are represented not just as 'high' and 'low', let alone as location of 'accent', but in terms of concrete pitch levels, namely the contraction found in itte'ru - itte ru 'is going', etc. Note that while the full form is unaccented, the contracted form has an accent in the place where the vowel was lost. I conjecture that this is not a true accent: rather, the non-distinctive drop in pitch which in the full form is spread over two syllables is compressed onto one syllable, creating an illusion of accent. Weltzman (1969) presents measurements of rises and falls in pitch in phrases of one to four short syllables. His measurements
show that, except for the rise in pitch from an unaccented first syllable to an unaccented second syllable, there is a drop in pitch between adjacent unaccented syllables, though generally a smaller one than between an accented syllable and the following unaccented syllable. The relevant data are as follows:

<table>
<thead>
<tr>
<th>Informant</th>
<th>Average Drop in Pitch after Accented Syllable</th>
<th>Average Drop in Pitch from 2nd to 4th Syllable in Unaccented Phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.5</td>
<td>4.7 semitones</td>
</tr>
<tr>
<td>2</td>
<td>3.7</td>
<td>2.1</td>
</tr>
<tr>
<td>3</td>
<td>8.1</td>
<td>5.9</td>
</tr>
<tr>
<td>4</td>
<td>5.9</td>
<td>3.7</td>
</tr>
</tbody>
</table>

If the contraction of itte iru left the same pitch drop between te and ru as occurs in the full form, it would be close enough to the drop in pitch heard after an accented syllable for it to be perceived as an accent.

The dialects of eastern Honshū, western Honshū, and eastern Kyūshū are textbook cases of a pitch-accent system. Kansai dialects involve a type of pitch accent system, since they are subject to rules inserting accent (=fall in pitch) with concomitant removal of other accents present; however, the 'preposed' accent (realized as low-initial) does not trigger nor undergo such elimination of accent. Kansal dialects thus involve a pitch accent system with an additional distinction that is realized in terms of pitch. The dialects of western Kyūshū do not appear to be pitch accent systems at all. At no stage of derivations in western Kyūshū need pitch information ever be expressed in terms of accent marks: lexical representation of pitch is in the form of a feature [± Falling] attached to the whole lexical item, and the rule associating surface pitches with underlying forms can associate them directly with surface pitches, without any intermediate stage of representation in which 'accent' is marked on specific syllables. This should be contrasted with the accentual behavior of standard Japanese verbs and adjectives. While the lexical information about pitch for standard verbs and adjectives is also in the form of a binary feature of the whole lexical item, it is necessary to convert that lexical feature into an accent mark on a specific syllable, since accented verbs and adjectives trigger the same deletion of a subsequent accent as do accented nouns:

(1) \( \text{tabe}^{[\text{+acc}]} \rightarrow \text{ta'be} + \text{ta'ra} \rightarrow \text{ta'be-tara}. \)

Facts about diglossia can be cited as evidence that at least one western Kyūshū dialect does not operate in terms of 'accents' on specific syllables. Kagoshima dialect forms of verbs generally have fewer syllables than do the corresponding standard forms, and the present tense of adjectives generally has one more syllable. Kagoshima speakers
pronounce standard forms with the same 'falling' or 'level' melody as the corresponding dialect forms, even though the high pitch is then generally on a different syllable:

(2) Kagoshima pronunciation of standard form

<table>
<thead>
<tr>
<th>Standard Form</th>
<th>Corresponding Dialect Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>agaru 'rise'</td>
<td>agat</td>
</tr>
<tr>
<td>kanasimu 'grieve'</td>
<td>kanasim</td>
</tr>
<tr>
<td>akai 'red' (Present)</td>
<td>akaka</td>
</tr>
<tr>
<td>akaku 'red' (Adverbial)</td>
<td>ako</td>
</tr>
</tbody>
</table>

If the reconstruction sketched at the end of §3 is correct, then proto Japanese is one degree further removed from a true pitch accent system than are modern Kansai dialects: not only is the place of fall in pitch distinctive, as well as whether the word starts high or low, but the place where a rise in pitch occurs is also distinctive. A system in which place of rise and place of fall are distinctive but nothing else is is what would develop if a system in which high and low were free in distribution underwent a change whereby L's surrounded by H's were eliminated, e.g. were assimilated up to the level of the H's. It is thus plausible to conjecture that at some time in the remote past, Japanese had a true tonal system involving high and low tones that were relatively unrestricted in distribution and that it has evolved by steps into the true pitch accent system found today in Tōkyō and much of the rest of Japan.

1.3. Miscellaneous typological points

Pitch accent in Kagoshima requires pitch to be treated 'suprasegmentally' in the sense that certain pitches are obligatorily present regardless of the number of syllables in the phrase affected (cf. McCawley 1970b, Leben 1973). The pitches assigned to a falling phrase are given by the formula L₀HL, i.e. there are obligatorily a high pitch and a low pitch at the end, and low pitches are assigned to whatever syllables are not used up in accommodating the obligatory high and low. The obligatory high and low go respectively on the penultimate and final syllables if the phrase has at least two syllables and are compressed into a falling pitch if there is only one syllable in the phrase.

Running down the middle of Kyūshū there is an area in which accentual contrasts have been lost. A comparison of the accentual systems on the two sides of this area suggests an explanation of why such an area should exist and provides confirmation of a conjecture made by Valdis Zeps regarding a similar situation involving Latvian dialects. The five accentual types of 2-syllable nouns recorded in the Ruijumyōgishō are manifested as follows in Kyūshū:
Under the assumption that 'falling' is the marked accentual type in western Kyūshū, eastern Kyūshū 2-syllable nouns never agree with their western Kyūshū counterparts in markedness: in types 1 and 2, eastern unmarked accent corresponds to western marked accent, and in types 3-5, eastern marked accentuations correspond to western unmarked accent. Zeps' paper dealt with three mutually contiguous dialect areas in Latvia: area A, in which both the glottalization opposition and the rising/falling opposition in stressed syllables were preserved, area B, in which the glottalization opposition had been lost, and area C, in which the rising/falling opposition had been lost. Around the border between B and C, a 'buffer zone' had developed, in which both glottalization and rising pitch occurred but were used non-distinctively. However, no buffer zone appears around the border between A and B or between A and C. Zeps conjectured that this was because the B-C border was the only one for which marked pronunciations on both sides of the border corresponded to unmarked pronunciations on the other side of the border:

The eastern/western Kyūshū border is an even more spectacular case of the kind of correspondence to which Zeps attributes the development of a buffer zone: practically no words have the same markedness on both sides of the border. Only in one small area in Shikoku is there an accentless area separating Kansai accentual systems from 'Tōkyō-type' accentual systems. The absence of a buffer zone around most of the border between Kansai and 'Tōkyō-type' accentual systems conforms to Zeps' conjecture: while the correspondences between Kansai and Tōkyō-type systems are quite involved, the correspondences are mainly of marked accentuation to marked and unmarked to unmarked.

5. Annotated bibliography.

Listed here are both items referred to in the paper and other works of importance in the study of Japanese accent. Annotations have been added where appropriate, to call to the reader's attention important features of the various works.
Akinaga, Kazue. 1968. Akusento syuutoku hoosoku (Rules for learning [standard] accent). Appendix to H. Kindaichi (editor), Meikai Nihongo Akusento Ziten (Japanese accent dictionary), Tōkyō: Sanseido. A 68-page summary of accentual paradigms, rules for compounds, etc., covering many topics not otherwise treated in the literature (e.g. the accentuation of 'classical' forms in modern Japanese). A particularly valuable feature of the dictionary is that each entry indicates what rules the word illustrates.


Hattori, Shirō. 1960. 'Bunsetu' to akusento (Phrase and accent). In S. Hattori, Gengagaku no hōhō (Methods of linguistics), Tōkyō: Iwanami, 428-446. Highly insightful analyses of the Kameyama (Mie pref.) dialect and standard Japanese, with attention to the interaction between emphasis and accent and some comparison with English stress.


Hirayama, Teruo. 1966b, 1967. *Ryūkyū hōgen no sōgōteki kenkyū* (Compendium of research on Ryukyu dialects) and *Ryūkyū sentō syōhōgen no sōgōteki kenkyū* (Compendium of research on the dialects of the outer Ryūkyū Islands). Tōkyō: Meiji-shoin. Includes discs with samples of the principal dialects.


McCawley, James D. 1968. *The phonological component of a grammar of Japanese*. The Hague: Mouton. Most of its material on accent has been incorporated into this paper.
McCawley, James D. 1970a. Some tonal systems that come close to being pitch
accent systems but don't quite make it. *Papers from the sixth regional
meeting.* Chicago Linguistic Society, 526-532.

Linguistics* 1.123-130.

dialect materials), 9 volumes. All regions of Japan are covered; many of the
dialect sketches contain little information about accent, but discs are included
containing excerpts from most of the texts.

Covers several interesting dialects (Narada, Matsue, Izumo, Toyama), with an
especially detailed treatment of Hiroshima.

Illinois dissertation.


Tokugawa, Munemasa. 1962. *Nihon syohōgen akusento no keifu siron* (Towards a
family tree for accent in Japanese dialects). *Kokugo-kokubun gakkaishi*
(Gakushuin University, Tōkyō) no. 6. English translation by James D. McCawley
family tree, based on coalescences among the 5 accentual types of 2-syllable
nouns distinguished in the Ruijumyōgishō.

Wada, Minoru. 1962. *Akusento* (accent). In *Hōgengaku-gaisetu* (Introduction to

instrumental study, giving fine details of pitch.

**FOOTNOTES**

*I am grateful to Samuel E. Martin and Teruhiro Hayata for valuable comments on
preliminary versions of this paper, and to Noriko A. McCawley for data on Kyōto
Japanese.

1 the overline indicates high pitch; unmarked syllables are low pitched. Rising
and falling pitch will be indicated by '/' and '\'. Except when I wish to draw attention
to details of the pronunciation, I will cite examples in terms of fall in pitch (written ')
rather than in terms of high and low pitches.
However, two consecutive vowels need not be in the same syllable; for example, e 'to' is always a separate syllable, even in a combination such as ude e 'to the arm', in which identical vowels abut. In what follows I will assume that I know where the syllable boundaries are, though I do not guarantee that I could justify the assumed syllabifications without circularity.

While -ra has nothing in common syntactically with /ni'/, /kara'/, etc., it behaves accentually like a 'particle'.

Morpheme boundary is marked by -. The underlying forms of verb endings are the same whether added to a consonant-stem verb or a vowel-stem verb. Endings and derivational affixes lose an initial consonant other than /t/ when a stem-final consonant immediately precedes: susum-ru → susum-u; susum-re'ba → susum-e'ba.

The examples in (16) may be pronounced with more vowels devoiced than is indicated here. A final short high vowel after a voiceless consonant is frequently devoiced; in particular, the adverbial ending -ku is frequently pronounced with a voiceless vowel. In cases where two or more consecutive syllables have vowels that meet the conditions for devoicing, they need not all be devoiced. See Han (1962) for data showing that (depending on whether the vowel is /i/ or /u/ and on what the surrounding consonants are) some syllables tend more strongly than others to devoice. When several consecutive syllables are potentially subject to devoicing, generally only alternate syllables actually devoice, with Han's hierarchy determining whether the even-numbered or the odd-numbered syllables devoice. However, in rapid speech, devoicing in adjacent syllables is quite common; indeed, the all-purpose swear-word /tikusyoo/ is generally pronounced with both of its first two vowels devoiced.

Dake can also be used with the accentuation of kara':
ma'kura dake, sakana dake'

This form is called the 'tentative' by Bloch 1946.

This form is not to be confused with the one in which soo + Copula is attached to a tensed clause (e.g. sinbun o yo'nda soo desu 'I hear that he read the newspaper'). The soo of the latter construction has an underlying accent which is lost after an accented word. It does not trigger accent-attraction.

This form is used both for analogues to the English passive and for 'affective' passives with meanings such as 'He was subjected to his son's dying'.
According to Akinaga (1958), compounds of a 1- or 2-mora native morpheme with a 1-mora native morpheme are unaccented if the first member is unaccented and have accent at the end of the first member if the first member is accented:

- hana' ti → hanazi 'nosebleed'
- tama' ko → tama'go 'egg'

-a> kosiyu 'hip bath'

- kosi + yu' → a'me + to → ama'do 'storm window'

-zin 'inhabitant of' bears an accent if attached to something final-accented but is preceded by an accent if attached to anything else:

- amerika 'America'
- do'itu 'Germany'
- supe'in 'Spain'
- nihon 'Japan'
- tyoose'n 'Korea'
- taiwa'n 'Taiwan'

- amerika'-zin
- doitu'-zin
- supein'-zin
- nihon-zin
- tyooseen-zin
- taiwan-zin

In this and some other compounds given here, the first member does not occur as an independent word.

I adopt here the makeshift of using @ to mark the boundary between separate phrases, % to mark the boundary between phrases of a phrase group, and " to indicate accented syllables that are on only a 'mid' rather than a high pitch.

See Schmerling (1973) for a demonstration that the nuclear stress rule corresponds to only one of several factors that are jointly responsible for sentence stress in English.

In one dialect, that of Narada (Yamanashi pref.), what is distinctive is not place of fall in pitch but the place of rise in pitch. The Narada system has evolved from a system essentially identical to that of the standard language by a change in which all moras up to the one after the accent inverted their pitch (i.e. high pitch becomes low and low becomes high):

- standard
- usagi ga
- kabuto ga
- kokoro ga
- atama ga
- Narada
- usagi ga
- kabuto ga
- kokoro ga
- atama ga
Hattori's data are for the dialect of Kameyama, Mie pref., which agrees with the data given so far except for having additional accentual types for 3-syllable nouns, namely final-accented forms such as /asita/ 'tomorrow' and /'tikame/ 'near-sighted' (p. 441).

The distinction is preserved in some Kansai dialects, for example that of Kochi (Southern Shikoku):

<table>
<thead>
<tr>
<th>Unaccented in standard language</th>
<th>Accented in standard language</th>
</tr>
</thead>
<tbody>
<tr>
<td>atu'i 'thick'</td>
<td>a'tui 'hot'</td>
</tr>
<tr>
<td>oso'i 'late'</td>
<td>na'gai 'long'</td>
</tr>
<tr>
<td>kura'i 'dark'</td>
<td>ku'roi 'black'</td>
</tr>
</tbody>
</table>

However, even in Kochi, the distinction is lost in longer adjectives.

See exx. (21)ff for illustrations of this and of a qualification that it is subject to.

In §3, I will argue that such forms existed in 11-th century Kyoto and in proto Japanese.

This form is pronounced akakereba. Kansai dialects allow composite words in which there are two falls in pitch within the same word, though each must be contributed by a separate morpheme. Accent deletion in such words is optional: 'aka'kereba is an alternative pronunciation. I do not know of any generalization that distinguishes the cases in which accent deletion is optional from those in which it is obligatory.

'si_yahen 'does not do' and de'yahen 'does not leave' also occur. -hen takes the form -hin after /i/.

I have no idea why this word is low-initial.

The data given here are taken from Wada 1962 and Tōjō 1961, vol. 4.

The data cited here are from Hirayama (1966a, 1967).

In this discussion I do not consider Amami and Ryūkyū data. In Amami-Ōshima, Yaeyama-Kuroshima, and part of the Miyako archipelago, set 3 falls together with sets 1 and 2 rather than with sets 4 and 5 (Hirayama 1966a, b, 1967).
A distinction between sets 2 and 3 is also found in a portion of Shizuoka prefecture, in one locality in the Noto peninsula, and in much of Shikoku and some islands in the Inland Sea. See Tokugawa 1962 for details.

Except for a couple of localities where sets 1, 2 and 3 remain mutually distinct: Togi (Noto peninsula) and Ibuki Island (Kagawa pref.). (ili) is a gross oversimplification. See Tokugawa (1962) for reasons why the coalescence of sets 2 and 3 was an independent development in several different regions.

In most 'Tōkyō-type dialects' (e.g. Numazu and Hiroshima, which otherwise agree closely with Tōkyō in accent), these words have 00′0. In Tōkyō, 00′0 is comparatively rare in 3-syllable nouns: aside from compound and derived nouns, only koko′ro has that accentuation. Tōkyō thus appears to have undergone a shift 00′0 → 0′00. The same change has occurred in Kyōto, though since Kyōto has not undergone the general rightward shift, it affects different nouns, namely types 2 and 4.

Though high-initial words in the Ruljumyōgishō are generally high-initial in modern Hyōgo and Kyōto, those type 2 nouns which are unaccented in Tōkyō are mainly low initial in Hyōgo and Kyōto.

The one exception is that in four-syllable phrases with first-syllable accent, three of Weitzman's four informants show a greater drop between the second and third syllables than between the first (accented) and second syllables.

Zeps' observations were contained in a paper read at the Dec. 1966 meeting of the Linguistic Society of America.
Towards a Reconstruction of Uto-Aztecan Stress

Pamela Munro

University of California, Los Angeles

Larry Hyman has discovered that initial, penultimate, and final stress are all about equally common in languages having non-phonemic stress, but that the fourth most common stress pattern observed, second-syllable stress, is much much rarer, occurring in only three percent of the languages he sampled. The stress patterns reconstructable for the Uto-Aztecan family of western North America, however, appear to be of this very rare type, or one very much like it. In this paper I will describe the evidence which leads me to reconstruct such an unusual stress pattern for Proto-Uto-Aztecan, discuss some of the ways the modern Uto-Aztecan languages have diverged from the earlier system, and suggest how such a system might have evolved in Proto-Uto-Aztecan. Part of Hyman's goal was to provide a "psychological" explanation for why some stress patterns occur more often than others—primarily because some patterns more effectively mark word boundaries. Rarer patterns are evidently less effective in this regard and thus, Hyman suggests, may be more susceptible to changes of various sorts. Uto-Aztecan offers some confirmation of Hyman's proposals, since so few of the modern daughter languages preserve the proto stress pattern unchanged. It would appear that any phonologists interested in the functional nature of stress should be interested in studying the Uto-Aztecan evidence.

On first glance, the Uto-Aztecan languages present an incredible diversity of stress patterns—it is probably because of this diversity that Voegelin, Voegelin, and Hale (1962), in their pioneering discussion of Proto-Uto-Aztecan phonology, declined to reconstruct a proto stress pattern (p. 34).

In his fairly conservative classification of the Uto-Aztecan languages (1964), Sydney Lamb recognized nine independent branches of the family. These are, roughly from north (Great Basin) to south (central Mexico): Numic, Tubatulabal, Giamina, Takic, Hopi, Pimic, Taracahitic, Corachol, and Aztec. (I have substituted more common names for some of Lamb's terminology.) In some earlier classifications (the most famous being that of Sapir (1929)—see Lamb (1964) for discussion) these nine sub-families are grouped into just three divisions—a northern group, known as Shoshonean, including the languages of Lamb's first five sub-families; a southern group, known as Sonoran or Piman, the next three sub-families; and a third group
consisting solely of the various Aztec dialects. The three-branch subgrouping is more useful for some purposes than the unwieldy nine-branch one, but the stress data to be presented below lend it no great support. The extinct language Giamina is known to us only through a brief description and list of about twenty words (without stress marking) presented by Kroeber (1906-07), and will not be further considered.

I consulted descriptions of twenty-seven other Uto-Aztecan languages and dialects (including in some cases my own field notes) to arrive at the conclusions presented here. The sub-grouping of these languages within the sub-families of Uto-Aztecan named above is given in Table 1. (Sub-families which do not branch—Túbatulabal and Hopi—consist of single languages without significant dialect differentiation.)

The bewildering array of stress patterns reported for these languages is presented in Table 2, in which languages are listed only by sub-family, without any indication of closer relationships. The source or sources I used to arrive at the listed descriptions appears in the table. When different sources present different descriptions of stress phenomena, each is listed, and attributed.3 If no source is listed (as for Kawaiisu and Pima), the description is based on my analysis of data from my own field notes.

Some commentary on the terminology used in the table is in order here. First of all, I consider a language to have "lexical" stress (I might have chosen the terms "phonemic" or "contrastive") if stress alone appears to play a role in differentiating significant numbers of classes of lexical items (e.g. if both CVCVC and CVVC are frequently occurring canonical forms, as in Serrano and Luleño, for example). A language like Northern Paiute, for which Nichols (1974) reports the possibility of a very small number of minimal pairs for stress, would not be considered to have lexical stress, since stress is predictable for the great majority of words in the language. (The Northern Paiute situation is not atypical, incidentally. Scattered exceptions to the general descriptions given in Table 2, for which stress must be lexically marked, are reported by Voegelin (1935) for Túbatulabal, Seller (1965, 1967) for Cahuilla, Lindenfeld (1973) for Yaqui, Preuss (1933) for Cora, and Boas (1917) for the Pochutla dialect of Aztec.)

Vowel length is contrastive in most modern Uto-Aztecan languages, and should be reconstructed for Proto-Uto-Aztecan.4 In some cases, vowel length is important in determining the placement of stress. The term "second-mora" which appears several times in Table 2 is used to describe a situation in which the first syllable of a word is stressed if it contains a long vowel or a diphthong, but the second syllable is stressed if the first syllable vowel is short. If long vowels and diphthongs count as two moras, or vowel-units, and short vowels count as one mora, it is clear that in such languages stress always falls on the second mora of the word.
Table I
Genetic Grouping of Twenty-seven Uto Aztecan Languages and Dialects

<table>
<thead>
<tr>
<th>Sub-family</th>
<th>Sub-sub-family</th>
<th>Language</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WESTERN</td>
<td>Northern Paiute</td>
</tr>
<tr>
<td>NUMIC</td>
<td></td>
<td>Mono</td>
</tr>
<tr>
<td></td>
<td>CENTRAL</td>
<td>Shoshoni</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comanche</td>
</tr>
<tr>
<td></td>
<td>SOUTHERN</td>
<td>Ute</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Southern Paiute</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Chemehuevi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kawaiisu</td>
</tr>
<tr>
<td>TUBATULABAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAKIC</td>
<td></td>
<td>Serrano</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Kitanemuk</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Luiseño</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cupeño</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cahuilla</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wánikik Cahuilla</td>
</tr>
<tr>
<td>HOPI</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Papago</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pima</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Northern Tepehuan</td>
</tr>
<tr>
<td>PIMIC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TARACAHITIC</td>
<td></td>
<td>Yaqui</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Varohio</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tarahumara</td>
</tr>
<tr>
<td>CORACHOL</td>
<td></td>
<td>Huichol</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cora</td>
</tr>
<tr>
<td>AZTEC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Classical Nahuatl</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Modern Aztec Dialects</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pochutla</td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>Subfamily</th>
<th>Language</th>
<th>Stress Pattern</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numic</td>
<td>Northern Paiute</td>
<td>Second mora</td>
<td>Nichols 1974</td>
</tr>
<tr>
<td></td>
<td>Mono</td>
<td>Penultimate</td>
<td>Lamb 1958</td>
</tr>
<tr>
<td></td>
<td>Shoshoni</td>
<td>Root–initial</td>
<td>Miller 1975</td>
</tr>
<tr>
<td></td>
<td>Comanche</td>
<td>Lexical</td>
<td>Smalley 1953</td>
</tr>
<tr>
<td></td>
<td>Ute</td>
<td>Lexical</td>
<td>Goss 1962</td>
</tr>
<tr>
<td></td>
<td>Southern Paiute</td>
<td>Second mora</td>
<td>Sapir 1930–31</td>
</tr>
<tr>
<td></td>
<td>Chemehuevi</td>
<td>Second mora</td>
<td>Press 1974</td>
</tr>
<tr>
<td></td>
<td>Kawaiisu</td>
<td>Penultimate</td>
<td></td>
</tr>
<tr>
<td>Tubatulabal</td>
<td></td>
<td>Final</td>
<td>Voegelin 1935</td>
</tr>
<tr>
<td>Takic</td>
<td>Serrano</td>
<td>Lexical</td>
<td>D. Crook</td>
</tr>
<tr>
<td></td>
<td>Kitanemuk</td>
<td>Word–initial</td>
<td>A. Anderton</td>
</tr>
<tr>
<td></td>
<td>Luiseno</td>
<td>Lexical</td>
<td>e.g. Bright 1965</td>
</tr>
<tr>
<td></td>
<td>Cahuilla</td>
<td>Lexical</td>
<td>Hill &amp; Hill 1968</td>
</tr>
<tr>
<td></td>
<td>Wánikik Cahuilla</td>
<td>Root–initial</td>
<td>Seiler 1965, 1967</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Word–initial</td>
<td>Seiler 1967</td>
</tr>
<tr>
<td>Hopi</td>
<td></td>
<td>Second–syllable</td>
<td>Kalectaca 1975</td>
</tr>
<tr>
<td></td>
<td></td>
<td>dominant</td>
<td></td>
</tr>
<tr>
<td>Pimic</td>
<td>Papago</td>
<td>Root–initial</td>
<td>Hale 1965</td>
</tr>
<tr>
<td></td>
<td>Pima</td>
<td>Root–initial</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Northern Tepehuan</td>
<td>Lexical tone</td>
<td>Bascom 1959, 1968</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Woo 1970</td>
</tr>
<tr>
<td>Taracahitc</td>
<td>Yaqui</td>
<td>Second syllablé</td>
<td>Lindenfeld 1973</td>
</tr>
<tr>
<td></td>
<td>Varohio</td>
<td>Lexical</td>
<td>Johnson &amp; Johnson</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1972</td>
</tr>
<tr>
<td>Corachol</td>
<td>Huichol</td>
<td>Lexical</td>
<td>McIntosh 1945</td>
</tr>
<tr>
<td></td>
<td>Cora</td>
<td>Lexical pitch</td>
<td>Grimes 1959</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Antepenultimate</td>
<td>Preuss 1932</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lexical pitch</td>
<td>R. W. Langacker</td>
</tr>
<tr>
<td>Aztec</td>
<td>Classical Nahuatl</td>
<td>Penultimate</td>
<td>e.g. Andrews 1975</td>
</tr>
<tr>
<td></td>
<td>Modern Dialects</td>
<td>Penultimate</td>
<td>e.g. Whorf 1946</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Milpa Alta)</td>
</tr>
<tr>
<td></td>
<td>Pochutla</td>
<td>Final</td>
<td>Boas 1917</td>
</tr>
</tbody>
</table>
Hyman chooses to describe the second-mora stress pattern as "second-syllable dominant"—the "dominant" stress placement, in his framework, is the "elsewhere" case: stress is placed on the second syllable except the first syllable contains a long vowel or diphthong. (Where there is a rule of "dominant" stress placement, in other words, stress may be deflected from the "dominant" syllable to a "heavy" syllable elsewhere in the word, with the heavy syllables those which are closed or contain either a long vowel or a diphthong.) I have used the specific term "second-mora" in Table 2 for languages with the stress pattern just described because this pattern is so common in the Numic sub-family of Uto-Aztecan and because it seems to me that Hyman's "second-syllable dominant" characterization is better reserved for more complicated stress patterns like that found in Hopi (as indicated in Table 2). The general Hopi stress rule presented in the lessons prepared by University of California, San Diego, Hopi grammar project (Kalectaca, 1975) is this: stress the second syllable of a word unless the vowel of the first syllable is long or followed by two consonants (the first of which might be a glide), in which case stress the first syllable—in other words, stress the second syllable unless the first is "heavy". The second-mora rule will not work for Hopi, since in a word with the form CVCCVCV, for instance, the second mora, the vowel of the second syllable, will not receive stress even if the first vowel is short.

The difference between "root-initial" and "word-initial" stress may be seen in a comparison of the same word in the Wánikik dialect of Cahuilla and other Cahuilla dialects, as described by Seller (1967). Stress is on the first syllable of the Wánikik word nētaxaw 'my body', but this word appears in the Desert and Mountain dialects of Cahuilla as netaxaw, with the second syllable stressed. The initial syllable of this word is a separate morpheme, the first-person singular possessive prefix nē-, so the stress pattern in other Cahuilla dialects may be described as root-initial. In a non-prefix word like nāxaniš 'man', stress falls on the first syllable in all dialects, since the first syllable of the root is the first syllable of the word. Prefixes like nē- occur on all possessed nouns and some verbs; such prefixes are one syllable long (Seller, 1967, pp. 141, 145). In a language like Shoshoni, unstressed pre-root elements in the phonological word are not prefixed but cliticized, but the general situation (as described by Miller, 1975) is similar to that of Wánikik Cahuilla.

In addition to the scattered lexically marked exceptions to the general stress rules outlined in Table 2 which I mentioned above, there are in some languages various sub-regularities which constitute minor stress rules. For instance, Southern Paiute and Hopi have a special rule by which all two-syllable words are stressed on the first syllable, regardless of their CV structure. Occasionally morphological structure may determine departures from the usual stress pattern: Miller (1975) reports that certain forms of the Shoshoni verb paradigm may depart
from the usual root-initial stress pattern, and it is well-known that the only class of exceptions to the penultimate stress rule of most Aztec dialects are vocatives, ending in the stressed morpheme -e.

A final regularity which is frequently found in various Uto-Aztecan languages is the presence of a rule of "alternating" stress, by which a number of secondary stresses are placed within a word, depending on where the primary stress falls. Generally, after primary stress is placed, every second mora or syllable away from the primary stressed vowel receives a secondary stress. Sapir's (1930) example of the rule of alternating stresses in his class description of Southern Paiute is a word meaning 'even if he says to me': ai-yu-sampa-apa-ni, phonetically [aiYUSampanaI]. Stress falls on the first syllable, a diphthong, since it contains the second mora of the word, and every second mora thereafter gets a secondary stress. (Final vowels may not receive primary or secondary stress, and unstress, and unstressed vowels are subject to various types of reduction--devoicing, as in Southern Paiute, or syncope, in some other cases.) The placement of secondary stress by the alternating stress rule moves away from the word boundary whose position initially determined the placement of primary stress. In Southern Paiute, as shown in Table 2 and the above example, primary stress is second-mora, positioned relative to the beginning of the word, and the alternation of stresses continues toward the end of the word--every second mora away from the primary stress toward the end of the word receives a secondary stress. In Mono, on the other hand, stress is placed on the penultimate syllable originally--i.e., the position of the primary stress is determined relative to the end of the word--and secondary stresses alternate away from the primary stress toward the front of the word. Regular alternation of stress is found in most Numic languages (probably not in Comanche or Kawaiisu--cf. Miller (1975, p. 12)), in Tübatulabal, and in Cahuilla, and to some degree in various other languages.

It is clear that these remarks on what the stress pattern descriptions given in Table 2 mean could be greatly expanded. I have chosen not to give examples from each of the languages discussed, because extensive exemplification is readily available in the original sources. Merely citing a form with penultimate stress marked is not enough to demonstrate that a language has penultimate stress, after all. For such demonstrations, the curious reader is referred to the sources named in Table 2.

From the point of view of deciding what the stress pattern of Proto-Uto-Aztecan must have been, the data in Table 2 looks extremely confusing. There are languages with predictable stress and languages with lexical stress. Among the languages with regular stress rules there are languages in which stress is reckoned from the beginning of the word (word-initial, stem-initial, second-mora, second syllable) and languages in which stress is reckoned from the end of the word (final, penultimate, antepenultimate).
Given Hyman's (1975) survey of stress patterns around the word, however, it is striking how many of the Uto-Aztecan languages appear to have some form of second-syllable stress. As noted above, roughly three percent of the languages for which Hyman had data had second-syllable or second-syllable dominant (including second-mora) stress. Yet Table 2 shows that the Uto-Aztecan languages Northern Paiute, Southern Paiute, Chemehuevi, Hopi, and Yaqui have stress patterns of this type. It is not entirely clear how many of the twenty-seven "languages" named in Tables 1 and 2 are in fact distinct—Southern Paiute and Chemehuevi, Cahuilla and Wánikik Cahuilla, Papago and Pima, Classical Nahuatl and Milpa Alta Aztec (and even perhaps Pochutla) are all possibly better considered dialects than separate languages although the evidence is not all in. Supposing, because of this uncertainty, we assume that we are dealing with data from, say, twenty-five languages, of which five have (in Hyman's terms) second-syllable stress. If Hyman's findings are interpreted to mean that there is a three-percent chance of any randomly selected language having second-syllable stress, it is clear that Uto-Aztecan has an extremely high number of languages displaying this rare stress pattern. In fact, the probability of finding five or more languages with second-syllable stress by chance in a sample of twenty-five languages can be calculated as just eight chances in 10,000 (.0008). It seems likely, then, that the incidence of second-syllable stress in Uto-Aztecan is not just a random occurrence, but due to some historical factor.

In other words, if Proto-Uto-Aztecan had second-syllable stress, it would not be too surprising to find that some of the modern Uto-Aztecan languages still should exhibit this pattern. If, on the other hand, Proto-Uto-Aztecan had some other stress pattern, and five of the languages in my sample independently innovated second-syllable stress patterns, this would be most unlikely (the odds against such innovation being, by the above calculation, greater than a thousand to one).

Notice that the non-second-syllable stress patterns listed in Table 2 are almost all reasonably common ones. Lexical, initial, penultimate, and final stress languages were all well represented in Hyman's study. We might suspect, in fact, that given as unusual a stress pattern as the second-syllable one we are beginning to hypothesize for Proto-Uto-Aztecan, it would be expected that the grammars of the various daughters of Proto-Uto-Aztecan might gradually change so that stress placement would follow a commoner and more natural rule.

The line of argumentation just outlined almost forces one to the conclusion that Proto-Uto-Aztecan must indeed have had second-syllable stress. And if this conclusion is taken as a starting point, a number of additional pieces of evidence emerge which lend support to the basic premise. There are a number of subregularities and exceptions to the general stress patterns presented above and in Table 2 which may be viewed as relics of an older second-syllable pattern. Taken together, the evidence for proto second-syllable stress is quite convincing.
In the remainder of this paper I will examine each of the eight branches of Uto-Aztecan for which we have data (as noted before, Giamina Is excluded), noting for each the various factors which support the reconstruction of an earlier second-syllable stress pattern, and describe some of the changes which have taken place in the evolution of the modern stress system.

Finally, I will consider in more detail the question of exactly what Proto-Uto-Aztecan stress must have been like, and how it might have gotten that way.

Within a given subfamily of Uto-Aztecan, any occurrence of second-syllable or second-syllable dominant stress must be taken as the strongest type of support for the reconstruction of such a stress pattern. Failing this, we should next hope to find some indication that stress might once have been placed on the second syllable. If this too was lacking, an initial stress pattern also might be considered to be consistent with an earlier second syllable stress rule, since initial stress could arise simply by a simplification of a "second syllable dominant" stress rule, like the second-mora stress rule or the Hopi stress pattern described above. The most problematic case, in fact, would be one in which stress was reckoned from the end of the word, with no indication of some earlier change from another stress pattern.

The northernmost sub-family of Uto-Aztecan, Numic, offers a great deal of support for the reconstruction of second-mora stress of Uto-Aztecan. This pattern is found today in two of the three Numic subgroups: within Western Numic in the Northern Paiute language, and within Southern Numic in the Southern Paiute and Chemehuevi languages. The Central Numic languages appear to have changed this pattern to one of regular root-initial stress, although in Comanche stress may be becoming phonemic. The examples cited by Smalley (1953) to demonstrate the contrastive function of Comanche stress include minimal pairs like nimikuhcu? 'buffalo'/nimikucu? 'our cow', for instance, which are consistent with a root-initial analysis: the second form includes the first-person plural clitic nimt-, which would not be expected to be stressed.

However, despite the tendency to a rule of root-initial stress in Central Numic, many words in these languages do have stress on the second syllable of the word. Since most of the possessive and subject clitics which may be prefixed to a root (in contrast to Comanche nimt- in the above example) are only one syllable long--Miller describes them, for Shoshoni, as "single mora unstressed elements" (1975:13)--stressing the first syllable of the root gives the same perceptual result as stressing the second syllable of the word. Furthermore, in Shoshoni there are a number of exceptions to the root-initial stress rule, for which stress is on the second syllable of the root--Miller (1975:12-13) mentions that some words with a long vowel in the second syllable will be stressed on that syllable, and many verbs, especially in the durative form, are stressed this way. These "exceptions" might be viewed as relics of an earlier (Proto-Numic) second-mora stress rule.
Lexical stress on the first or second syllable is also reported by Goss (1962) for Ute, within Southern Numic. It seems clear that stress became contrastive in Ute at the same time as the vowel length distinction present in other Numic languages was lost, since Goss reports that Ute has no contrastive long vowels. The only minimal pair for stress which he cites, $\text{pa-kê}\ddagger 'fish'/\text{pa-kê} 'trout'$ is fairly easily accommodated within this historical account. The cognate stem from closely related Southern Paiute, $\text{pa-gu} 'fish'$, is a compound, containing the common Uto-Aztecan stem for 'water', $\text{pa-}$, which has in Southern Paiute two variant forms, $\text{pa-}$ and $\text{pa-}$ (Sapir 1930:597). If we assume pre-Ute to have had competing forms $\text{pa-kê}\ddagger$ and $\text{pa-ki-tê}$ for 'fish', with the less common of these, the second, becoming specified to mean 'trout', the path of development is clear. With vowel length and a second-mora stress rule, the words would be $\text{pa-kê}\ddagger 'fish'/\text{pa-ki-tê} 'trout'$; on the loss of vowel length, stress would become phonemic because of the resultant contrast between $\text{pa-kê}\ddagger$ and $\text{pa-kê}\ddagger$.

The most problematic cases within Numic and Mono and Kawaiisu, both of which have penultimate stress patterns. Since Mono and Kawaiisu represent separate branches of Numic, and since the other evidence just summarized fairly clearly supports the reconstruction of second-mora stress for Numic, there cannot be a genetic reason for the fact that speakers of both Mono and Kawaiisu reckon stress from the end rather than the beginning of the word. But there may be a historical connection nonetheless. The only other northern Uto-Aztecan language in which stress is reckoned from the end rather than the beginning of the word is Tübatulabal (which constitutes a separate sub-family of Uto-Aztecan by itself). It is noteworthy that the three end-stressing languages among the northern Uto-Aztecan groups, Mono, Kawaiisu, and Tübatulabal, are all spoken within a relatively confined territory in south and east central California—Mono is the southernmost language in the Western branch of Numic, Kawaiisu is the westernmost language in the Southern branch of Numic, and Tübatulabal is spoken in area north of Bakersfield, along the Kern River. All three languages, consequently, are spoken very close to the area occupied until quite recently by speakers of Yokuts, a language of the Penutian family spoken in the southern part of the San Joaquin Valley in central California. All dialects of Yokuts (most now extinct) had regular penultimate stress (Newman, 1944). The Yokuts, then, determined stress placement from a final word boundary, rather than from an initial one. It is not unreasonable to suppose that speakers of the three Uto-Aztecan languages in closest contact with the Yokuts might have borrowed from them their method of marking word boundaries by stress. Kroeber (1906-07) cites other evidence of borrowing from Yokuts into these languages, so it is clear that there was linguistic contact between these groups. 7

If this hypothesis as to the origin of end-stressing in Mono, Kawaiisu, and Tübatulabal is correct, it is probably the case that Tübatulabal developed its present pattern of final stress with a few lexically marked exceptions after going through a period in which stress was penultimate.
Within the Takic subfamily, initial stress is reported for the Cahuilla and Kitanemuk languages; Kitanemuk and most Cahuilla dialects have root-initial stress, while the Wânikik dialect of Cahuilla can be shown to have innovated a pattern of word-initial stress (Seiler, 1967).

Seiler (1967) proposed that the Cahuilla pattern of root-initial stress was reconstructable for Proto-Cupan, the subgroup of Takic which includes Cupeno and Luiseño as well as Cahuilla, but this suggestion has been convincingly refuted by Hill and Hill (1968), who demonstrated that the Cahuilla pattern must be a regularization of a pattern of lexical stress in Proto-Cupan. They showed that in a number of words a particular lexically marked stress may be reconstructed not just for Cupan but for Proto-Takic, since there is confirmatory data from Serrano.

Serrano stress, according to Donald Crook, is not predictable by rule, although there is a tendency for long vowels or initial vowels to be stressed. (Since Serrano lexical stress agrees in some cases with Luiseño and Cupeno lexical stress, it seems that the development of initial stress in Kitanemuk must have been an innovation.)

Cupeno has a fairly unusual synchronic stress system, according to the analysis of Hill and Hill (1968): most roots have first- or second-syllable stress lexically marked, and there are certain suffixes which have the effect of causing stress to fall on the last vowel of a preceding root. In addition, there is a small set of "stressless" roots with no inherently marked stress: a word containing one of these is stressed initially (on the first syllable of the root, or on a prefix, if present) unless it contains one of the stress-attracting suffixes.

The problem of stress in Luiseño is quite tantalizing. The placement of stress in most words may be determined by rule: stress the first vowel of a stem unless the second vowel is long. This rule works for most of the languages, and interacts believably with other phonological processes in Luiseño (Munro & Benson, 1973); it does not, however, handle certain groups of problematic "exceptions". One such group is a type of verb stem reduplicated to show intensity: from an initially stressed CVCC is formed a CVCC-CVCC derivative stressed on the third syllable, e.g. cipi 'break'/cipi-lipi 'shatter'. Problems concerning the interaction of reduplication processes and other rules frequently arise cross-linguistically (cf. Wilbur, 1973); one may after all regard the first CVCC of the reduplicated form as something added onto the beginning of the CVCC base after stress has been placed, by rule, on its first syllable.

The remaining sets of stress problems are more interesting for the purposes of our present investigation. First of all, there is a fairly substantial number of CVCC noun roots in Luiseño in which both vowels are short which are stressed, unexpectedly, on the second syllable. Compare, for instance, kwas-ı-t 'fog', which follows the rule given above, with the exceptional wiʔé-t 'grasshopper' (the suffix in both words is an absolutive ending).
As given above, the Luiseño stress rule determines stress placement within the root: addition of a possessive prefix, for example, does not change the placement of stress: ku:-ca 'house'/po-ki: 'his house', etc. (absolutives drop in the possessed form). For a small group of lexically marked words, however, a singular possessive prefix may receive stress: pó-gwi 'his forehead'. The curious thing is that this third group of exceptions to the stress rule intersects to some degree with the second: in the non-possessed forms of these words which are stressed on a singular possessive prefix, the second syllable is stressed, often irregularly—c.f. cuwi-l 'forehead'. (Note that the first vowel of the stem cuwi- is syncopated in the possessed form.)

This group of exceptions, then, shows not root-initial stress but rather an alternation between word-initial (prefix) stress and second-syllable (of root) stress—a determination of which of the two patterns is basic in this group of exceptions to the regular rule may reveal something about earlier stress patterns in Luiseño.

Consider the possessed forms of the second-syllable stressed words sulá-t 'claw, fingernail' and qisi:vi-s 'tail', which are illustrated by pú-sla 'his claw' and pí-sliv 'his tail'. The quality of the possessive prefix vowel is important here. The underlying form of the prefix is clearly po-, although when unstressed it is often heard as pu- (e.g., po-ki: 'his house' may be heard as pú-ki:—compare, for instance, the stressed independent (primarily relative) pronoun po (from the general Proto-Uto-Aztecan pronominal element *pi-; *p is regularly reflected in Luiseño as o). The appearance of the vowels u and i in the stressed prefixes in pú-sla and pí-sliv, then, needs to be explained, and I believe that the explanation lends some support to the idea that the second-syllable stress pattern is basic in these words.

There seems to be a low level tendency in Luiseño for unstressed vowels to "harmonize" somewhat with other nearby vowels in the word under certain conditions—some such vowel harmony rule appears to have operated historically to produce the unexpected vowels in the possessive prefix in words like 'his claw' and 'his tail'.

From an underlying po-sula and po-qisi:vi(l), then, low-level vowel harmony would produce pu-sula and pi-qisi:vi(l). The synchronic harmony process described above only operates on unstressed vowels, however; elsewhere in Luiseño, a number of rules operate to alter the quality of unstressed vowels, but I know of no rule which changes the quality of stressed vowels. It seems likely that the historical rule may have been similarly restricted. 9

This evidence, then, suggests that the second-syllable stress preserved in the nonpossessed forms of the words in question is older: we may hypothesize a historical development of po-sulá to pu-sulá preceding the shift of stress to the initial syllable, giving pu-sula; a regular syncope process produces the synchronic pú-sla.
I believe that this group of Luiseño noun stems reflects a much earlier pattern of regular second-syllable or actually second-mora stress in pre-Proto-Takic. The earlier second-syllable stress is retained in their non-possessed forms, and the evidence of the vowel harmony process which has operated in their possessed forms indicates that the present stress in these forms is innovative. (An interesting question for which I can offer no solution at present is that of why stress should have shifted to the first syllable in these possessive forms. As noted above, except for such words, possessive prefixes are never stressed in Luiseño.) Hill and Hill (1968) have characterized this group of stems synchronically for Luiseño as "stressless"; many of these are cognate to the stressless roots they list for Cupeno. It appears that these small groups of "stressless" roots in both Cupeno and Luiseño have preserved an old regular second-mora stress rule—the unaffixed Cupeno roots all have long initial vowels, which receive stress, while the Luiseño roots, with short initial vowels, have second-syllable stress (cf. Hill & Hill, 1968:240). 10

The Takic evidence, then, although complex, tends to support the hypothesized development from an earlier second-syllable (or -mora) stress pattern. I believe that a detailed reconstruction of Proto-Takic stress will reveal a lexical stress system similar to those preserved in Luiseño, Cupeno, or Serrano, with an "initial-dominant" stress rule to which there are many lexical exceptions. The nature of the exceptions, like those just described for Luiseño, will, I predict, reveal relics of an earlier second-syllable stress pattern in the process of changing to an initial stress pattern. In Cahuilla and Kitanemuk, then, regular initial stress patterns of various sorts have resulted from a regularization of the irregular pattern of lexical stress with a tendency toward initial stress at the Proto-Takic level.

The Hopi pattern of "second-syllable-dominant" stress described above (stress the first vowel in a word if it is long or followed by two consonants, otherwise stress the second vowel) is good support for the postulation of Proto-Uto-Aztecan second-syllable stress. Such a stress pattern could have easily developed from a second-mora stress rule—it is identical with such a rule except for the requirement that the first vowel of a word be stressed if it is followed by two consonants. Almost any consonant cluster in a word in any present-day Uto-Aztecan language may be taken as evidence that some vowel deletion must have occurred—-the proto-language, according to all suggested reconstructions, was strictly CVCV. 11 I believe that the synchronic Hopi stress pattern reflects the operation of a very common process in Uto-Aztecan: the syncope of an unstressed vowel in a syllable following a stressed syllable, particularly one containing a long vowel. In Hopi, then, words longer than two syllables may have three different stress configurations for the first two syllables: CVVCV-, CVCV-, and CVCCV-. The first two are consistent with a second-mora stress pattern, and the third is presumed to reflect a development from an originally second-mora-stressed *CVVCVCV to CVVCVCV to CVCCV. 12
Further confirmation for the idea that Proto-Uto-Aztecan must have had a second-syllable dominant stress system comes from the Pimic subfamily. In Papago and Pima, stress is regularly root-initial (cf., for example, Hale, 1965), a system which, as we have seen, could have arisen as a regularization of a more marked way of reckoning stress from the beginning of the word. 13

In the Northern Tepehuan language there is a system of phonemic tone (Bascom, 1959), with a four-way contrast between low, high, rising, and falling tone on long vowels. Woo (1970) has shown, however, that the placement of tone is predictable given historical information. By Woo's formulation, high tone is assigned to the first vowel of the second syllable in words of more than two syllables. (Note that high tone never occurs on any but the first or second vowel cluster in a word.)

Woo's abstract description of a synchronic phonology of Northern Tepehuan has most value as a suggestion of how such a system could have evolved historically. It seems clear that the present tonal system in Northern Tepehuan has developed out of an earlier stress system, with primary stress being replaced, historically, by high tone. The stress system from which a tonal system like that of Northern Tepehuan could have developed would have been a second-syllable dominant one. So this evidence from Northern Tepehuan shows that the Pimic sub-family of Uto-Aztecan also reveals traces of an earlier second-syllable stress pattern.

The development of tone contrasts in Northern Tepehuan is paralleled by comparable developments in the other Sonoran sub-families of Uto-Aztecan—indeed, the tendency toward the development of contrastive tone may constitute a characteristic of the southern Uto-Aztecan groups in general, since such developments are found in Taracahitic and Corachol as well as Pimic—every southern sub-family except Aztec.

For instance, tonal as well as stress contrasts are reported for Yaqui (Taracahitic sub-family) by Johnson (1962) and Crumrine (1961), although both investigators appear to conclude that the placement of tone is determined by the grammatical structure of the phrase. Although it appears that Yaqui must be analyzed as having lexical stress, my most recent source on Yaqui, Lindenfeld (1973:4) reports that most words have second-syllable stress.

This development appears to be definitely secondary in Taracahitic; only Yaqui has innovated any tonal phenomena. TARAHUMARA, for instance, has a system of lexical stress which probably merits further study from a comparative point of view—both Lionnet (1972) and Burgess (1970) report that stress may fall on any of the first three syllables of a Tarahumara word. A system of lexical stress appears to exist in the little-known Taracahitic language Varohio, in which stress is apparently predominantly final or penultimate. There is a suggestion, however, in the Varohio words cited by Johnson and Johnson (1947) that the unstressed syllables before the stressed vowel in a Varohio word might be grammatical prefixes (or perhaps proclitics)
of various sorts, and that, in fact, stress placement might actually be on the first vowel of the root. (More grammatical work on Varohlo is certainly called for.) All in all, the data from Taracahitic suggests that the system of lexical stress in these languages might have developed from an earlier system in which stress placement was determined from the beginning rather than the end of the word, at least. Lindenfeld's observation that the predominant stress pattern in Yaqui is second-syllable is at least consistent with the hypothesis that Taracahitic could have developed from an earlier language with a second-syllable stress rule.

The synchronic situation in the Corachol sub-family is the least clear in Uto-Aztecan. Early sources on Cora and Huichol report lexical stress in both languages. Preuss (1932) in fact found Cora to have a predominantly antepenultimate stress pattern (extremely rare cross-linguistically, according to Hyman's survey). McIntosh (1945) claimed that the majority of words in Huichol are stressed on the penult. These two observations, taken together, would suggest that Proto-Corachol must have reckoned the placement of regular stress from the end rather than the beginning of the word. However, both these descriptions have been called into question by more recent investigators, who report instead that both languages have phonemic pitch or tone rather than stress.

Eugene Casad, for instance, has recently reported contrastive pitch for Cora (Ronald Langacker, personal communication); Grimes (1959) posits a basic lexical pitch for Huichol and shows that, given the pitch within a phrasal foot, the placement of stress is predictable. It is interesting to note, however, in terms of the present study, that placement of this secondary sort of stress in the Huichol word in Grimes' formulation is apparently determined relative to the beginning rather than the end of the word, which may later prove to be significant. Clearly, however, more work needs to be done on the comparative Carachol accentual system.

At first, the Aztecic sub-family appears to be a major problem for the proposed reconstruction. Classical Aztec had a strict penultimate stress rule, and most modern dialects of Aztec have preserved such a rule (cf., for example, Whorf, 1946). According to Boas (1917), the divergent Pochutla dialect of Aztec has predominantly final stress, which appears to have been derived from an earlier system like the Classical one (Pochutla is not a descendent of Classical Aztec). In many words final stress in Pochutla falls on the same vowel as in the penultimately stressed cognate in other dialects—simply because the original final vowel has been lost in Pochutla—cf. Pochutla apázt, standard Aztec apáztili 'olla'. (In some other cases, apparently, the stress has actually been shifted from the penultimate to the final syllable, presumably on analogy with forms like apáztl in which vowel loss has made an originally penultimate stress appear to be final.) In all cases, at any rate, stress appears to be determined with regard to the final rather than the initial word boundary.
There is some indication within Aztec, however, which suggests that at some much earlier date stress may have been determined with regard to the beginning of the word; in fact, it may have been regularly initial. Sapir (1913-15:419) determined, primarily on the basis of internal reconstruction within Aztec, that historical processes of vowel syncope in the language operate according to the following "law": "The first mora of a word is strong, the second weak, the third strong, the fourth weak, and so on alternately. A short vowel standing in a weak syllable is syncopated." Thus, for instance, the base form te-tli (root te plus absolutive -tli), with a strong-weak structure, becomes tetl 'rock', while the base *teni-tli (strong-weak-strong) becomes tentli 'lips'. The placement of stress on the derived penult in the last word (tentli) confirms the antiquity of this syncope process.

But the similarity of the strong-weak syllable marking and the placement of primary stress and development of an every-other-mora alternating stress system from that primary stress to the end of the word is too great to be ignored, particularly since the placement of alternating stresses in the northern languages which exhibit this phenomenon is so frequently followed by devoicing or even deletion of the vowels which are left unstressed. It seems most likely to me that the historical syncope law stated by Sapir reflects an earlier state of Proto- or Pre-Proto-Aztec at which stress was initial, with regularly alternating stress toward the end of the word. So even for this sub-family there is some indication that stress was once reckoned from the beginning rather than the end—and as we have seen, initial stress may quite easily arise as a normalization of the uncommon second-syllable dominant stress patterns.

A highly interesting question, which I do not have the evidence to consider further here, is that of how the habit of reckoning stress from the end of the word was innovated in such southern Uto-Aztecan languages as Aztec, Varohio, and possibly Cora and Huichol. I have no suggestions at this time about possible contact sources for this phenomenon such as I made in connection with final-stress reckoning in the northern languages Mono, Kawaiisu, and Túbatulabal.

In the previous discussion I have shown that there are many independent traces of an earlier second-syllable stress system among the various synchronic Uto-Aztecan languages, and that, further, when such direct evidence is lacking, there is still conclusive evidence in almost every sub-family that at some earlier period stress must at least have been reckoned in some other way (i.e., initially) from the beginning of the word. While this latter type of evidence does not confirm the second-syllable stress theory, it is at least consistent with it, since as Hyman has shown, it is likely that a marked second-syllable stress system might easily change to an initial one.

Two points remain to be considered at this time. First, there is the question of what sort of second-syllable stress system should be assumed for Proto-Uto-Aztecan.
Second, are there any indications of why such a highly marked stress pattern should occur in Uto-Aztecan at all? The answers to these two problems may be related.

Let us begin with the first question, that of what kind of second syllable stress pattern Proto-Uto-Aztecan might have had. Was it, for instance, an exceptionless second-syllable stress pattern, in which every word was stressed on the second syllable? Or was it some kind of "second-syllable dominant" system, to use Hyman's term, such as, perhaps, a second-mora stress rule, by which words would be stressed on the second syllable only when the first syllable did not contain a long vowel?

In my opinion, the evidence strongly favors the reconstruction of second-mora rather than second-syllable stress for Proto-Uto-Aztecan. First of all, in all the modern Ute-Aztecan languages which preserve some sort of second-syllable stress system, there is no language which has a regular second-syllable stress rule (Lindenfeld reports predominantly second-syllable stress for Yaqui, but there are many exceptions to such a generalization). However, a number of languages of the Numic sub-family (often argued on phonological and other grounds to be the most conservative sub-group within Uto-Aztecan) do preserve regular second-mora stress. We have seen how the slightly different second-syllable dominant stress pattern of Hopi may be seen as a fairly recent development from a second-mora stress system, and relic forms in Cupan seem to point to an earlier second-mora pattern.

Secondly, many languages with lexical stress (e.g., Ute, Luiseno, Cupeno, Serrano) have the restriction that main stress may occur only on the first or second syllable of a word. This clearly seems to be in accord with an earlier system in which stress fell by rule on either the first or second syllable, but it is harder to relate it to an earlier system in which initial stress was the rule.15 Taken together, I believe that both these facts support the claim that Proto-Uto-Aztecan stress was not second-syllable, but second-mora.

I believe that a further reason for the postulation of second-mora rather than second-syllable stress for Proto-Uto-Aztecan emerges from a consideration of our second question, that of why the proto language would have evolved so highly marked a system.

In a discussion of Siouan, another language group with the unusual second-syllable stress pattern, Hyman reports a suggestion of Wallace Chafe regarding the historical origin of this system. Chafe has proposed that some pre-Proto-Siouan stage had an unremarkable (root-) initial stress rule, but that many words were preceded by various grammatical prefixes, one syllable in length, which were unstressed. Even at this stage, then, many words appear to have second syllable stress. The adoption of a regular second-syllable stress rule for all words would simply constitute a generalization of the stress pattern perceived in words with prefixes to the rest of the language.
A similar argument could be made for Uto-Aztecan. A set of dependent pronominal elements is clearly reconstructable for Proto-Uto-Aztecan, descendents of which morphemes now function as subject agreement markers on verbs and possessive markers on possessed nouns in various synchronic Uto-Aztecan languages. It has not yet been determined, I believe, whether subject prefixes on the verbs of main clauses are reconstructable for the proto language, but I believe it is non-controversial to say that the use of possessive prefixes on nouns and probably certain dependent verbs (e.g., the verbs of complement and relative clauses) may be assumed for the proto language. At least the singular pronominal morphemes all have the shape of CV so one might hypothesize that second-syllable stress could have developed from earlier initial stress in Uto-Aztecan by the same process as that suggested by Chafe for Siouan.

The change from root-initial to regular second-syllable stress is argued to represent a generalization, a simplification of a perceptually confusing system. I believe, however, that a change from a root-initial system like that we have been describing for some stage of pre-Proto-Uto-Aztecan to a second-mora stress rule may also be seen as a generalization, and that, perhaps, for languages in which contrastive vowel length is important, it may even be a more likely sort of generalization to have occurred.

Consider first the change from root-initial to second-syllable stress diagrammed in Table 3. At Stage 1, there are words stressed on the first syllable (type B) and words stressed on the second syllable (type A— the + represents a prefix boundary), while at Stage 2, all words are stressed on the second syllable:

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. CV+CV</td>
<td>A. CV+CV</td>
</tr>
<tr>
<td>B. CVCV</td>
<td>B. CVCV</td>
</tr>
</tbody>
</table>

Vowel length plays no role here. What is involved is the reinterpretation of all the words with the structure B as having, for stress purposes, the structure A.

Now consider the change from second-syllable to second-mora stress. Since vowel length is crucial for determining the placement of stress in a second-mora stress system, we must consider whether the unprefixed forms have a long (type C) or (type B) vowel in the first syllable. However, there is no evidence that any of the CV pronominal prefixes reconstructed for Uto-Aztecan have long vowels, so the representation CV+CV
may still be used for all prefixed forms (type A). Again, Stage 1 shows root-initial stress, while stage 2 shows second-mora stress.

Table 4
Root-initial to Second-mora Stress

<table>
<thead>
<tr>
<th>Stage 1</th>
<th>Stage 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. CV+CV</td>
<td>A. CV+CV</td>
</tr>
<tr>
<td>B. CV.CV</td>
<td>B. CV.CV</td>
</tr>
<tr>
<td>C. CV.CV.CV</td>
<td>C. CV.CV.CV</td>
</tr>
</tbody>
</table>

Here, just as in the change shown in Table 3, simplification of a perceptually confusing pattern has taken place, a simplification which has the result of making stress placement a purely phonological matter for which morphological information (such as the location of a + boundary) is not necessary. The change from root-initial stress to second-mora stress involves the reinterpretation of words having the structure B as being like words having the structure A (with which they share the feature of having a first syllable whose vowel is short). Table 4 illustrates a smaller change than does Table 3, since fewer forms are affected by it: fewer words have a different stress at Stage 2 than their original stress at Stage 1 following the change to second-mora stress than following a change to second-syllable stress. The group of prefixed forms (type A in both tables) includes the same number of words at all times, but the group of words whose stress changes (B in both tables) is larger when the change is to a regular second-syllable stress pattern (Table 3) than it is when the change is simply to a second-mora stress pattern (Table 4). The change to second-mora stress from root-initial stress thus is a simpler change since it affects fewer words. This discussion provides an additional piece of inferential support for the reconstruction of second-mora rather than second-syllable stress for Proto-Uto-Aztecan.

Proto-Uto-Aztecan was a language in which vowel length contrasts were often important for distinguishing lexical items. Although little has yet been done toward reconstructing vowel length in many cognate sets, it seems clear that such contrasts are most often found in the first vowel of a CVCV root. If a language which has vowel length contrasts occurring most often in the first syllable of a word, such contrasts would be heightened if the stress on the word varied according to the length of the first syllable; in other words, if stress provided an additional cue to length, as in a second-mora system. However, if stress was regularly on the second syllable, while these contrasts were in the first syllable, the importance of such contrasts would be minimized. Therefore, I think that the morphological structure of Proto-Uto-Aztecan also makes the reconstruction of a second-mora stress pattern seem plausible.17
I believe I have presented enough evidence to show that Proto-Uto-Aztecan must be assumed to have had some kind of second-syllable stress pattern, despite the cross-linguistic rarity of such systems; further it seems correct to assume that this stress pattern must actually have been a second-mora stress rule. Of course, the statement that Proto-Uto-Aztecan had a second-mora stress pattern can undoubtedly be further refined—-it is likely, for instance, that an alternating stress rule and a special rule of initial stress for bisyllabic words will also prove to be reconstructable. Much remains to be done, too, to clarify the historical evolution of the rules used in the various daughter languages—-particularly important, it seems to me, are the development of lexical stress in Takic and lexical pitch in Corachol, and the source of Aztec's penultimate stress rule. Hopefully all these problems will be addressed in the future by Uto-Aztecanists, and by other phonologists interested in the phenomenon of second-syllable stress.

FOOTNOTES

1 I am grateful to Villiana Hyde (Luiseño), Lida Girado and the late Bertha Goings (Kawaiisu), and Ethleen Rosero (Pima) for their patience and generosity in teaching me about their languages. My work on Luiseño and Kawaiisu was supported by the Department of Linguistics, University of California, San Diego; my work on Pima, by the Department of Linguistics, University of California, Los Angeles.

I thank Allen Munro, Ian Maddieson, and Ronald Langacker for helpful comments about various parts of this paper. I am also grateful for the valuable suggestions and criticism I received from the participants in the Fourth Friends of Uto-Aztecan Working Conference (Long Beach, June, 1976), particularly Jeffrey Heath, Wick Miller, and Michael Nichols.

My understanding of Luiseño phonology owes much to discussion with Peter Benson.

2 I am familiar only with the preliminary version of Hyman's paper "On the Nature of Linguistic Stress" (1975), circulated prior to the Stressfest. The figure three percent was arrived at because Hyman reported second-syllable stress for twelve languages out of a total of "about 400" which he sampled.

3 Donald Crook's knowledge of Serrano is based on work with the last fluent speaker of that language, the late Mrs. Sarah Morongo Martin. I am grateful to him for providing me with a detailed demonstration that Serrano stress is contrastive.

I thank Alice Anderton for providing me with her impressions of what stress must have like in the extinct Takic language Kitanemuk, based on her study of recordings by the late J. P. Harrington, in which stress is not always marked.
Ronald Langacker has informed me that Eugene Casad of the Instituto Lingüístico de Verano believes that modern Cora has a contrastive pitch-accent. I am grateful to Langacker for giving me a copy of the manuscript of Casad and de Jesús (1975), a Cora text in which high and low pitch are marked.

4 Voegelin, Voegelin, and Hale (1962:34), in the study referred to above, list vowel length as well as stress among the "series generating components which ... remain to be reconstructed for Proto-Uto-Aztecan."

5 By the Binomial Theorem, the equation is

\[ P(5 \text{ or more}; 25, .03) = 1 - \sum_{k=0}^{5} \binom{25}{k} (.03)^k (.97)^{25-k} = .0008 \]

(I am grateful to Allen Munro and James Cunningham for help with the statistics.)

6 The exception to this generalization is Cora, whose antepenultimate stress pattern is even rarer than the second-syllable type, according to Hyman's study. Note, however, that Eugene Casad, a modern Cora scholar, claims that Cora has a lexical pitch-accent system (see footnote 3). Preuss (1932) claims that "sehr häufig fällt der Akzent auf die drittletzte Silbe", but lists a number of counterexamples to this generalization. Perhaps internal reconstruction will sometime clarify the Cora situation.

7 For instance, Kroeber notes that the Kawaiisu and Tübatulabal words for 'seven' and the Mono word for 'man' were borrowed from Yokuts. Incidentally, "Kawaiisu" is originally a Yokuts name.

8 Bright (1965) observed that Luiseno stress was predominantly second-syllable (in contrast to Munro and Benson's rule given in the text), I believe because the data he had examined for that article seems to have consisted almost entirely of nouns. There are, in contrast, almost no verbs with stressed second syllables containing short vowels.

9 Several examples could be cited to show that such a rule has operated in Cupan historically. Consider, for instance, the word 'one', e.g. Luiseno supul (one of the second-syllable stress nominal exceptions), Cupeño suplawit, Cahuilla supłyi 'one'/ supul 'other' (cf. Hill and Hill (1968:241), Miller (1967:#507-508). This word apparently derives from a Proto-Uto-Aztecan source something like *si-pul, where *si means 'one' and the morpheme *pul, according to the cognates suggested by Miller,
probably meant something like 'alone'. Since a stressed *u does not regularly produce \( u \) in any Cupan language, it would seem that this example too reflects (1) original stress on the second vowel (sipú), (2) vowel harmony (supú), and (3) in Cahuilla-Cupeño, a stress shift followed by syncope of the second vowel and other changes.

10 The notion of the stressless root, as described by Hill and Hill (1968), may be valid even for Proto-Takic. The comparative evidence suggests rather strongly, I think, that this concept is most useful in a primarily synchronic description, to help describe a stress system in a state of flux.


This statement ignores the central problem in Uto-Aztecan historical phonology, that of the well-known alternations between "spirantized", "nasalized", and "geminated" or "unaltered" forms of intervocalic consonants (cf. Sapir 1913-15, 1930-31; Voegelin, Voegelin, & Hale, 1962, etc.), which must have been present in some form even in pre-Proto-Uto-Aztecan (cf. Langacker, 1975). However, all clusters which reflect these alternations are invariably homorganic at the phonetic level. There is no homorganicity requirement on the consonant clusters in Hopi like those being discussed, or in examples like the Luiseño one given in footnote 13.

12 An example from Luiseño shows exactly the same sort of development: consider tūka-t 'wildcat' and tūkwut 'mountain lion'. tūkat contains a root tūka- and an absolutive suffix -t; in tūkwut the augmentative suffix -wu- intervenes between these two morphemes. The derivation of tūkwut must be assumed to have followed the following steps:

\[
\begin{align*}
\text{tūka-wu-t} & \quad \text{Stress} \\
\text{tūkwut} & \quad \text{Syncope} \\
\text{tūkwut} & \quad \text{Pre-cluster vowel shortening}
\end{align*}
\]

This suggests that the historical processes which produce the CVCC- words in modern Hopi must include a rule shortening long vowels in closed syllables.

13 Some Papago and Pima words with first-syllable diphthongs differ as to which vowel of the diphthong is stressed and which is glided (Kenneth Hale originally pointed this out to me). The first vowel is always stressed in Papago, but in Pima the second vowel is stressed if it is the same height as or lower than the first vowel: Papago múa/Pima muá 'kill'. It is possible that the Pima rule reflects some relic of an earlier second-mora stress rule.
Ronald Langacker has told me that speakers of some modern Aztec dialects which he observed in Mexico in early 1976 seemed to have a tendency to let stress wander from the penult under certain circumstances—for instance, stress might be attracted to some long vowels. It still seems to be the case that Aztec stress is generally penultimate, however.

The restriction that high tone in Northern Tepehuan may only occur on the first or second syllable of the word should also be mentioned here. If Woo's (1970) synchronic analysis of Northern Tepehuan tone actually reflects a historical development from an earlier stress system, as I have claimed above, then perhaps Northern Tepehuan does provide evidence that some Uto-Aztecan language at one stage may have had a regular second-syllable stress rule (this does not contradict the claim in the previous paragraph of the text, but should be evaluated along with it).

The reconstruction of these pronominal elements, and their uses, has been described by Steele (e.g., 1976) and Langacker (e.g., 1976). The discussion in the text seems to provide some indication of the time depth at which at least some prefixation must be reconstructed for Proto-Uto-Aztecan.

In connection with this argument, see also the discussion by Hyman (1975:21).

If the proto-language indeed had a rule moving the stress to the first vowel of a two-syllable word, we can in fact more easily suggest mechanism for the development of penultimate stress in languages like Aztec. A very large number of the words in most Uto-Aztecan languages, and undoubtedly in the proto-language also, have either two or three syllables underlyingly (CVCV or CVCVCV). By a modified second-mora stress rule, with the addition just described, such words would be stressed CVCV and CVCVCV. This data is consistent with the second-mora plus special bisyllabic stress rules, but it also sounds exactly like penultimate stress. It is easy, therefore, to imagine this pattern of penultimate stress for two- and three-syllable words being generalized to longer words as well (resulting in a simplification of the grammar, with two stress rules being replaced by one). What is surprising, in fact, is that so natural a development seems not to have happened more often.
REFERENCES


Casad, Eugene H., and Santos de Jesús. 1976. The Coyote ate the Rabbit: T'uí nu?u t'wa'avingo t'ácu?u wa-t'-áh-k?/u-t'-áh-k?/u. Ms.


Kalectaca, Milo, et al. 1975. Hopi lessons prepared at the University of California, San Diego, for use in classes at California State University, San Diego. Ms.


STRESS IN HINDI

Manjari Ohala

California State University, San Jose

Introduction.

Word stress is a rather controversial area in Hindi phonology, not only because there are competing algorithms for assigning stress to words but also because there are some who doubt that stress even exists in Hindi. And among those who think stress is present, few have paid attention to its phonetic correlates which makes it difficult to verify the stress assignment rules that have been offered. Those who do think Hindi has stress all agree that it is far weaker than in English, perhaps because unstressed syllables are not detectably reduced as they are in English. Also, unlike English, stress plays a very marginal role in Hindi: no words are differentiated solely by stress (however, see the discussion below). This does not mean that there is complete anarchy with regard to stress on Hindi words: stress can be used contrastively and certainly for the vast majority of words it seems that there is only one syllable that can be made prominent, and even in a few three (or more) syllable words where the placement of accent is doubtful there is at least one syllable where the accent may not be placed.

In any case it is of general linguistic interest to investigate the rules and the phonetic correlates of word stress in Hindi no matter how "subtle" it may be since:

(a) We can contribute to the discussion of universals of stress patterns.
(b) We can use the rules to improve the teaching of Hindi as a second language.
(c) We can make use of the rules in machine synthesis of Hindi.

Although a large number of studies dealing with stress in Hindi have appeared in the past 100 years (Kellog 1875, Grierson 1895, Qadri 1930, Bailey 1933, Scholberg 1940, Firth 1944, Rudin 1958, Arun 1961, Dixit 1963, Mehrotra 1965, Dimshita 1966, Tiwari 1966, Ray 1966, Kelkar 1968, Sharma 1969, Jones 1971) there is still little that is understood about it. Amongst these writers, Grierson, Qadri, Rudin, Dixit, Dimshita, Tiwari, Mehrotra, Ray, Kelkar, Sharma, and Jones give rules for stress assignment. The primary problem with all of these studies is that they are entirely impressionistic: no attempt was made to check the predictions about stress placement with numbers of Hindi speakers nor, in most cases, are any phonetic correlates of stress suggested so that someone else could check the predictions.
In this paper I will first review five of these algorithms proposed for assignment of word stress in Hindi and then present some instrumental phonetic data on selected words for which the stress patterns have been predicted by these writers, thus showing that the probable phonetic correlates are what has been identified as "stress".

**Phonemic stress in Hindi?**

Writers such as Bailey, Mehrotra, Arun and others have claimed that there are a handful of words in Hindi which are distinguished solely by stress. They cite examples such as:

\[(1) \begin{align*}
\text{[bɜha]} & \quad \text{'flowed' (past tense)} \\
\text{[gəla]} & \quad \text{'throat'} \\
\text{[bɜha]} & \quad \text{'cause to flow' (imperative)} \\
\text{[gəla]} & \quad \text{'melt' (imperative)}
\end{align*} \]

Where those in column 1 are claimed to have stress on the first syllable and those in column 2 on the second syllable. However, I would claim that these are homophones as far as lexical stress is concerned: rather it is the peculiar intonation pattern of imperative sentences which may make the words in the second column phonetically different from those in the first.

This is an empirical question, however, and can easily be tested. To do so I tape recorded the following two sentences spoken twice each by a native speaker of Hindi.

\[(2) \begin{align*}
\text{ye hE vska gala} & \quad \text{'This is her throat'} \\
\text{tu hi bərəf gala} & \quad \text{'You melt the ice'}
\end{align*} \]

Since in these sentences the initial word is emphasized, both cases of [gəla] get similar intonational contours, thus eliminating any differences due to the imperative intonation itself. These four instances of [gəla] were copied twice and spliced out and randomized onto a separate tape. These 8 samples were then played back to the same informant who was asked to identify which [gəla] belonged to which of the two sentences. The subject reported that they all sounded alike to her and was unable to differentiate them. This shows that the difference between words such as those in (1) is not due to lexical stress.

**Algorithms proposed for stress assignment in Hindi.**

The five linguists whose stress assignment algorithms I will focus on here are Grierson, Dixit, Mehrotra, Kelkar and Sharma, although I will discuss Grierson's proposals only very briefly and concentrate mostly on the latter four. The rules given by other writers are in general similar to the rules of one or more of these five. Most of the proposed algorithms require reference to the concept of syllable
weight which is a function of vowel length and whether the syllable is open or closed.

The earliest algorithm for stress assignment in Hindi was that given by Grierson. Actually his rules were intended to apply not only to Hindi but to Modern Indo-Aryan languages in general. He claims that "The Indo-Aryan vernaculars closely follow the rules of the Sanskrit stress accent" (139). He is referring to the rule of Sanskrit accentuation used by European scholars; however, many linguists today question whether such a rule ever existed in Sanskrit or in Modern Indo-Aryan. The rule he gives, in brief, amounts to:

assign stress to the penultimate syllable if it is long, if it isn't to keep moving backward till one finds a long syllable. (If no such syllable is found, the accent is thrown back as far as possible but no further than pre-antepenultimate if the word ends in a short vowel and antepenultimate if it ends in a long vowel.)

This is certainly reminiscent of the Latin stress rule.

Dixit and Sharma explicitly recognize two degrees of syllable weight: light (L) and heavy (H) (however, Sharma implicitly recognizes 3 degrees of syllable weight as can be seen from his rule 3). Heavy = VC₁ and VC₀, all others are light (i.e. V is light). Kelkar recognizes three degrees of syllable weight: light, medium and heavy. His light is the same as that of Dixit and Sharma, his medium is V or VC and his heavy is any other syllable type (i.e. VC₂ or VC₁).

Given below are my versions of the stress assignment rules of Kelkar, Sharma and Dixit.

Kelkar:

1) If there is only one maximally heavy syllable, assign stress to it.
2) If there is more than one maximum, the last but one of the maximally heavy syllables gets the stress.

Sharma:

1) If there is only one heavy syllable, it gets the stress.
2) If all the syllables are light, the penultimate is stressed.
3) If there is more than one heavy, then the last V or VC which is followed by a consonant or syllable gets stress (i.e. if the V were word final it would not get stress but if it were followed by a C it would be stressed e.g. VC#; similarly final VCC would be stressed but VC# would not).
Dixit:

1) If all syllables are heavy, or all are light, stress the penultimate.
2) If there are some light syllables and some heavy, stress the penultimate if its heavy. If not, keep moving back one syllable till a heavy is found.  
3) If only one syllable is heavy it gets the stress.

Mehrotra:

1) In two syllable words, stress the second syllable if it is $\tilde{V}C_1$ or $\tilde{V}C_2$, otherwise stress the first syllable.
2) In a three syllable word, if all three vowels are short, stress the first.
3) In words with three open syllables, if the penultimate is long, stress it, if not stress the first vowel.
4) In words where the first two syllables are open and the last is closed: if there is only one long vowel, it is stressed, otherwise the first vowel gets the stress. (Mehrotra allows here for variation of the following sort: if the final syllable has a $\tilde{V}$, then optionally it could be stressed rather than the first syllable.)
5) In all other types of words including those with more than three syllables, stress either a $\tilde{V}$ or a closed syllable, which ever comes first, starting with the penultimate.

In spite of some superficial differences, most of these rules are very similar. For example, Kelkar's and Sharma's first rules are the same as Dixit's third. For words having all syllables of equal weight, Dixit's rule 1, Kelkar's rule 2, and Sharma's rules 2 and 3, would all assign stress in the same way, namely, to the penultimate syllable.

Not surprisingly, these rules give similar results for many word types; a sample of these are listed in appendix A. Nevertheless there are cases where, due either to the rule used or the differing degree of syllable weight recognized, different predictions are given; some examples are given in appendix B.

Phonetic correlates of stress.

Having clarified how various writers' stress algorithms for Hindi differ, it would be of interest to take those cases where their predictions on stress assignment to a particular syllable agree, and see how that syllable differs from other syllables in order to find out what the phonetic correlates of stress in Hindi are. Previous studies of stress in various languages have shown that it can be manifested in a variety of ways: by pitch, duration, intensity, vowel reduction, degree of aspiration, etc. (Lehiste 1970). In searching for the correlates of stress in Hindi it is reasonable to start by looking at duration and pitch as prime candidates since these are the most
salient correlates in some languages. Moreover Kelkar, one of the very few to address the problem of the phonetic correlates of stress in Hindi, has claimed "phonetically the prominence of a tonic or subtonic syllable consists of extra length to the coda nonsyllabics or (in case there is no coda) to the nucleus". This requires some interpretation, however. First, if it was the case that stress was correlated with the heaviest syllable in the word, then it would be pointless to speak about stress, we could just as well talk about short and long syllables. But I think by "extra length to the coda" Kelkar means that the same coda or nucleus would be longer when the syllable containing it gets word stress than when it occurred in a syllable which didn't. Kelkar apparently did not actually make any measurements to test this.

To test this claim, stress was located on 36 words according to Kelkar's rules (and for most of these words stress would be assigned similarly by the other's rules as well). The words were spoken by a native speaker of Hindi in isolation, and also in the frame: aapne ______ kaha 'you said ______'. The list was read once backwards and once forwards. The utterances were taped on a high quality tape recorder and were processed by a pitch and amplitude meter. The processed signals were written out on paper by an oscillograph. The coda was measured by hand. (The words had been chosen so that the coda or nucleus would be acoustically distinct from their environment and could be easily measured.)

The results did not reveal any consistent increase in duration for the tauto-syllable consonant after stressed vowels. For example the [t] in [máṭā] 'meaning' averaged 125 ms. whereas that in [prátyék] 'each' was 177 ms. (both uttered in isolation), just the opposite of what would be predicted by Kelkar. Nor was any consistent increase found in the duration of the vowel which would be considered to be 'stressed' according to Kelkar's rules.

Of course one might object that such measurements do not take into account how duration would be influenced by position in the word and sentences, how many syllables there are in the word, the nature of adjacent segments, etc. This is true, but if all these factors were controlled for we would not be able to carry out the measurements since some of these factors determine where stress falls. If we kept them constant then we couldn't vary stress.

The next step was to see if pitch was a correlate of stress. Some 38 words were used. Although a total of 4 informants were used, not all of the words were said by all of the informants, thus the analysis for many words is based on only one informant. Again all the informants were native speakers of Standard Hindi from Uttar Pradesh. The words were put in a context designed to give the word sentence stress. For example, for the word [tambaku] 'tobacco' the informant would be asked to read the following dialogue:
mera nökər tambaku pita ḥə. 'My servant smokes tobacco'
kya pita ḥə? 'What does he smoke?'
təmbaku pita ḥə. 'He smokes tobacco'

The utterances were taped on a high quality tape recorder. For three of the informants the data were processed by a pitch and amplitude meter and written out on paper by an oscillomink. For the fourth informant narrow band spectrograms were made.

The data suggest that in general the syllable which was identified by the above-mentioned writers as the stressed, i.e., the prominent syllable, has a rising pitch on it and a falling pitch on the syllable following it, e.g. [dərvəza] 'door' as given in Figure 1. Sometimes this "stressed" syllable did not have a rising pitch on it but the following syllable did have a falling pitch, e.g. [pətələ] 'last one' given in Figure 2. Also sometimes if the syllable identified as the stressed one was short it occasionally only had high pitch on it, i.e. higher than immediately surrounding syllables. There were a number of words which had ambiguous pitch patterns, e.g. [əməsta] 'a proper name',7 as can be seen in Figure 3. One problem encountered was how to decide when a high falling pitch contour was due to the intended accent and when it was due to the effects of a preceding voiceless stop. Nevertheless, it is possible to offer the generalization that what has been identified as the 'stressed' syllable in a polysyllabic Hindi word is generally the syllable which has a high or rising pitch on it with a falling or low pitch on the immediately following syllable.

More research on this question is needed, however. This paper is offered as only a first step towards finding out the nature of stress in Hindi. To verify that pitch is the primary cue, or at least an important cue for stress in Hindi, tests using synthetic speech are required wherein Hindi subjects can judge the acceptability of synthesized Hindi words where the pitch pattern can be varied without any variation in the other phonetic characteristics of the word.

Acknowledgements: I am extremely grateful to John Ohala for his helpful comments on this paper as well as for making available to me the facilities of the Phonology Lab, University of California, Berkeley. I would also like to thank my subjects for their patience.

Footnotes

1 For example Scholberg says "The author is not prepared to say it does not exist, but he is more and more convinced that what there is is largely of foreign origin." (1962: 14). Also Jules Bloch (1930) thinks that New Indo-Aryan in general has no accent.
Figure 1. Parameters from top to bottom: fundamental frequency, rectified and integrated microphone signal, (unprocessed) microphone signal, 0.1 sec timing pulse. These data show that the second syllable of [dərvaza] "door" has a rising pitch and the syllable following a falling pitch.
Figure 2. Parameters from top to bottom: rectified and integrated microphone signal, fundamental frequency, (unprocessed) microphone signal, 0.1 sec timing pulse. These data show the second syllable of [plčʰlə] "last one" has a falling pitch.
Figure 3. Tracing of the fundamental frequency curve from a narrow-band spectrogram of the word [amīta] (proper name). Here either the first or the third syllable of the word could be said to have high pitch.
Another problem is that with few exceptions, those writers offering stress rules ignore whatever was done by others in this area. The result is an unnecessary proliferation in the literature of differently phrased but functionally equivalent rules.

Grierson seems to equate long and short syllables with long and short vowels.

Kelkar also gives rules for locating syllable boundaries in medial consonant clusters, but these need not concern us here.

This rule of Dixit's is similar to that given by Grierson.

Mehrotra doesn't really give an algorithm but gives a list of words exemplifying different syllable structure and lists where the stress would go for each type. I am here giving the rules I extrapolated from his lists.

On this word Mehrotra's rules assign stress to the first syllable, and those of the others on the last syllable.

Appendix A

Examples of words where the algorithms of Kelkar, Sharma, Dixit and Mehrotra give the same results.

1. páthára 'stone'
2. pátíla 'stony'
3. gréstári 'arrest'
4. pakxstání 'Pakistani'
5. dvmánzíla 'two storied'
6. mêsargly 'suffix'
7. ádálet 'court'
8. kímét 'price'
9. kímí 'expensive'
10. púrvej 'ancestors'
11. beboš 'unconscious'
12. rášvat 'bribe'
13. áta 'flour'
14. kúrsi 'chair'
15. můskal 'difficulty'
16. tábáku 'tobacco'
17. mámla 'affair'
18. avásýak 'necessary'
19. gusélkána 'bathroom'
20. kêtárba 'melon'
21. patámpra 'tradition'
22. dýasalá 'matchstick'
23. múláb 'meaning'
24. ád rak 'ginger'
25. rásol 'kitchen'
26. késárya 'saffron colored'
27. mésányst 'humanity'
28. támáša 'exhibition'
29. čápási 'peon'
30. děrzá 'door'
31. spáhi 'policeman'
32. körkí 'window'
33. lákar há 'wood cutter'
34. bárhyá 'old woman'
35. kóbáfr 'junk collector'
36. hálvál 'sweet seller'
37. tálya 'towel'
38. dixál 'seen'
39. hészyst 'position'
Appendix B

Examples of words where the algorithms of Mehrotra, Dixit, Kelkar and Sharma assign stress differently.

1. Kelkar: k'ubsurat  'pretty' others: k'ubsurat
2. Kelkar: rezgari  'small change' others: rezgari
3. Kelkar: rajniti  'politics' others: rajniti
4. Kelkar: b'artly(e)  'Indian' others: b'artly(e)
5. Mehrotra kólom  'pen' others: kólom
6. Mehrotra 'zmxta  'proper name' others: 'zmxta
7. Mehrotra gorzrya  'grass cutter' others: gorzrya
8. Mehrotra 'amzya  'raw mango' others: 'amzya
9. Mehrotra kála  'art' others: kála
10. Sharma: aršivád  'blessing' others: aršivád
11. Sharma: ztzhaskár  'historian' others: ztzhaskár
12. Dixit: pakístan  'Pakistan' others: pakístan
13. Dixit: kélákár  'artist' others: kélákár
14. Dixit: kémzor  'weak' others: kémzor
15. Dixit: áktrój  'walnut' others: áktrój
16. Dixit: púčkar  'cooling sound' others: púčkar
17. Dixit: góbár  'cow dung' others: góbár
18. Kelkar/Dixit rozgár  'job' Sharma: rozgár
19. Kelkar/Sharma břeštačár  'bad behavior' Dixit/Mehrotra břeštačár
20. Kelkar/Mehrotra ándolan  'movement' Dixit/Sharma ándolán
21. Sharma tehsildár  'govt. official' Dixit/Mehrotra tehsildár

Bibliography

Bloch J. 1930. Introduction to Qadri's Hindustani Phonetics.
Firth, J. R. 1944. Introduction in Harley's Colloquial Hindustani.


