

ISSUES IN THE PHONOLOGY OF PROMINENCE

by

Scott Meredith

B.S., Georgetown University
(1983)M.S., The George Washington University
(1985)Submitted to the Department of
Linguistics and Philosophy
in Partial Fulfillment of
the Requirements of the Degree of

DOCTOR OF PHILOSOPHY

at the

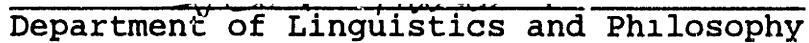
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May, 1990


© Scott Meredith, 1990

The author hereby grants to MIT permission to reproduce and
to distribute copies of this thesis document in whole or in
part.

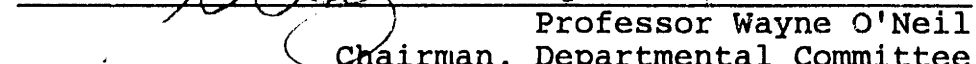
Signature of Author


Department of Linguistics and Philosophy
May 7, 1990

Certified by


Professor Morris Halle
Thesis Supervisor

Accepted by


Professor Wayne O'Neil
Chairman, Departmental CommitteeMASSACHUSETTS INSTITUTE
OF TECHNOLOGY

JUN 12 1990

LIBRARIES

ISSUES IN THE PHONOLOGY OF PROMINENCE

by

SCOTT MEREDITH

Submitted to the Department of Linguistics and Philosophy
on May 7, 1990 in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy in Linguistics

ABSTRACT

A theory of the interaction of phonology and phonetics in the prosodic domains of fundamental frequency, duration, and intensity is developed. Metrical constituent structure in the framework of Halle and Vergnaud 1987 is shown to be the fundamental representation of information for these prosodic domains, integrating phonological categories of tone and syllable weight with phonetic parameters of pitch, duration, and intensity.

The theory is illustrated by reference to three detailed case studies in the prosody of Tibetan, Beijing Mandarin Chinese, and English. For Tibetan, it is shown that the status of syllable nuclei in abstract metrical structure correlates with complexity of surface tonal realization. For Beijing Mandarin Chinese, it is shown that abstract metrical structure describing the location of stress is partially dependent on lexical tone quality of syllables. For English, it is argued that tonological primitives are not interposed between abstract metrical structure and surface fundamental frequency generation.

Acknowledgements

I would like to thank Morris Halle, without whom I would not have come to MIT, stayed at MIT, or benefited greatly from my study at MIT. The debt I owe to Morris cannot be summarized here, and can never be repaid, so I will content myself with this statement of profound gratitude.

Thanks to Ken Hale for serving on my committee, for providing endless obscure references, and for many subtle hints that always concealed deep insight, when I was subtle enough to pay attention.

Thanks to Jim Harris for serving on my committee, and for never missing a chance to explore a potential weakness. Though many weaknesses doubtless remain in this treatment, they cannot be laid at Jim's door.

Thanks to Francois Dell, for serving on my committee, and for bringing many very significant points to my attention.

Thanks to Stefanie Shattuck-Hufnagel, for serving on my committee and for strong and interesting challenges graciously presented.

Thanks to Noam Chomsky, for setting an astonishing example of endeavor and excellence at what must be the limits of intellectual and moral capacity.

Thanks to my class: San Duanmu, Michael Hegarty, Lori Holmes, Kate Kearns, Peter Kipka, Itziar Laka, and Anoop Mahajan. Though I am thanking them as a group, I regard each one of them as an outstanding individual and as a life-long friend.

Thanks to Dennis Klatt, for teaching me phonetics and challenging me, by example, always to work harder and look deeper.

Thanks to my two outstanding phonology professors, who got me some distance toward seeing what it is all about: Donca Steriade and Michael Kenstowicz.

Thanks to my two good friends at the AI Lab: Michael Kashket and Eric Ristad.

Thanks to Wayne O'Neil and Sylvain Bromberger, for encouragement and services too numerous to mention.

Thanks to my teachers: Robert Berwick, James Higginbotham, Richard Kayne, Richard Larson, and Ken Stevens.

Thanks to Sharon Manuel at the Speech Lab, for all kinds of phonetic education and friendship.

Thanks to all my linguistic schoolmates: Zhiming Bao, Eulalia Bonet, Lisa Cheng, Christopher Collins, Hamida Demirdache, Mika Hoffman, Sabine Iatridou, Bill Idsardi, Peter Ihionu, Betsy Klipple, Utpal Lahiri, Paul Law, Harry Leder, Yafei Li, Kelly Sloan, and Chris Tancredi. Knowing you has been a great privilege.

For Tibetan: thanks especially to Ngawang Jordan, for serving as my informant and for teaching me Tibetan and much besides. Thanks also to John Dunne, Masatoshi Nagatomi and Moria Yip.

For Chinese: thanks to Zhiming Bao, Lillian Chen, Francois Dell, San Duanmu, and Yunhui Zhao.

For English: thanks to Noam Chomsky, Irene Heim, Alec Marantz, Chris Tancredi, and William Poser.

Thanks to Dave Boor and Beverly Knystautas.

Thanks to IBM corporation for summer funding.

Thanks to Shane Snowdon for being there.

Thanks to my parents for everything.

This thesis is dedicated to Shao-Ti and Shao-Ling, with gratitude for their love and support.

Table of Contents

| | |
|---|----|
| Abstract | 2 |
| Acknowledgements | 3 |
| Chapter 1: Tone, Stress, and Intonation | 8 |
| 1.1 Introduction: The Sources of Phonological Prominence | 8 |
| 1.2 Universals of Phonetic Interpretation of Metrical Constituent Structure | 13 |
| 1.2.1 Fundamental Frequency | 13 |
| 1.2.2 Duration and Length | 22 |
| 1.2.3 Gesture | 27 |
| 1.3 Overview of the Case Studies | 31 |
| 1.3.1 Tibetan | 31 |
| 1.3.2 Beijing Mandarin Chinese | 36 |
| 1.3.3 English Intonation | 38 |
| 1.4 Summary | 40 |
| Chapter 2: Tibetan Tone and Stress | 42 |
| 2.1 Introduction | 42 |
| 2.2 RST Nouns | 44 |
| 2.2.1 The Tonal Inventory of RST Monosyllabic Nouns | 44 |
| 2.2.1.1 Syllable Coda Quality | 44 |
| 2.2.2.2 Glottal Stop Coda | 47 |
| 2.2.1.3 Surface Tone Properties | 48 |
| 2.2.1.4 Distribution of Tonal Shapes | 49 |
| 2.2.1.4.1 High Register Tone Shapes | 51 |
| 2.2.1.4.1.1 The Level Class | 52 |
| 2.2.1.4.1.2 The Falling Class | 53 |
| 2.2.1.4.2 Low Register Tone Shapes | 53 |
| 2.2.1.5 Phonological Form of Tone Specifications | 54 |
| 2.2.2 Tone Patterns of Compound Nouns | 57 |
| 2.2.2.1 Disyllabic Compounds | 57 |
| 2.2.2.1.1 Tonal and Segmental Effects of Compounding | 58 |
| 2.2.2.1.2 A Restriction on Tonal Output | 61 |
| 2.2.2.1.3 Metrical Dependencies (MS1) | 63 |
| 2.2.2.1.4 Disyllabic Compound Derivations | 65 |
| 2.2.2.2 Trisyllabic Compounds | 71 |
| 2.2.2.2.1 Morphological Types | 71 |
| 2.2.2.2.2 Pitch Interpolation | 72 |
| 2.2.2.2.3 Trisyllabic Compound Derivations | 74 |
| 2.2.2.2.4 Rise Insertion | 79 |
| 2.2.2.3 Quadrisyllabic Nouns | 81 |
| 2.2.2.3.1 Quadrisyllabic Noun Derivations | 81 |
| 2.2.3 Stress Patterns of Nouns | 84 |
| 2.2.3.1 Relation of Stress Metrics to Abstract Tonal Metrics | 85 |
| 2.2.3.2 Basic Units of Stressability | 86 |
| 2.2.3.3 Parameters of Noun Stress (MS2) | 89 |
| 2.2.3.4 Noun Stress Derivations | 90 |

| | | |
|---|--|-----|
| 2.2.4 | Head-shifting Determiners in NP Syntax | 95 |
| 2.3.1 | Verbal Complex Tone | 105 |
| 2.3.1.1 | The Verb Stem | 105 |
| 2.3.1.2 | Metrical Dependencies | 106 |
| 2.3.1.3 | Verb Stem Alternations: Level/Fall | 108 |
| 2.3.1.4 | Verbal Complex Tone Derivations | 110 |
| 2.3.1.5 | Classification of Affixes | 113 |
| 2.3.1.6 | The Honorific Auxiliary Verb [nan] | 114 |
| 2.3.1.7 | Negation in the Verbal Complex | 122 |
| 2.3.2 | Verbal Complex Stress | 126 |
| 2.3.2.1 | Verbal Complex Stress Parameters | 126 |
| | | |
| Chapter 3: Lexical and Phrasal Stress in Beijing | | |
| | Mandarin | 132 |
| 3.1 | Introduction | 132 |
| 3.2 | The Tone/Stress Correlation | 133 |
| 3.2.1 | The Tones of Beijing Mandarin | 133 |
| 3.2.2 | The Relative Strength of Tones | 134 |
| 3.2.3 | Formalization of Stress | 136 |
| 3.2.4 | Derivations | 138 |
| 3.3 | Alternative Approaches | 144 |
| 3.3.1 | Universal Phonetics | 145 |
| 3.3.2 | Independence of Stress and Tone ? | 147 |
| 3.4 | Compound Noun Stress | 149 |
| 3.4.1 | Compound Noun Stress Parameters | 150 |
| 3.4.2 | NP Derivations | 151 |
| 3.4.3 | Four-syllable Compounds and Stress Clash | 154 |
| 3.5 | An Alternative Treatment (Yip 1989) | 165 |
| | | |
| Chapter 4: On the Representation of Intonation in | | |
| | Linguistic Theory | 171 |
| 4.1 | Introduction | 171 |
| 4.1.1 | The Focus Model | 171 |
| 4.1.1.1 | Focus Structure, Accent Placement, and Interpretation | 172 |
| 4.1.1.2 | A Focus Structure Derivation | 175 |
| 4.1.1.3 | The Formulation of Accent Placement | 179 |
| 4.2 | The Intonation Model | 184 |
| 4.2.1 | Laryngeal Generation | 184 |
| 4.2.2 | Relation to Metrical Structure | 186 |
| 4.3 | Issues of Duration | 197 |
| 4.4 | Comparison with a Discrete-level Theory | 200 |
| 4.4.1 | Overview of Pierrehumbert 1980 | 201 |
| 4.4.2 | The Tonal Inventory of English | 203 |
| 4.4.2.1 | PF-external Considerations | 203 |
| 4.4.2.1.1 | Syntax | 203 |
| 4.4.2.1.2 | Semantics | 204 |
| 4.4.2.2 | PF-internal Considerations | 209 |
| 4.4.2.2.1 | The Basic H/L Distinction | 209 |
| 4.4.2.2.2 | Differences in Interpolation | 216 |
| 4.4.2.2.3 | Downstep | 218 |
| 4.4.2.2.4 | Impossible Contours | 220 |
| 4.5 | Cross-dialect Intonation | 223 |
| 4.6 | Intonation and Gesture | 226 |

| | |
|--|-----|
| 4.7 Intonation and Linguistic Theory | 229 |
| Appendix 1: Metrical Constituent Structure | 232 |
| 1. The elements of metrical structure | 232 |
| 2. Metrical constituents | 233 |
| 3. Well-formedness statements for grid lines | 235 |
| 4. Derivational Procedure | 239 |
| 5. Sample Derivation | 240 |
| References | 241 |

Chapter 1

Tone, Stress, and Intonation

1.1 Introduction: The Sources of Phonological Prominence

Tone, stress and intonation act in many languages as phonetic manifestations of abstract prominence relations. Though these abstract prominence relations may have a number of distinct sources, they must be interpreted in speech output. A simple example of a relation of relative prominence among elements of an utterance is provided by the syntactic/semantic phenomenon of focus or emphasis. Consider the following dialog:

- (1) a. John: What did you buy?
 b. Bill: I bought [+focus a dog].

In (1), we observe that 'a dog' in Bill's answer is labelled [+focus]. This labelling corresponds to the fairly direct intuition that the content word that 'fills in' the question word in a dialog such as (1) is the most important element in a response such as Bill's above¹. The remainder of Bill's response, 'I bought', is felt to be understood, given the dialog situation, and may even be omitted. Here, then, is an apparently simple example of a prominence relation within an utterance. We observe now that 'dog' in (1b) is also felt to be the phonologically most prominent element in (1b). Though

¹ Cf. Rochemont 1986 for a survey of relevant claims.

the exact phonetic nature of this phonological quality is hard to pin down (and part of the research discussed in the following chapters is an attempt to untangle that question), most native speakers can agree that it would be somewhat odd, given John's question, to produce an extremely high or low pitch, say, on the word 'I', coupled with a relaxed or neutral pitch on 'a dog'. Or, taking a different phonetic parameter, it would be strange to pronounce 'bought' in a very loud voice, relative to 'I' and 'a dog', again in the given context. We see then that prominence relations are somehow reflected in the phonetics of an utterance.

The focus prominence exemplified above is an example of a prominence relation that originates outside the phonology, though it is respected by the phonology and phonetically interpreted. Focus prominence ultimately derives from discourse properties of utterances, mediated by the syntax, and must be accounted for by a theory of the information structure of discourse. This property of focus will be discussed in more detail below and in chapter 4. But discourse relations and utterance context are not the only sources of abstract prominence relations. Consider English word stress, indicated in the list below by an asterisk above the prominent syllable:

(2)

*
elephant

*
solution

*
emptiness

*
mirage

Although the exact phonetic parameters that are activated in the production of the prominence relation of English word stress are still disputed (this question will be taken up to some extent in chapter 4, also cf. Beckman 1986 for a survey of experimental evidence), any native speaker of English will agree that there is some phonetic quality in the pronunciations of (2) that corresponds to the placement of the asterisks. Thus, the asterisks in (2) are markers of abstract prominence relations among syllables. This kind of prominence is different from the focus-related prominence exemplified in (1). There is nothing in the discourse situation that requires the prominences indicated, nor is there any syntactic basis for them, given that the items in (2) represent X^0 projections in syntax. English word stress is an abstract prominence relation among syllables determined in the lexicon by principles of morphology and phonology.

Nevertheless, to take one final example of a prominence relation, relative metrical strength of syllables within a word is not necessarily the exclusive province of lexical morphological and phonological principles. Discourse determined prominence can interfere, as in the following 'correction' utterance:

(3) I didn't say ^{*}employ^{*}EE, I said employER.

The lexical stress is indicated by an asterisk over the strong syllable, as in (2) above. The actually realized prominence is on the capitalized syllables, however. Here the relative metrical strength of syllables is determined by discourse situation, a response to the hearer's apparent misunderstanding.

We have seen that prominence relations can originate at different levels of the grammar, but ultimately they must be phonetically interpreted. This research starts from the premise that at some point prior to phonetic production, all abstract prominence relations, whether contributed by discourse, syntax, morphology, or phonology, must be integrated in one or more planes of **metrical constituent structure**, the basic elements of which are selected from the units of phonological representation of words in the lexicon, e.g. syllables, morae, or segments. The formal properties of metrical constituent structure are developed exhaustively in Halle and Vergnaud 1987, and are summarized in appendix 1. A very simple example of a prominence

relation that is easily captured by this system is French word-final stress. If we build a grid with an asterisk above each syllable of a French word, mark a constituent that spans all asterisks so placed, and select the rightmost asterisk for augmentation on the next higher grid line (thus, the head of the line 0 constituent), we can derive the following representation for stress in a typical French word:

```

          *   line 1
(* * * * * *) line 0
originalité

```

The formal statements of possibilities of constituent construction on a grid line and head selection for constituents are severely restricted in this system, and the reader is referred to appendix 1 for a more comprehensive treatment. Representations built in accordance with the theory developed in Halle and Vergnaud 1987 will be taken as the abstract unifying mechanism at PF for all the prominence relations (deriving from a number of distinct sources) that surface in phonetic interpretation.

The case studies presented in the following chapters are concerned with certain issues that arise under this conception of abstract prominence and its relation both to other components of the grammar and to non-grammatical motor

output systems. Before summarizing some of the particular questions and results presented in the following chapters, it is useful to get a broad overview of the relation between abstract metrical structure in the phonology and the possibilities for its surface phonetic realization.

1.2 Universals of Phonetic Interpretation of Metrical Constituent Structure

1.2.1 Fundamental Frequency

In sketching an account of the possible interactions between abstract phonological representations of relative prominence, on the one hand, and surface phonetic specification of prominence on the other, we will initially consider the phenomenon of pitch or tone, and then generalize to other phonetic qualities. It will be assumed that there exists at the level of universal phonetics an abstract component of pitch control. To minimize unproductive terminological controversy, we will use the term 'F0 generator' (FG) for this component. Although results from research in this area are far from conclusive, we will assume that FG makes use of subglottal pressure, vocalis muscle tension, cricothyroid cartilage positioning, and other physiological mechanisms to achieve control of fundamental frequency (cf. Stevens and Hirano 1981, Fink and Demarest 1978, Lieberman 1966, and Fromkin 1978, among many

others). The physiological details are not especially important for this abstract consideration.

We are concerned with FG as a 'black box', that is, with its inputs and outputs. The output of FG we have specified as the surface F0 contour of a word, phrase, or sentence. Several elements from the phonology could conceivably be relevant as inputs. One obvious candidate is lexically specified tone. Thus, for example, in Beijing Mandarin Chinese the surface F0 shape of a syllable rime is obviously influenced by the quality of the tone associated with that syllable in the phonology (lexicon). Therefore, lexical tones are a possible source of determination (input) for FG.

In addition, we find that in many languages, including, continuing our example, Beijing Mandarin Chinese, syllables that are specified as more prominent in the derivation (say by a discourse process such as focus or emphasis, introduced above) are realized with a markedly greater pitch range than corresponding non-prominent syllables, all else being equal. It is therefore clear, at least in some pre-theoretic sense, that focus or emphasis specifications should also be accepted as inputs to the general mechanism of FG, since they originate 'earlier' (in the model of grammar) than the phonetics (FG), and have an effect on surface F0 contour².

² In using the term 'earlier', we are assuming a model of phonology consisting of ordered application of rules, with rules of detailed phonetic interpretation ordered last or 'latest' in a derivation. Focus specifications are assumed to be input to the phonology and passed along in some form

Another distinct source of potential variation in F0 contour arises from considerations of speaker attitude, emotion, expressiveness, etc. Although this kind of determination of F0 may appear indistinguishable from the focus or emphasis specifications just introduced, this is not the case. It should be clear that focus labelling is required within the linguistic component or module, because it has a direct and simultaneous effect on logical interpretation and syntactic form, as well as on phonological prominence (cf. Rochemont 1986, Horvath 1981, Culicover and Rochemont 1983, and discussion below). Considerations of speaker attitude and emotion, while they must respect the focus-determined properties of the phonological representation, are not constrained by correlations with other language-module factors, and, lacking such correlation, will be seen as determined within a separate module. There will be further discussion below, and especially in chapter 4 of this conceptualization of the organization of the grammar.

For the moment, to summarize, we have isolated three potential abstract inputs to our FG abstract phonetic component: 1) lexically specified pitch (tones); 2) focus representation; and 3) expressive or affective input. With this conception, certain interesting questions immediately arise. For example: Are all three inputs always required for

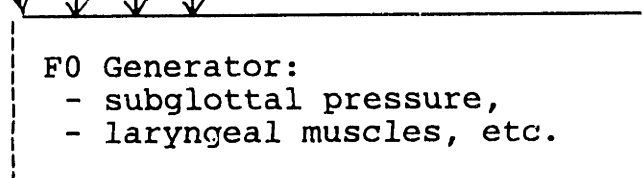
(to be made precise in chapter 4) to processes of phonetic interpretation.

proper generation in FG? What are the relations between the different inputs? Can conflicting specifications arise among them, and, if so, how does FG reconcile them? Are there other abstract phonetic components with similar or identical properties, and what are they, how could they be motivated, and how do they relate to FG? Answers to these questions, however partial or tentative, will be advances in our understanding of phonology, phonetics, and the relation between them.

The three case studies presented in this thesis, on Tibetan, Beijing Mandarin Chinese, and English, are offered as first steps on the road to this understanding. In this introductory discussion, it is appropriate to summarize the broad picture that emerges from the data analysis carried out in the remaining chapters. Consider then, FG and its inputs, as schematized in (4) below:

(4) Phonetic F0 Generator, as a 'black box':

Inputs:

a. Metrical constituent structure
(required)b. Lexical tones
(optional)c. Expressive qualities
(optional)d. Segmental melody
(required)

Output:

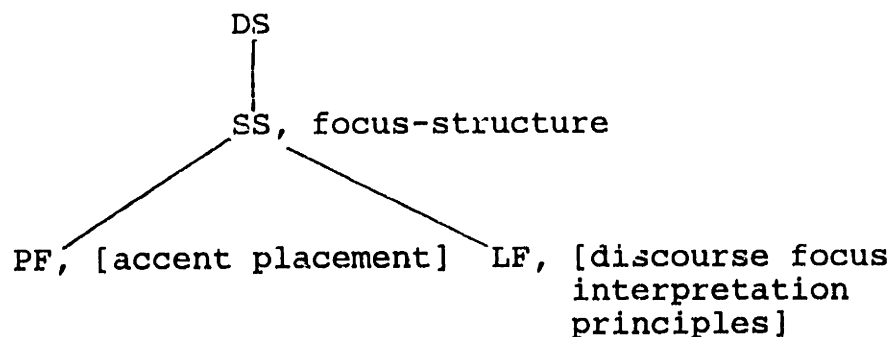
[surface intonation contour, F0]

(4) makes several claims, which will be examined in turn. First, notice that the first input, a., is labelled 'Metrical constituent structure'. This actually includes what was referred to in the discussion above as 'focus representation'. In fact, focus or emphasis relations are only one input to an early phonological process of metrical constituent construction. The metrical structure derived in the phonology is assumed here to be the fundamental organizing factor in F0 generation (and much else besides, as we will see below and in the case studies to follow), and

is therefore a required input to FG. The metrical representation crucially assumed throughout this work is that developed by Halle and Vergnaud 1987, henceforth HV. See appendix 1 for an elementary exposition of the important features of this framework. For purposes of the present general discussion, it is sufficient to note that a phonological process exists which consists of a battery of parameter settings which indicate primitive formational units of metrical constituents, size of constituents, direction of constituent construction across the formational units, degrees of projection, and placement of heads within constituents. This battery of parameter settings, and the representations it constructs, will be referred to in this introductory section henceforth as 'MCC' for 'metrical constituent construction'.

The model of grammar we assume is the standard T-model, with focus annotations of s-structure, as justified by Selkirk 1984 and Rochemont 1986, which is shown below:

(5) Model of the Grammar



Within PF, MCC constructs metrical constituent structure by means of its own battery of associated metrical parameter settings and simultaneous reference to: focus-structure (from the syntax); and segmental / syllabic / morphological properties of lexical items, where relevant (this differs from language to language, detailed examples will be provided in the case studies below).

The output of MCC, then, is input a. in (4) above, the required metrical constituent structure input to the FG component. We claim in this thesis that a formal structure adhering to the properties of MCC is, universally across languages, an input to FG (which itself is taken as a provision of universal phonetics). This claim applies to languages that are not (generally) thought of as tone languages, such as English, to 'syntagmatic' tone languages such as many of the Bantu languages of Africa (cf. Schuh 1978, Sietsema 1989), and to 'paradigmatic' tone languages (cf. Schuh 1978 for origin and discussion of the term) where tones are typically lexically specified for every syllable, such as many tone languages of East Asia (e.g. Thai, Vietnamese, Tibetan, and Mandarin Chinese).

The phonetic interpretation of the MCC relations that are input to FG is a complex question, and this question is treated in detail in the case studies to follow. We can make the initial observation that metrically prominent items tend to exhibit expanded pitch range (eg. higher pitch for

metrically stronger items), cf. Shih 1987 for discussion relevant to Mandarin Chinese.

Returning to our discussion of (4), then, the next input to FG, b., is lexical tones. While MCC is a linguistic universal in phonological organization, lexical tones, it is claimed here, are not. This is an uncontroversial claim, in the sense that a distinction in just these terms is generally recognized between, say, Cantonese and French. However, with the introduction of tonal units as primitives of intonational analysis (cf. Pierrehumbert 1980), the claim that such units need not be specified (in the sense that the presence of this input is a parametric choice among language types), or cannot be specified (if that parameter cannot be demonstrated to be active for a given language) becomes potentially more interesting and controversial. More specifically, the claim in (4) ('optional') is not just that a given utterance in a given language may happen to lack specification for lexical tonal contrast (including here the 'intonational lexicon' of Liberman 1975 and Pierrehumbert 1980), but that languages may entirely fail, at any level, to specify tonal organization in phonological terms, as an input to F0 generation. There will be much more discussion of this distinction in chapter 4, but the general outline of the claim should be clear. It is therefore held here that while Beijing Mandarin Chinese, for example, takes the option that essentially every syllable has an associated

tonal (b.) input to FG, English, on the other, provides no such input to FG, at the level of the mora, rime, syllable, phonological word, phonological phrase, intonational phrase, utterance, or any other level (cf. Selkirk 1984, Nespor and Vogel 1986 for some analysis of post-lexical levels in phonology).

Moving to the expressive input (c.) of FG, we observe that this is optional, in the sense that an utterance may be uttered without affect or implied attitude. Note that even the most mechanical speech, say a sentence read simply as a list of words, entirely lacking affective or expressive intonation, nevertheless has, as required, a metrical representation. In such a metrical representation, every word is a metrical subconstituent, and the metrical head of the utterance as a whole, or greatest prominence, is on the last such subconstituent (an insight covered by the traditional Nuclear Stress Rule, cf. Chomsky and Halle 1968, and Liberman and Pierrehumbert 1984 for some phonetic analysis of list intonation). Expressive considerations may expand the pitch range beyond what is required by the MCC mandatory input (a.). Here we assign the source of the expressive factors to the 'E-faculty' (for emotion or expression), while all other inputs to FG have their origin in the 'L-faculty' or language faculty. We will return to this distinction in discussion below and in chapter 4.

Finally, the segmental input (d.) to FG, is not of great concern here, but it clearly plays a role in processes such as the effect of voicing quality of onset obstruents on pitch, a well-documented phonetic effect with interesting but irrelevant diachronic consequences, cf. Hombert 1978, Ohala 1978, and Bao 1990 for survey and discussion.

1.2.2 Duration and Length

Going beyond the F0 generator, it is not unreasonable to consider whether a similar organization can be motivated for other abstract phonetic components that appear to play a role in the surface expression of abstract prominence relations. A number of studies have shown (cf. Beckman 1986 and references therein) that relative segmental or nuclear duration plays a key role in the phonetic spell-out of abstract prominence in English (lexical stress, in the relevant cases). Suppose this to be the case. Then it may be reasonable to assume the existence of what for consistency we might refer to as a 'Duration Generator'. This again is an abstract phonetic component, with various subtended articulatory mechanisms of the familiar sort, e.g. subglottal pressure, vocal fold mechanisms, oral articulators, etc. This is therefore another 'black box', of the same general character as FG discussed above. The output of this abstract phonetic processor is segmental duration (possibly relevant only for moraic (rime) dominated

segments, but see Kipka 1987 for possible counter-examples). We expect that the inputs might be selected from a set analogous to that specified for FG in (4) above, specifically:

(6) Duration Generator, DG, as a 'black box':

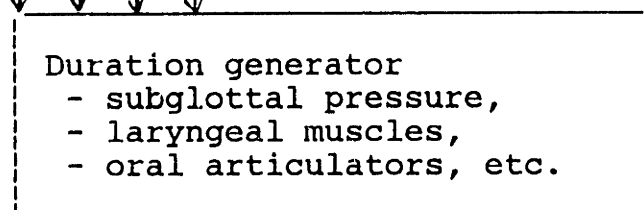
Inputs:

a. Metrical constituent structure
(required)

b. Lexical length, weight
(optional)

c. Expressive qualities
(optional)

d. Segmental melody
(required)



Output:

[surface duration of segments, especially rime]

Based on the results from the studies on the phonetic realization of stress in English (summarized in Beckman 1986) and a number of other languages, which have duration as the primary phonetic exponent of abstract metrical structure, we will make the strong claim that the phonetic

expression of nuclear or rime segmental duration in every language universally refers to a metrical constituent representation, though the surface reflex of this may be subtle in language such as Japanese, which expresses metrical prominence primarily by means of the pitch component, FG. Note, however, that even for Japanese interesting claims have been made for durational contrasts in prominence relations subsidiary to the main accent (the primary accent is expressed by pitch, cf. Yamada 1990 and references therein).

Many languages, such as Japanese and Tibetan, have lexically specified nuclear length distinctions. This input for DG is thus analogous to the (b.) input, lexical tones, to FG, and again, obviously, it is optional on a language-specific basis.

The expressive input is clearly relevant for all languages, and again we will claim that this input is the only one that, crucially, originates outside the L-faculty, just as argued for FG above.

Segmental melody refers to the phonetic substance of the segments that compose the word. This is relevant in that, for example, length may be maintained differently for sonorants as opposed to vocalic segments, and miscellaneous information of that type. This input is not especially interesting for the present discussion.

We see then, that a structure entirely parallel to the FG abstract phonetic component, (4), can be derived for the durational control component, (6).

Furthermore, there is no principled reason why this treatment cannot be extended to an abstract phonetic component of intensity control (loudness). While recent work has tended to deny the relevance of intensity to the phonetic realization of lexical stress (cf. Beckman and references therein), clearly intensity can co-vary with the prominence relations, and can be the carrier for expressive qualities. It has not been shown that intensity is ever a factor in the phonology of any language. Nevertheless, in the general model we are developing the non-metrical phonological input (e.g. tones in (4)) is always optional. Therefore, one might take the position that, for the 'intensity generator', IG, non-metrical phonological input is an option that is not ever realized (and could be disallowed by stipulation, though we see no reason to do so).

A case could further be made that realization of vowel quality, another standard candidate in the search for phonetic exponents of lexical stress, can be modelled along exactly analogous lines. Clearly, where vowel reduction is seen as a consequence of metrical status, as often in English (though see Bolinger 1986 for a contrary view), the

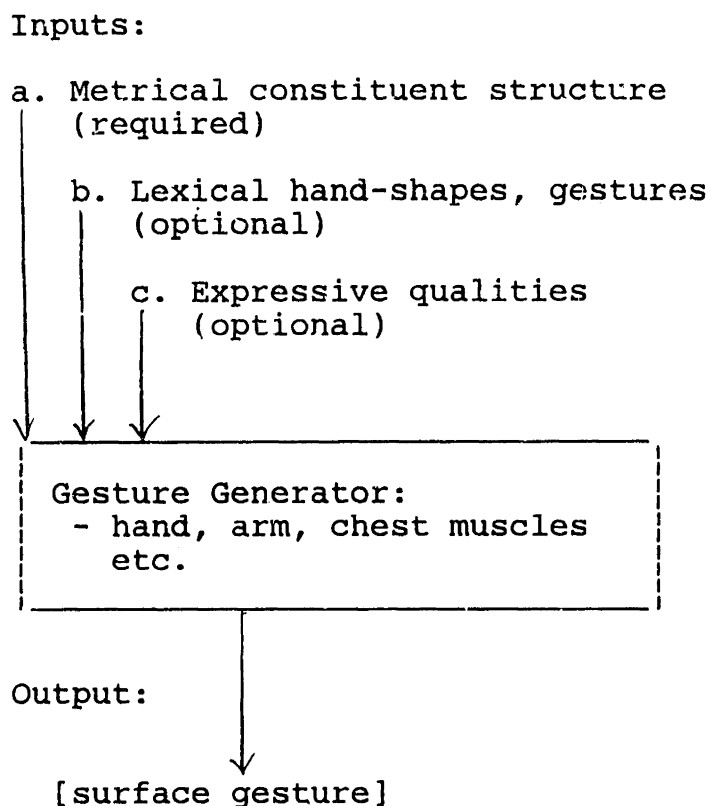
universally required input for all the abstract phonetic components developed so far (FG, DG, IG, and now VG 'vowel quality generator') is present, namely, the standard stress grid (metrical constituent structure) that is simultaneously input to FG and DG. More interestingly, in at least one language, Tiberian Hebrew, the MCC that is input to VG can be shown to be formally distinct from that which is input to the other phonetic stress realization components (presumably at least FG, possibly DG). This is an interesting confirmation of the discrete structure of the phonetic level in the model presented here. Were it the case that the phonetic expression of metrically encoded prominence relations was always a homogenized output where pitch, duration and vowel quality varied uniformly in all languages, the case for discrete phonetic components in the realization of metrical constituent structure, and the subsequently predicted possibility of distinct metrical representations as inputs to those discrete components, would be greatly weakened. We will see in chapter 2 below that Tibetan is another such case, though involving a different, and therefore interesting, partition of phonetic components. See Rappaport 1984 and HV for further discussion of Tiberian Hebrew.

1.2.3 Gesture

FG, DG, IG, and VG are all output systems which see a particular 'end-point', or external interface, of the grammar, in this case PF. In principle, since they lie outside the language model proper, there is no a priori reason to deny them reference to LF, if that can be motivated, though this discussion would take us too far afield. Though they are claimed to lie outside the language faculty in the specific sense discussed in this section, they are closely enough related to the expression of linguistic categories (tone, stress, etc.) that we may have some trouble making a clear distinction in their placement. Let us then consider another 'black box' output system, that shares many of the properties of FG, DG, IG, and VG, but which generally lacks the confusing situation of the essentially invisible laryngeal and oral articulatory mechanisms, which are simultaneously affected by segmental, supra-segmental, metrical and expressive factors.

Gesture will be claimed, in the present model, to instantiate a non-oral/laryngeal abstract output component of essentially the same type motivated in the general model presented so far. The relevant specifications for this universal output component are:

(7) Gesture Generator, as a 'black box':



This generator has essentially the same structure as the generators we saw above. It has the required metrical structure input (though under certain conditions of environmental constraint the effects of this may not always be visible). Obviously, expressive qualities can be relevant to gestural output as well. In its role as supporting system to spoken language, as accompaniment to the speech of a speaker of American English, say, the non-metrical linguistic input (analogous to lexical tones in FG's input) is absent, and is therefore labelled 'optional' in the model, as in all the other diagrams for the various

components introduced above. But if the speaker is employing ASL, for example, the linguistic input is activated. The possibility that this component can function with or without linguistic input, in the most obvious sense, fortifies our general conception of these motor systems, the 'black boxes', as lying outside the L-faculty proper, though they have access to at least one 'end-point' of that faculty.

Clearly, the model presented here raises as many questions as it answers, and as always theory is underdetermined by available evidence. The remaining chapters are devoted to case studies of three languages that shown various concrete instantiations of the possible interactions between elements within a given phonetic component, relations between the phonetic components and their linguistic inputs, and among the phonetic components themselves.

In studies of gesture and speech coordination, it is not uncommon to find the position that the rhythmic structure of both speech and gesture is based on the workings of the same generalized abstract rhythmic component (cf. Condon and Ogston 1967). This view is probably correct, if understood in terms of two discrete properties of the model sketched here: first, the internal processing of all the motor system boxes diagrammed above (on the phonetic side of the line) doubtless has much in common at some level of abstract description, where, as Hadar 1989 puts it:

Reduction of computational load requires that motor programmes should not compute the forces of individual muscles in detail; instead, within a fairly abstract representation of the task (a 'schema' for example, see Schmidt, 1975) the related muscles may be constrained to operate synergistically ... by fixing the relationship between them.

Thus, internal to the motor boxes, it would be surprising if the representational properties of these control specifications differed crucially across motor components, and in this sense the same rhythmic generator may be said to underly both speech and gestural properties.

On the other hand, the representational system of metrical constituent structure is also a mandatory input to all motor systems in the model suggested here, and this is crucially and quintessentially a linguistic structure, originating outside the motor control system. To the extent it encodes linguistic properties that will have rhythmic expression, it is another source of computational parallelism between speech and gesture, but it should not be confused with the lower-level commonality of an abstract 'schema' for muscle control and coordination.

We turn below to brief overviews of the phonological organization of tone, stress and intonation in the three languages considered in the case studies of chapters 2, 3, and 4.

1.3 Overview of the Case Studies

1.3.1 Tibetan

Rappaport 1984 has provided an intriguing analysis of Tiberian Hebrew, showing that distinct orthogonal metrical planes must be used to represent the phonological phenomena of stress and vowel reduction in this language. Commenting on the general significance of Rappaport's findings, Halle and Vergnaud 1987 observe that:

- (1) ... more than one metrical constituent structure may be associated with a given central line of phonemes,
- (2) ... metrical constituent structure need not always be interpreted in stress terms, and
- (3) ... a special relationship holds between the head of a metrical constituent and the rest of the elements in the constituent.

These conclusions are important, and it would clearly be valuable if another system employing distinct, orthogonal metrical planes could be adduced, both for comparison with the particular phonological/phonetic properties of Tiberian Hebrew and to strengthen our confidence in the general conclusions stated above. While recent studies of certain Bantu languages have tended to confirm point (2) above (cf. Sietsema 1989), these analyses have not provided evidence for the existence of distinct, orthogonal metrical planes in these languages. Let us consider each point listed above in

turn, and briefly summarize how Tibetan strengthens our confidence in the general conclusion and provides interesting new data for further theoretical advances.

First, Tibetan provides the first well-motivated example beyond Tiberian Hebrew that indeed multiple metrical constituent structures can coexist. In Tibetan, one plane of metrical constituent structure (MS1 in discussion below) represents metrical rank among syllables that is relevant for tonal processes. High tone insertion, contour simplification, and tone deletion all depend on a plane of metrical structure. In this respect, Tibetan somewhat resembles, at a very coarse grain of description, the Bantu languages analyzed by Sietsema 1989. Sietsema shows that tonal operations in four Tanzanian Bantu languages, Kimatuumbi, Ci-Ruri, Digo, and Sukuma are dependent on underlying metrical structure, also within the formal framework of Halle and Vergnaud 1987.

In Tibetan, a distinct plane of metrical constituent structure (MS2 in discussion below), is constructed by reference to the adjusted tonal specifications, syllable weight, and independent metrical parameter settings. The phonetic exponent of this plane of abstract metrical structure is what is traditionally termed 'stress'. Just as in English, the actual phonetic implementation of the stress prominence relations among syllables in lexical items is subject to some debated, and will require extensive and

detailed instrumental analysis. Though it is not necessary to take a firm position here (any more than a firm position on the phonetic implementation of English stress is taken in standard analysis of English lexical stress), it is clear that at a minimum, stress is associated with larger pitch range in the realization of syllable tone in Tibetan. In any case, it is very clear that the phenomenon of stress represented by MS2 is entirely distinct from the metrical relations encoded in MS1, which have a precise interpretation in terms of the phonological structure of tone.

Moving to the second of Halle and Vergnaud's points given above, that metrical constituent structure need not always be interpreted in stress terms, it should be clear from the discussion just above that, because MS1 has an interpretation strictly in terms of phonological tone structure, Tibetan provides a dramatic confirmation of the claim that Kappaport advanced, as do the Bantu studies mentioned earlier.

As for the third of Halle and Vergnaud's observations, its application to Tibetan is quite interesting, in that in this language we find direct correlation between the metrical status of a given syllable and its phonological tone specification. The head will undergo high tone insertion under its 'register' (in the sense of Yip 1980) node, a syllable that is not a metrical head (not

represented on line 1, cf. appendix 1 for formal details of the framework) must undergo contour deletion (i.e. any lexically specified contour shapes, such as fall or rise, must not appear on the surface in such syllables), and a syllable which fails to belong to any metrical constituent (by the operation of conflation, to be covered in chapter 2) lacks any phonological specification for tone and is phonetically interpreted by a default process. Thus, we observe that in Tibetan, not only is the metrical status of head respected by phonological processes, other, subsidiary metrical possibilities are directly reflected in the tonology³.

The study of Tibetan reported on in chapter 2 thus represents not only a significant advance in the understanding of a relatively little-studied language, but also a rare instantiation of the complex and highly interesting phenomenon of distinct orthogonal planes of metrical structure.

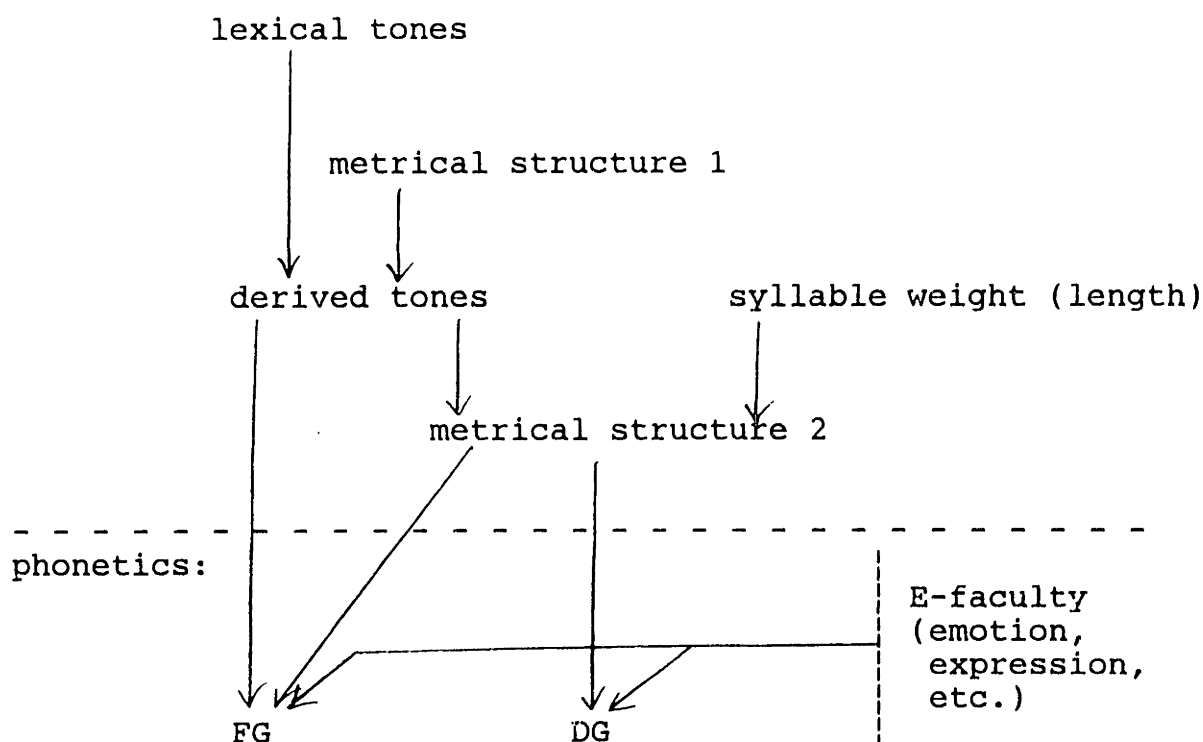
We show below a diagram depicting the various relations between the lexical tones, lexically specified rime length or weight distinctions, and two distinct planes of metrical constituent structure in Tibetan. Detailed development of this picture is presented in chapter 2. The diagram below is intended only as an abstract guide to the phonological and

³ This phenomenon is rather different in quality, of course, from the metrical processes that motivated Halle and Vergnaud's claim (3) for Tiberian Hebrew.

phonetic organization of tonal and prominence relations in this language. An arrow between boxes indicates an input/output relation:

(8) The logical organization of Tibetan tone and stress

phonology:



In (8), FG, with the inputs shown, determines the ultimate surface realization of tone and intonation, while DG, with the inputs shown, determines the surface realization of relative rime duration⁴. The notion 'stress' in Tibetan is a derivative concept that signifies the perceptual effect of the combined outputs of FG and DG (both of which have metrical structure 2 as input, where metrical

⁴ Note that the segmental inputs to FG and DG are assumed but not shown.

structure 2 is the metrical expression of that prominence which most closely fits the pre-theoretic term 'stress' for this language).

1.3.2 Beijing Mandarin Chinese

Many researchers on the Beijing city dialect of Mandarin Chinese have noted the existence of lexical stress in this language (Hockett 1947, Rygaloff 1956, Xu 1960, Kratochvil 1969 1974, Ho 1976, Hoa 1983). While the phonetic exponents of stress in Mandarin are understood to involve expanded pitch range and increased relative duration (Shih 1987), the proper phonological characterization of Beijing Mandarin stress has remained unclear. Thus Hoa 1983, in a pioneering study of phrasal and sentential stress patterns of Beijing Mandarin reflects the general consensus that lexical stress is an arbitrary lexical diacritic, with perhaps a statistical preference for main stress on the last syllable. In Hoa's study then, words emerge from the lexicon with an essentially arbitrary marking for stress, and the stress location so identified is respected and propagated by more regular and principled processes of phrasal and sentential stress assignment, corresponding roughly to the Nuclear Stress Rule of Chomsky and Halle 1968, though cast in a different formal framework.

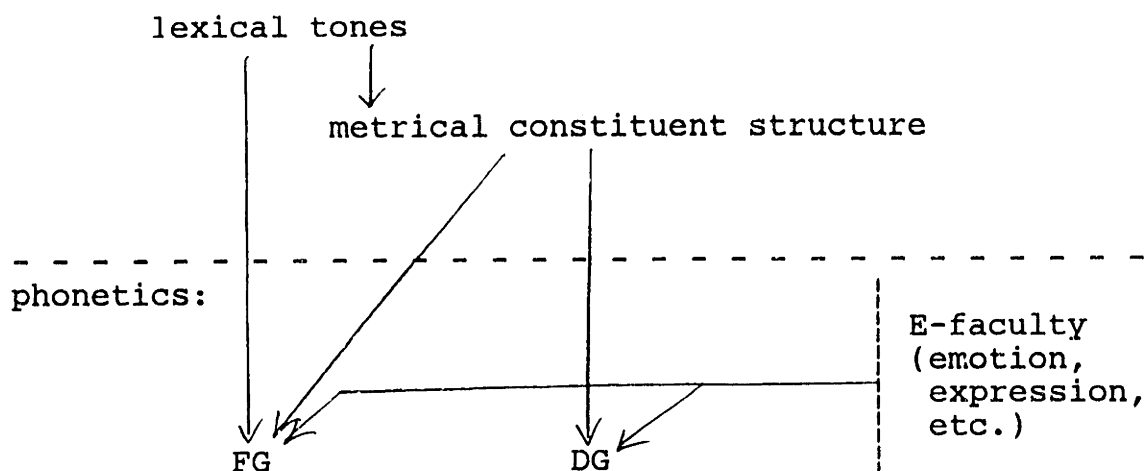
Interesting questions, then, can be raised concerning the status of Beijing Mandarin lexical stress. The most interesting question is whether some underlying regularity could not in fact be uncovered, which would help explain acquisition of stress (though of course it is not a priori unthinkable that stress is a lexical diacritic). In the study undertaken here, it is shown that stress in Beijing Mandarin is fundamentally analogous to stress in many other languages, in that its location is formally characterized by bracketed metrical grids based on parameter settings, in the framework of Halle and Vergnaud 1987. This representation serves to illuminate considerable, and quite interesting, regularity in Beijing Mandarin stress patterns.

The metrical organization of Beijing Mandarin Chinese is somewhat simpler than Tibetan, though rather surprising in light of that fact that while tonal operations and realization have been shown to be dependent on metrical structure in a number of languages (cf. Sietsema 1989), the opposite case, metrical structure built on the basis of lexical tone quality, has been previously demonstrated only in the analysis of Molinos Mixtec (Yip 1981). Thus in Beijing Mandarin, stress is partially determined by lexically assigned tone of syllables, and the tones themselves can be seen as organized along a scale of the metrical strength they induce, with the high falling tone 4

inducing the greatest metrical strength, and the low dipping tone 3 the least.

(9) The logical organization of Beijing Mandarin tone and stress

phonology:



1.3.3 English Intonation

There has been longstanding controversy in the field of English intonation regarding the proper form of phonological description for pitch variation associated with meaning and expression. One popular traditional approach culminating in the synthesis of Liberman 1975 has emphasized the description of contours in terms of a number of discrete tone levels, and four such levels are claimed by Liberman to exist in American English. Pierrehumbert 1980 has taken what appears to be the final step in reduction of the primitive

elements, casting all contrasts in terms of two primitive tone specifications, H (high tone) and L (low tone), with a complex battery of phonetic interpretation rules deriving all sub-tonemic variations.

The analysis presented in chapter 4 takes issue with the use of tonal primitives for description of English intonation, claiming that there is a lack of the expected correlation between these phonological elements and other linguistically motivated terms. To the extent 'intonational meaning' exists, then, it is a phenomenon similar to gestural meaning, lacking discrete characterization in the grammar. Therefore, the use of the phonological feature primitives H and L (abbreviations, presumably, for [+high register] and [-high register] in the sense of Yip 1980 or Bao 1990) is entirely inappropriate and misleading. The use of these primitives for intonational description can be analogized to use of the phonological tongue position feature [back] in description of the oral dynamics of chewing food, an obvious category error. In the area of intonation, the distinction appears to be somewhat more subtle, because intonational meaning exactly analogous to gestural meaning does in fact exist. Chapter 4 attempts to analyze this subtle distinction in some detail. The conclusion presented there is that pitch specification in English is affected only by focus prominence structure, correlated with lexical stress specification, and expressive or affective factors, but none of these inputs to FG is

of theoretical biology. While theoretical constructs may turn out to be mistaken, there is nothing incoherent in such a position. In particular, it appears that a view of the language faculty as a strictly modular system with discrete, restricted visibility to other cognitive systems, a view of phonological features as primitives of lexical organization rather than as arbitrary diacritics for annotation of dynamic systems, and a view of metrical constituent structure as the fundamental system of organization at the interface between PF and motor output can combine to yield new theoretical and descriptive insight. The remaining chapters are devoted to the detailed motivation and exposition of the processes whose interaction is schematized in (8), (9), and (10).

Chapter 2

Tibetan Tone and Stress

1. Introduction

While the Lhasa dialect of Tibetan has been fairly well-studied, with poorly understood and often conflicting results (cf. Hari 1979 chapter 3 for a survey), the various other dialects and sub-dialects of Tibetan have not been well-investigated. The present study reports on the results of intensive investigation of the speech of a speaker of what we will refer to throughout this work as Refugee Standard Tibetan (RST). This dialect is the lingua franca of the Tibetan refugee community in Nepal and India, particularly among those younger members of the community who have been brought up primarily outside Tibet. RST is loosely based on Lhasa dialect, with many elements of Tsang speech occurring as well. Much of the treatment of RST presented here is relevant for Lhasa dialect. The informant for this study was born in རྟོ་ལོ་ལྷོ་རྫོང་, of parents who had been born and raised there. At age two, he went with his family to Nepal, where he lived until his mid-twenties. The refugee community with whom he and his parents primarily associated consisted of families from throughout Tibet, especially Tsang. At the age of 10, the informant entered a Tibetan Buddhist monastery in Nepal, where he associated with monks and novices from all over Tibet and the refugee community.

This study, in addition to presenting some descriptive clarifications of the phonology of an important Tibetan dialect, proposes a new analysis of the tone/stress relation in Tibetan, which shows that certain tonal processes are dependent on abstract metrical structure. The formal system employed in stating these metrical dependencies is presented in Halle and Vergnaud 1987. Familiarity with the bracketed metrical grid system presented there (outlined in appendix 1) will aid understanding of the analysis¹.

The contributions of this study are as follows: first, an empirically adequate description of a broad array of phenomena in a standard dialect of Tibetan (Refugee Standard Tibetan) are presented for the first time; second, a novel metrical analysis is presented for Tibetan tone and stress phenomena, which, to the extent they are paralleled in the better-studied Lhasa dialect, have resisted principled treatment; third, a claim is made concerning surface underspecification of tone resulting on the surface in an interpolation effect analogous to recent claims for Japanese (cf. Pierrehumbert and Beckman 1987); fourth, evidence is presented for a particular interpretation of conflation in metrical theory (line deletion and subsequent unmetrification) and the empirical consequences of this conception are shown; fifth, interesting examples of two

¹ All claims regarding surface pitch contours are based on auditory impressions by the author (a trained phonetician).

distinct uses of metrical structure in a single domain are provided (metrical planes for tonal adjustment and stress are related but separate).

We follow Sprigg 1954 in observing that literate Tibetans employ three quite distinct styles of speech: "[Tibetan] utterances may be considered in terms of three styles, 'spelling style', 'reading style', and 'speaking style'", and, as in Sprigg 1954, "all utterances appropriate to the reading and quoting of written texts, 'reading-style' utterances, have been excluded from [this study]."

Because there has been some disagreement about the basic tonal structure of Tibetan, a substantial part of the following discussion is devoted to motivating the underlying tonal specifications, and showing their relations to syllable type.

2. RST Nouns

2.1 The Tonal Inventory of RST Monosyllabic Nouns

2.1.1 Syllable Coda Quality

Table (1) below shows the correspondences between orthographic, surface segmental, and tonal forms in RST. All possible syllable codas, as reflected in the orthography, are represented. The column headed 'Orthography' shows the shape of the syllable as represented by the 30 letter

Tibetan alphabet². Register qualities determined by orthographic prefixing, superscripting, and subscripting are not shown. Vowel quality, whether orthographically indicated or unlauded³, is not relevant to our discussion. All forms in the 'Orthography' column may occur with zero or non-zero onsets. We are abstracting away from this distinction.

² All examples will be cited in Uchen script. Correlations between orthography and spoken form are given to aid the reader in establishing a common frame of reference, and for testing the conclusions presented here with other RST native speakers. The letters used in Sanskrit transliteration and wazur conjunct letters are not employed.

³ With coronal orthographic suffix.

(1) RST Syllable Codas and Surface Tonal Correlations

| Orthography | Surface | L-reg | H-reg | Examples | Gloss |
|-----------------|---------|-------|-------|--------------|-----------------|
| v | v | 23 | 5 | ཅ་ཀ' | 1st;3rd letters |
| v _p | [pʔ] | 23 | 5 | ཡམ་ཅམ་ | father;water |
| vt | [ʔʔ] | 23 | 53 | བདུད་རྗེད་ | demon;high part |
| vk | [kʔ] | 23 | 5 | རག་རྩེག་ | brass;roof |
| vs | [v:] | 31 | 52 | བས་ཅས་ | barley;costume |
| vm | [m] | 24 | 55 | ཟུག་ཕུག་ | box;mango |
| vn | [n] | 24 | 52 | རྩེ་ལྷོ་ | religious dance |
| | [ṽ:] | | 55 | ཟུག་ | price eye |
| v _ŋ | [ŋ] | 24 | 55 | རྩེ་རྩེ་ | ravine;donkey |
| | | 31 | | ཟུག་ | story |
| v _l | [l] | 24 | 55 | རྩེ་ལྷོ་དུལ་ | jade;money |
| v _r | [r] | 24 | 55 | ཟུག་གསེར་ | camp;gold |
| v _{ks} | [kʔ] | 23 | 5 | བས་ཅམ་དུག་ | forest;breath |
| v _{ps} | [pʔ] | 23 | 5 | ཟུག་ཟུག་ | depth;method |
| v _{ms} | [m] | 31 | 52 | རྩེ་ལྷོ་ཟུག་ | disease;prison |
| v _{ŋs} | [ŋ] | 31 | 52 | ཟུག་དུག་ | copper;length |

The column headed 'Surface' in table (1) requires some explanation. This column is an indication of the surface phonetic form of the given syllable type in monosyllabic (non-compounded) citation form. This refers to ordinary non-casual speech. We are assured that illiterate speakers of RST have these forms as well as our literate informant, hence these forms are relevant for the 'speaking style'

considered here. Where a syllable has a surface phonetic form that lacks any coda consonant, the surface shape is given as 'V' for a short vowel and 'V:' for a long vowel. Where a syllable has a surface-realized coda consonant, a canonical phonetic shape is given, e.g. [p]. Where no further diacritics exist on an obstruent, it is to be understood as released. Unreleased obstruents will be so annotated, e.g. [k̚]. Coda obstruents are unvoiced in citation forms. For clarity, here is another example: where the 'Surface' column lists [m], that is to be understood as a segmentally realized labial nasal coda, not any sort of nasalized long vowel. Likewise, the orthographic lateral coda is segmental [l] in RST, not an unlauded long vowel (differing in this respect from Lhasa Standard Tibetan, henceforth LST).

Since table (1) is organized by orthographic form, it is important to note that the omission of Vts and Vns forms results from the systematic lack of such orthographic forms in the noun inventory⁴.

2.1.2 Glottal Stop Coda

A special note is in order regarding the glottal stop. Table (1) indicates that [ʔ̚] occurs only as the synchronic reflex of the /t/ orthographic coda. We take this to be the

⁴ The syllables which are orthographically VV, such as those formed by the addition of the diminutive suffix Q' , will not concern us here.

correct statement of the underlying forms. The orthographic forms [V], [Vk], and [Vks], however, all have [ʔ¹] coda realization on the surface freely varying with the surface coda realizations given in table (1). Because this variation is correlated with speech style and rate, and because it has tonal effects that will be important in the assignment of lexical items to tonal categories, discussion of this issue is postponed to section 2.1.4.1 below.

2.1.3 Surface Tone Properties

The tonal transcriptions are given for each syllable shape in table (1) under two headings: Low register and High register. Register in Tibetan refers to systematic differences of tone level and voice quality. The Low register is correlated in RST with breathy voice. The High register is correlated with clear, tense, or modal voice. The voice quality component of the register contrast is quite salient in RST⁵.

The tone melody for each syllable type in table (1) is indicated by the Chao tone numbering system, commonly used by Chinese researchers on East Asian tone. 5 is the highest level and 1 is the lowest in this system. There is some redundancy in the tonal forms indicated in table (1). In particular, the high level tone is represented as '5' for short syllables and as '55' for long (bimoraic) syllables.

⁵ Cf. Ossorio 1982 for discussion of the perceptual correlates of register in Tsang Tibetan.

This redundancy will be eliminated when we come to consider the phonological formulation of the tonal inventory below.⁶

2.1.4 Distribution of Tonal Shapes

On the basis of table (1), which gives the tonal possibilities for every surface syllable shape of RST, it can be observed that there are eight surface tonal shapes, partially correlated with syllable weight. These surface tone forms and the associated syllable types are shown in table (2) below:

(2) Surface tone shapes and their segmental affiliations

23 low register short rise:

| Orthography | Surface coda realization |
|-------------|--------------------------|
| V | V |
| Vp | [pʔ] |
| Vt | [ʔʔ] |
| Vk | [kʔ] |
| Vps | [pʔ] |
| Vks | [kʔ] |

⁶ Notice that table (1) does not attempt to correlate tonal register or contour with properties of syllable onsets. This diachronically indisputable correlation has been imported into the synchronic phonology with varying degrees of persuasiveness (Kjellin 1975, 1976; Civera 1971). We reject such correlations as synchronically relevant principles. See Hari 1979 for a persuasive refutation of attempts to derive register from orthography. In any case, the derivation of register is not our concern in this research. The reason table (1) is organized in terms of coda orthography is to 'set a level' and establish the degree of orthographic/phonetic correspondence for segments and tonal shapes in RST (which is more conservative in this respect than LST).

24 low register long rise:

| Orthography | Surface coda realization |
|-------------|--------------------------|
|-------------|--------------------------|

| | |
|----|-----|
| Vm | [m] |
| Vn | [n] |
| Vŋ | [ŋ] |
| Vl | [l] |
| Vr | [r] |

31 low register long fall:

| Orthography | Surface coda realization |
|-------------|--------------------------|
|-------------|--------------------------|

| | |
|-----|------|
| Vs | [V:] |
| Vn | [n] |
| Vms | [m] |
| Vŋs | [ŋ] |

5 high register short level:

| Orthography | Surface coda realization |
|-------------|--------------------------|
|-------------|--------------------------|

| | |
|-----|--------------------|
| V | [V] |
| Vp | [p ^ɹ] |
| Vk | [Vk ^ɹ] |
| Vps | [Vp ^ɹ] |
| Vks | [Vk ^ɹ] |

55 high register long level:

| Orthography | Surface coda realization |
|-------------|--------------------------|
|-------------|--------------------------|

| | |
|----|-------|
| Vm | [m] |
| Vn | [V̆:] |
| Vŋ | [ŋ] |
| Vl | [l] |
| Vr | [r] |

52 high register long fall:

| Orthography | Surface coda realization |
|-------------|--------------------------|
|-------------|--------------------------|

| | |
|-----|------|
| Vs | [V:] |
| Vms | [m] |
| Vŋs | [ŋ] |
| Vm | [m] |

53 high register short fall:

| | |
|-------------|--------------------------|
| Orthography | Surface coda realization |
| Vt | [ʔʔ] |

From table (2) we can begin to observe the correspondences between tone shape and syllable weight, and begin to speculate about possible reductions of the surface inventory, the standard procedure in discussing the tonal phonology of any East Asian tone language. We will also want to consider whether any of the surface shapes can be plausibly derived by universal phonetic conditioning, that is, the effect of segmental quality on tone, which is another standard source of tonological inventory reduction for East Asian languages.

2.1.4.1 High Register Tone Shapes

We will analyze the high register tones as falling into two categories: level and falling. This partitioning of the inventory cannot be done solely on the basis of synchronic syllable quality, and therefore every high register syllable must have its contour type specified in its lexical representation. However, there are some correlations with orthographic and surface syllabic quality that are brought out in the discussion below.

2.1.4.1.1 The Level Class

The following (surface) syllable qualities are invariably specified as level in lexical (underlying) representation: [\check{V} :] (from orthographic [Vn]), [V \check{v}], [Vl] and [Vr]). Syllables of shape [Vm], however, may be lexically specified as bearing either level or falling tone (cf. table 1), and this is the first indication that the level/falling distinction is lexical, rather than predictable from syllable quality or phonetic conditioning.

Furthermore, we analyze the following (high register) syllable types as underlyingly level: [V], [Vp] (from orthographic [Vp] and [Vps]), [V \check{v}] (from orthographic [Vt]), [V \check{k}] (from orthographic [Vk] and orthographic [Vks]). [V] and [V \check{k}] may optionally be realized with [ʔ] syllable codas, and when [ʔ] appears, the tonal realization is falling rather than level. The falling terminal contour is a common phonetic effect of a glottal stop syllable coda (cf. Hari 1979 and references therein⁷). This analysis in terms of phonetic conditioning explains the lack of an optional falling contour on [Vp \check{v}] syllables (there is no optional glottal coda realization of these), and the lack of a level realization of [V \check{v}] (from

⁷ We note in passing that while some aspects of Hari's analysis of Lhasa are relevant for the present discussion, the essential claim of her work, that 4-way tonal contrasts are found on short, coda-less syllables in LST (two contrasts in each register) has not been upheld by work with the RST-speaking informant.

orthographic [Vt]): these are invariably realized with glottal codas.

Finally, where '55' is given in table 1, as opposed to '5', this indicates that a sonorant mora is available for extended length realization of the level tone.

2.1.4.1.2 The Falling Class

Orthographic shapes [Vs], [Vms], and [Vns] are invariably falling tone on the surface in high register. However, since the synchronic realizations of orthographic [Vms] and [Vns] are segmentally identical with [Vm] and [Vn], these syllables must be underlyingly distinguished with falling tone specification. This is another indication that, despite some measure of predictability, there is indeed a level/falling lexical contrast in the high register, rather than simply level tone underlyingly, with the falling shape derived predictably by syllable quality or phonetic conditioning.

2.1.4.2 Low Register Tone Shapes

In the low register, the following syllable types will be grouped together as specified for level tone in lexical representation: [V], [Vp̣] (from orthographic [Vp] and [Vps]), [Vʔ̣] (from orthographic [Vt]), [Vḳ] (from

orthographic [Vk] and [Vks]), [Vm], [Vn], [Vl], and [Vr]. On the surface, these appear as low rising tone, with greater length in the rise (tone shape 24 as opposed to 23 in table (1)) for syllables with sonorant coda morae. The rise will be inserted between the lexical representation and the surface, by a process to be described in section 2.2.2.4 below. Syllables of shape [Vŋ] may be lexically specified as bearing either level or falling tone, and this is yet another consideration in analyzing RST as having a lexical level/falling contrast, rather than predictable alternation between the shapes.

Orthographic shapes [Vs], [Vms], and [Vns] are invariably falling tone on the surface in low register. However, since the synchronic realizations of orthographic [Vms] and [Vns] are segmentally identical with [Vm] and [Vŋ], these syllables must be underlyingly distinguished with falling tone specification. This again reflects the lexical nature of the level/falling contrast.

2.1.5 Phonological Form of Tone Specifications

In considering possible formal specifications of the tone shapes discussed in section 2.2, we will adopt as a general mechanism of description the tonal machinery proposed by Bao 1990. In this system, the tonal root node will be notated as 't'. It dominates the register node ('r') on its left branch, which in turn dominates the feature

[stiff vocal folds]. On the tonal root node's right branch, it dominates a contour specification node 'c', which in turn dominates the feature [slack vocal folds]. The 'c' (contour) node may branch, which allows for the representation of contour tones as adjacent feature specifications. For purposes of this discussion, we will take a single tonal root node to be attached to each syllable⁸.

This system affords us a conveniently discrete representation for register, which we will assign to underlying forms of lexical items, foregoing any attempt to derive register from orthography or synchronically observable onset voicing qualities⁹.

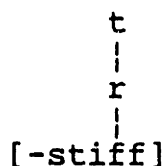
The full abstract phonological shapes of the RST tonal inventory, represented using Bao's machinery, are shown in (3) below:

⁸ An argument can be made that RST is a mora-tone language, but while we are not hostile to this view, the issue is not central to our present concerns, and the syllable (or rime) tone view is more convenient for expository purposes.

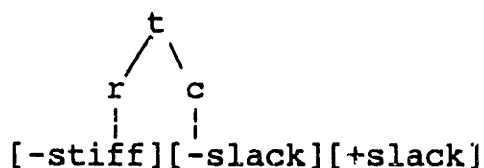
⁹ We are not concerned with the representation of the voice quality feature correlated with low register (breathy voice). Cf. Yip 1989 for some discussion.

(3) Phonological Tone Specifications for RST

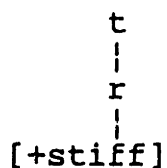
a. Low register, 23/24, rise



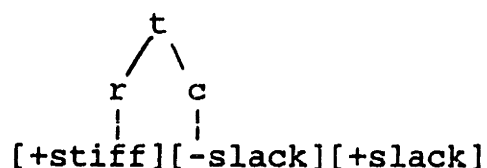
b. Low register, 31, fall



c. High register, 5/55, level



d. High register, 52, fall



It will be observed that tone (3)a., the low register rising tone, is not underlyingly specified as a rise. Since it does in fact appear as a low rise in surface citation form, there will be a process of rise insertion somewhere in the derivation. The exact phonological triggering of this process will depend on metrical representation, and is discussed more fully in the following sections.

For clarity of exposition, the tonal representations above will be modified somewhat in derivations throughout the rest of this chapter. Since the feature names, while they represent important claims regarding universals of the phonologic/phonetic interface, are not crucial for the discussion to follow, a cover feature [fall] will be used as short-hand for the 'c' subnode complex: [-slack][+slack]; [rise] will be used for [+slack][-slack]; H will be used for the 'r' subnode feature [+stiff]; and L will be used for

[-stiff]. Again, the use of this 'shorthand' is not in any sense to be construed as embodying a theoretical claim (e.g. on unitary contour tone). We are at liberty to do this because the tonal system of Tibetan, unlike the more complex systems considered by Bao 1990, does not happen to make use of all the degrees of freedom in contour representation provided by the geometry. The full expository representations that will be used are given below:

a. Low register, 23/24, rise

```

t
|
r
|
L

```

b. Low register, 31, fall

```

  t
 / \
r   c
|   |
L   fall

```

c. High register, 5/55, level

```

t
|
r
|
H

```

d. High register, 52, fall

```

  t
 / \
r   c
|   |
H   fall

```

2.2 Tone Patterns of Compound Nouns

2.2.1 Disyllabic Compounds

The bulk of the noun inventory of RST consists of disyllabic stem compounds. These are formed by joining two potentially independent syllables. The component syllables may not always have independent lexical existence outside of any compound, but in general the semantic contribution of each stem to the compound can be clearly identified (cf. Goldstein 1984 for discussion of the semantic typology of

LST disyllabic noun compounds). When the two participating syllables are joined, there are certain tonal and segmental effects that are characteristic of the compound as a whole. That is, a member of the compound may exhibit segmental or tonal effects that are not predicted by the simple juxtaposition of two syllables. The segmental effects include a certain degree of voicing neutralization and place assimilation at the juncture¹⁰, which will not be a central concern in this discussion.

2.2.1.1 Tonal and Segmental Effects of Compounding

In this section, we will be primarily concerned with the tonal qualities of disyllabic compounds. In particular, we will examine how the citation tones of syllables change under compounding, and what the optimal formal statement of these effects might be. Table (4) below, summarizes the segmental and tonal effects of compounding on each syllable type. The syllables are organized by coda orthography as in table (1), to bring out correspondences between orthographic, segmental, and tonal forms and to facilitate comparison between this description of RST and descriptions of other Tibetan dialects.

In table (4) the 'Orthography' and 'Surface' columns are as in table (1). The column headed 'S1 surface' gives an indication of the coda quality of the relevant syllable type

¹⁰ Voicing neutralization at the juncture is more prevalent in LST than in RST.

when it is in initial position in a compound. This representation is not intended to show neutralization and assimilation effects on the S1 coda at the juncture. It is intended to show whether the coda consonant, if any, of an S1 syllable is preserved, in some form. In the general case, an S1 syllable retains its coda consonant in some form, though that segment may be affected by place assimilation with a following S2 onset (particularly nasal codas). For obstruent codas, it can be observed from table (4) that these do not delete, they are clearly audible (and, in the case of [p], visible as well). As for long vowels, note that high register [V:] long syllables in citation are [V:] as S1 as well.

The column headed 'S2 tone' shows the tonal quality of the relevant syllable type in second compound position. The high register tones are not affected by placement in S2 position, while the low register tones are each mapped to a high register tone.

Table (4) retains the pretheoretic distinction between 5/55 and 23/24, to show orthographic/segmental and tonal correlations with maximum clarity. This organization of the data does not affect the theoretical discussion to follow.

(4) Surface tone shapes and their compound behavior**23 low register short rise:**

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|-------------------|------------|---------|---------|
| V | V | V | 2 | 5 |
| Vp | [p] | [p] | 2 | 5 |
| Vt | [ʔ ¹] | [V:] | 22 | 5 |
| Vk | [k ¹] | [k] | 2 | 5 |
| Vps | [p ¹] | [p] | 2 | 5 |
| Vks | [k ¹] | [k] | 2 | 5 |

24 low register long rise:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|---------|------------|---------|---------|
| Vm | [m] | [m] | 22 | 55 |
| Vn | [n] | [n] | 22 | 55 |
| Vŋ | [ŋ] | [ŋ] | 22 | 55 |
| Vl | [l] | [l] | 22 | 55 |
| Vr | [r] | [r] | 22 | 55 |

31 low register long fall:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|---------|------------|---------|---------|
| Vs | [V:] | [V:] | 22 | 52 |
| Vms | [m] | [m] | 22 | 52 |
| Vŋs | [ŋ] | [n] | 22 | 52 |
| Vŋ | [ŋ] | [n] | 22 | 52 |

5 high register short level:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|-------------------|------------|---------|---------|
| V | [V] | [V] | 5 | 5 |
| Vp | [p ¹] | [p] | 5 | 5 |
| Vk | [k ¹] | [k] | 5 | 5 |
| Vps | [p ¹] | [p] | 5 | 5 |
| Vks | [k ¹] | [k] | 5 | 5 |

55 high register long level:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|---------|------------|---------|---------|
| Vm | [m] | [m] | 55 | 55 |
| Vn | [n] | [n] | 55 | 55 |
| Vŋ | [ŋ] | [ŋ] | 55 | 55 |
| Vl | [l] | [l] | 55 | 55 |
| Vr | [r] | [r] | 55 | 55 |

52 high register long fall:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|---------|------------|---------|---------|
| Vs | [V:] | [V:] | 55 | 52 |
| Vms | [m] | [m] | 55 | 52 |
| Vŋs | [ŋ] | [ŋ] | 55 | 52 |
| Vm | [m] | [m] | 55 | 52 |

53 high register short fall:

| Orthography | Surface | S1 surface | S1 tone | S2 tone |
|-------------|-------------------|------------|---------|---------|
| Vt | [ʔ ⁷] | [V:] | 55 | 53 |

Table (4) is intended to show (1) that syllable codas do not delete in S1 position (this possibility would be relevant to a mora tone analysis, because the loss of a mora would be expected to have tonal effects), and (2) that certain tonal qualities are systematically retained in S1 and S2 position, while others are lost. This issue is taken up in the section just below.

2.2.1.2 A Restriction on Tonal Output

Table (5) below shows, on the left, every possible combination of S1 and S2 tonal types. On the right is the unique tonal output for the compound, based on the changes given above in table (4):

(5) Disyllabic noun compound tonal combinations:

| Input: | | Compound form: | |
|---------|---------|----------------|---------|
| S1 | S2 | S1 | S2 |
| H level | H level | H level | H level |
| H fall | H level | H level | H level |
| L level | H level | L level | H level |
| L fall | H level | L level | H level |
| H level | H fall | H level | H fall |
| H fall | H fall | H level | H fall |
| L level | H fall | L level | H fall |
| L fall | H fall | L level | H fall |
| H level | L level | H level | H level |
| H fall | L level | H level | H level |
| L level | L level | L level | H level |
| L fall | L level | L level | H level |
| H level | L fall | H level | H fall |
| H fall | L fall | H level | H fall |
| L level | L fall | L level | H fall |
| L fall | L fall | L level | H fall |

Given the 4 distinct lexical tones on monosyllables, there ought, by calculation, to be 16 tonal shapes of disyllable compounds. In fact only the following shapes occur:

(6)

| 1st syllable | / | 2nd syllable |
|--------------|---|----------------------|
| H level | | H level or H falling |
| L level | | H level or H falling |

We can see from (6) that only 4 of the 16 logically possible combinations occur. The first syllable (henceforth S1) is realized as H level whenever the monosyllable tone

borne by S1 is H register (tones (3)c. and (3)d. above), and L level when the tone of S1 is L register (tones (3)a. and (3)b. above). The second syllable (henceforth S2) is realized as H falling when the monosyllable tone borne by S2 is falling (tones (3)b. and (3)d. above), while S2 is realized as H level whenever the monosyllable tone is non-falling (tones (3)a. and (3)c. above).

2.2.1.3 Metrical Dependencies (MS1)

The tonal properties of the disyllabic nouns given in (5) are derived as shown in (7) below, employing abstract metrical structure. This metrical structure controls tonal neutralization and surface realization. A subsequent metrical process, to be given in a later section, takes the output of (7) as input for stress assignment. There are thus two distinct metrical planes functioning in lexical compounding: the metrical representation that underlies tonal adjustment and realization, and the metrical representation that derives surface stress. In using two distinct metrical representations, we are adding support to the conclusions reached by Halle and Vergnaud 1987, based on the biplanar analysis of Tiberian Hebrew proposed by Rappaport 1984:

- ... more than one metrical constituent structure may be associated with a given central line of phonemes
- ... metrical constituent structure need not always be interpreted in stress terms

- ... a special relationship holds between the head of a metrical constituent and the rest of the elements in the constituent.

[Halle and Vergnaud 1987, p. 69]

In Tiberian Hebrew, there is a battery of parameter settings for stress that constructs a metrical plane which encodes qualities that are required for the subsequent operation of the metrical process governing Vowel Reduction. The Tibetan system is analogous in that two distinct metrical representations are motivated. The metrical process that governs tonal realization will be shown to precede the process governing stress. However, on subsequent cycles, reapplication of the metrical process governing tonal realization may occur, in which case the second application takes the stress plane into account.

We begin with the statement of metrical settings for tonal adjustment and realization. The settings for stress and the interaction between the two systems will be taken up below.

(7) Parameter Settings for Compound Tonal Assignment (MS1)

(This battery of settings will be referred to in various places below as 'Metrical Structure 1' or 'MS1')

- a. Syllables project on line 0.
- b. Line 0 parameter settings are:
(+BND,+HT,right,left to right).
- c. Mark heads of line 0 constituents on line 1.
- d. Line 1 parameters settings are:
(-BND,+HT,left).
- e. Mark heads of line 1 constituents on line 2.
- f. Conflate (delete line 1 in the grid)¹¹.

Tonal processes are dependent on the metrical structure as follows:

- g. Insert [+stiff vocal folds] under 'r' subnode of the tonal root node of metrical head (now represented on line 1, post-conflation).
- h. Any unmetrified syllable has its tonal root node delinked (deleted, with all dependents).
- i. Any metrified syllable that is not the head of its constituent has its 'c' (contour) tonal subnode delinked (deleted, with all dependents).

2.2.1.4 Disyllabic Compound Derivations

Before we consider the interesting theoretical ramifications of (7), we will examine some sample derivations, to get a feel for the operation of the system:

¹¹ The notion of conflation adopted here is that motivated by Halle and Kenstowicz 1990, whereby a line can be entirely deleted from the grid, and any lower grid lines that may lose their heads by this deletion become unmetrified grid positions. This is specially noted because it differs from the formalization of conflation in Halle and Vergnaud 1987.

(8) Metrical Derivations of Disyllabic Noun Tone Patterns¹²

a. Tibet + person, 'Tibetan' ལྷན་པོ་

```

line 0:      *           *
word:      phöö        mi
            |           |
            t           t
            |           |
            r           r
            |           |
            L           L

```

Above is the initial representation of two syllables that are to be joined in a compound. S1 and S2 both bear low register, non-falling tones. Each syllable projects a grid position to line 0.

```

line 1:      *
line 0:      (*         *)
word:      phöö        mi
            |           |
            t           t
            |           |
            r           r
            |           |
            L           L

```

Above, we observe effects of constituent construction on line 0, followed by head marking on line 1.

¹² The segmental line of all Tibetan examples is given in the pedagogical phonemic system developed by Goldstein and Nornang 1984. This system was based on Lhasa dialect rather than RST, so there are some minor discrepancies. Furthermore, due to typographical limitations, some vowel quality distinctions have been suppressed. All segmental qualities that are relevant for the discussion of metrics and tone are retained. In addition, the Tibetan script for all examples is provided, for native informant checking.

```

line 2:          *
line 1:          (*)
line 0:  (*      *)
word:   phöö    mi
        |      |
        t      t
        |      |
        r      r
        |      |
        L      L

```

Above, we observe head marking on line 2 ((7)d. and e.).

```

line 1:          *
line 0:  (*      *)
word:   phöö    mi
        |      |
        t      t
        |      |
        r      r
        |      |
        L      L

```

Above, we see conflation ((7)f.), which does not result in unmetrification of any line 0 positions, because the line 2 head matches the line 0 parameter settings. (After deletion of line 1 above, line 2 is renumbered as line 1).

```

pitch: |          —
        |          —
        |—————|
line 1: |          *
line 0: |  (*      *)
word:  | phöö    mi
        | |      |
        | t      t
        | |      |
        | r      r
        | |      |
        | L      H

```


Above, we see the tonal change in S2's register, based on its status as the metrical head ((7)g.). We see that in the

final form, S1, though it is not the head, has no 'c' node to be delinked, so that (7)i. does not apply, or applies vacuously. The output realization thus has S1 low and level, and S2 high. (The representation above will be input to a stress marking process, to be discussed in a separate section below.)

b. banner + iron, 'iron banner fixture' ᄃᆞᆫ ᄃᆞᆫ ᄃᆞᆫ

| | | | |
|---------|------|------|------|
| line 2: | | * | |
| line 1: | | (*) | |
| line 0: | (*) | * | |
| word: | thuu | caa | |
| | | | |
| | t | t | |
| | | | |
| | r | c | r |
| | | | |
| | H | fall | H |
| | | | fall |

In b. above, we see a word with a high falling tone on both S1 and S2. The representation above shows the compound after all metrical processes have applied (except conflation).

pitch: | 

| | | | |
|---------|------|-----|------|
| line 1: | | * | |
| line 0: | (*) | * | |
| word: | thuu | caa | |
| | | | |
| | t | t | |
| | | | |
| | r | r | c |
| | | | |
| | H | H | fall |

Above, we observe the tonal effects of the metrical process: the S1 contour has been delinked ([+stiff] register

insertion in S2 has applied vacuously, as S2 previously bore [+stiff] register). The surface realization will have a high level tone on S1, and a high falling tone on S2.

c. cotton + robe, 'cotton robe' ɪ ʌ ' ɜ ɹ

| | | | |
|---------|-----|------|-----|
| line 2: | | | * |
| line 1: | | | (*) |
| line 0: | (* | | *) |
| word: | ree | | see |
| | | | |
| | t | | t |
| | | | |
| | r | c | r |
| | | | |
| | L | fall | L |

In c. above, we see the representation of the compound after metrical processes have applied (except conflation).

pitch: |

| | | | |
|---------|-----|-----|---------------|
| line 1: | | * | |
| line 0: | (* | | *) |
| word: | ree | see | cotton + robe |
| | | | 'cotton robe' |
| | t | t | |
| | | | |
| | r | r | |
| | | | |
| | L | H | |

Above, the contour subnode of the non-head, S1, has been deleted, while a [+stiff] specification has been inserted under the register subnode of the head, S2.

d. mother + great, 'mother' (honorific) युम'कैव

| | | |
|---------|-----|------|
| line 2: | | * |
| line 1: | | (*) |
| line 0: | (*) | *) |
| word: | yum | chéē |
| | | |
| | t | t |
| | | |
| | r | r |
| | | |
| | L | H |

In (8)d. above, the input consists of a (lexically) low-toned syllable followed by a high-toned syllable. The metrical structure is constructed as shown above (prior to conflation).

pitch: |

| | | |
|---------|-----|------|
| line 1: | | * |
| line 0: | (*) | *) |
| word: | yum | chéē |
| | | |
| | t | t |
| | | |
| | r | r |
| | | |
| | L | H |

Above we see that (7)g. has applied vacuously, and since 'yum' had no 'c' tonal subnode, it is not changed.

In the disyllabic compounds so far considered, we have not had occasion to observe unmetrification ((7)h.), because the head has always been the rightmost element. In the

following section on trisyllabic compounds there will be occasion for (7)h. to apply.

2.2.2 Trisyllabic Compounds

There is a good stock of 3-syllable compound nouns and borrowed words in Tibetan. We will see that these succumb to the metrical analysis given above for disyllables, and supply important confirming evidence for aspects of that analysis. Within the general class of trisyllabic nouns, there are both true compounds and synchronically unanalyzable words. The internal morphological bracketing of these different types is relevant and will be treated below, but we will continue, for convenience, to refer to the class as a whole as compound words (though some are not really compounds in the strict sense).

2.2.2.1 Morphological Types

It is necessary to distinguish two types of 3-syllable compounds, based on morphological structure:

(8) Morphological Bracketing Types of Trisyllabic Compounds

- a. [S1 [S2 S3]] and [[S1] S2 S3]
 b. [S1 S2 S3] and [[S1 S2] S3]

The forms with morphological bracketing (8)a., in which S1 (the first syllable in the compound) does not form a sub-constituent with S2, will have the accent realized on S1

itself. We therefore stipulate that S1 in morphological types (8)a. is marked with a line 1 asterisk in the lexical entry. This stipulation is not entirely arbitrary, obviously, as there is a correlate of morphological structure¹³.

2.2.2.2 Pitch Interpolation

Before we consider sample derivations of trisyllabic nouns, it is appropriate to consider the implications of tonal process (7)f. and (7)h., conflation and subsequent unmetrification, in greater detail. The tonal reflex of unmetrification, as prescribed in (7)h. above, is loss of the tonal root node. The surface realization of such toneless syllables in unmetrified positions is generally a mid-low falling contour, a 'trailing away'. We therefore ask the question of what principles determine the tonal realization of toneless syllables.

Pierrehumbert and Beckman 1987 (henceforth PB87) have proposed and meticulously motivated a process of 'linear interpolation to low' for Japanese unaccented and accented minor phrases. They show that in unaccented phrases there is a linear interpolation from a phrasal H tone to a L% at the right edge of the phrase. This subsumes the previously proposed rightward H-spreading account for Japanese, as the

¹³ This effect could possibly be seen as a consequence of cyclic compounding, but for present purposes it is sufficient that S1 be distinguished in the final morphological bracketing.

phonetic facts do not require (in fact, argue against) any such full phonological tone specification, and the phonology does not require or independently motivate such a process.

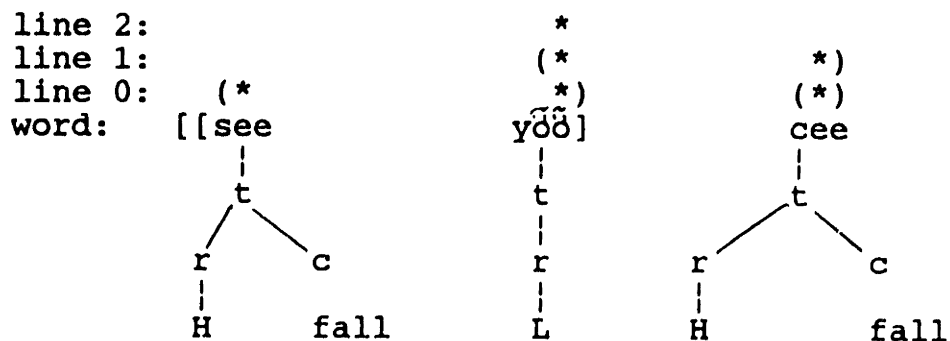
In accented phrases, on the other hand, where the accent is represented as HL, PB87 show that the L of the accent is systematically higher than a right boundary L%, and there is linear falling interpolation between the L tones, rather than rightward L-spread.

In Tibetan, a similar process occurs, with the difference that the applicability of interpolation is based on delinking of underlying tone (rather than non-application of spread), and this delinking is determined by abstract metrical structure, as specified in (7)h. We therefore posit a right boundary L% at the end of lexical domains such as those we have encountered so far (we will have occasion to examine refinements of the domain specification in later sections). There is a process of linear interpolation of pitch specification between the final tone value of the metrical head and the L% right boundary tone, across any syllables that have had lexical tone delinked and deleted due to their unmetrified status.

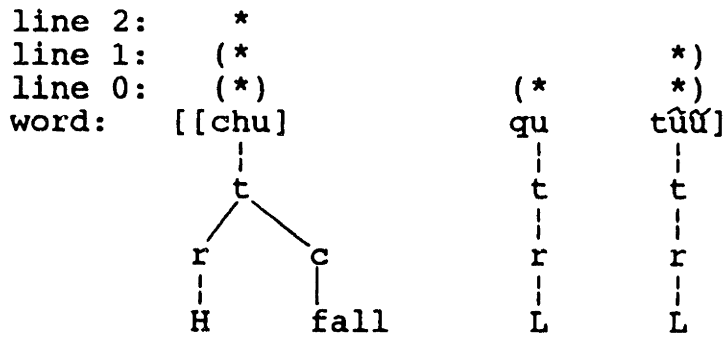
In the derivational stage above, conflation has occurred, resulting in unmetrification of 'ghan', the final syllable. To review: the reason 'ghan' is unmetrified here (appears within no metrical constituent boundaries) is that the former line 1 (see above) has been deleted, under our interpretation of conflation. The previous line 2 head, on 'su', coincides with the line 0 parameter settings, and hence can now appear as a line 1 head of a line 0 constituent. The small constituent embracing 'ghan' however has now lost its head, and under our interpretation of conflation this entails loss of inferior constituent structure, on line 0. There is only one such unmetrified syllable in (9)a., hence of necessity the interpolation is short in its time course. The phonetic beginning point of the interpolation will be approximately around the mid-tone level, based on the end point of a short high-register falling contour. We will examine longer unmetrified stretches in later examples.

b. knowledge + possessor, 'intellectual'

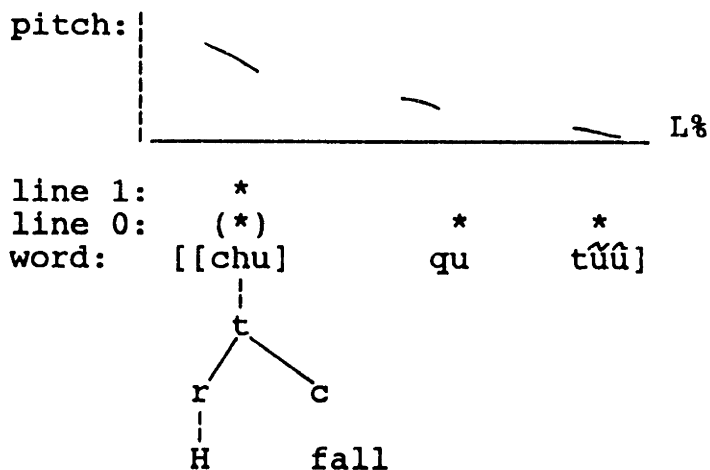
ज्ञान' धर्म' श' व'



c. water + certain kind of grape 'grape' (general) ཐུ་ཚུལ་འབྲུག་



In (9)c. above, we observe for the first time the effect of a prespecified line 1 asterisk, based on the morphological bracketing. The metrical constituent construction has had to respect the prespecified head, resulting in the metrical head shown on the leftmost syllable, prior to conflation.



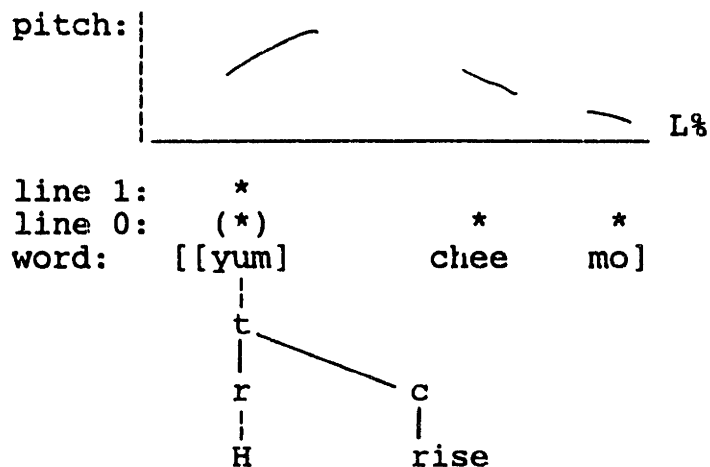
When conflation occurs, deleting (previous) line 1, the right-most binary constituent is left without a head, resulting in unmetrification, and subsequent loss of tone. The representation above is then input to the interpolation process, between the end of the short high register fall on 'chu' to the L% tone at the right boundary. As for the

initial syllable itself, we see that it does not lose its falling contour, as the initial syllable has in all the examples above, and this is due to its status as metrical head ([+stf] register insertion, ((7)g.), applies vacuously).

d. mother + great, 'mother' (honorific) ཡུལ་ཚེ་མོ་

| | | | |
|---------|--------|------|-----|
| line 2: | * | | |
| line 1: | (* | | *) |
| line 0: | (*) | (* | *) |
| word: | [[yum] | chěě | mo] |
| | | | |
| | t | t | t |
| | | | |
| | r | r | r |
| | | | |
| | L | H | L |

In (9)d., above, we see an interesting contrast with (8)d., farther above. 'chěěmo' is an alternate form of 'great', which adds a syllable to the representation, with the bracketing as shown. Given the bracketing, a line 1 asterisk is prespecified, resulting in the metrical headship falling on 'yum' (cf. (8)d., where 'chee' is head). The representation is shown prior to conflation.



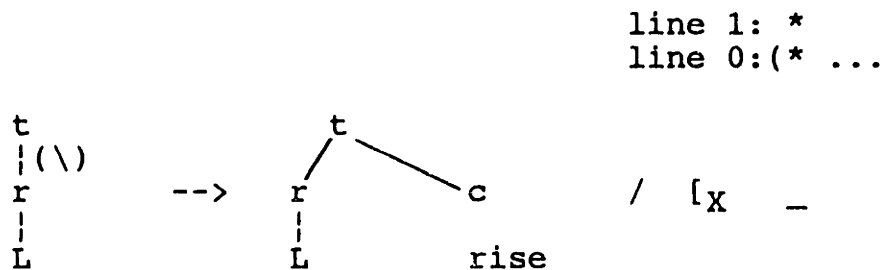
Above, we see the result of applying conflation, which unmetrifies 'chee mo', and causes these syllables to lose their tonal specification. However, there is an unexpected element in the representation: a contour node has been inserted under the accented syllable. This is discussed below.

2.2.2.4 Rise Insertion

The contour insertion in (8)d. is an instance of a general process in RST, whereby domain-initial, accented, low-register syllables acquire a rising contour. This process must apply subsequent to accent placement and prior to [+stiff] insertion ((7)g.), because accented, domain-initial high register syllables do not acquire a rise. Clearly, being a domain-initial low register syllable is not sufficient trigger: consider (8)d., 'yumchěé', with its word-initial low register 'yum' which does not acquire a rise. In (9)d., however, all the conditions are met, and

'yum' thus acquires a rise. The trigger is specified as 'domain-initial' because we do not yet want to commit ourselves to a particular domain, such as 'word' or N^0 (hence the domain bracket is just given as 'X' in the formalization below) Note that this process results, on the surface, in a rise in the high ([+stiff]) register because of the independent (7)g. ([+stiff] insertion), distinct from a possible rise in the low ([-stiff]) register. Rise insertion, which also accounts for the citation rise observed on a class of low register monosyllables (cf. table (1)), will be discussed further below.

(10) Rise Insertion¹⁴



(The parenthesized null right branch in (10) indicates that the rise insertion takes precedence over any possible pre-existing contour specification, we will see this provision in action below).

¹⁴ There is a rule with this name and somewhat overlapping effect specified in Dawson 1980, but since that account is not based on a metrical analysis, detailed comparison is not especially illuminating and will not be undertaken.

2.2.3 Quadrisyllabic Nouns

There are a number of 4-syllable nouns in the RST inventory. These are mostly place names or transliterations from Sanskrit. They provide important confirming evidence for the analysis given above. In all cases of true N^0 (non-phrasal) 4-syllable nouns, because of the essentially unanalyzable nature of the words (from the synchronic point of view), the initial syllable never carries a pre-specified line 1 asterisk. The morphological bracketing will always be [S1 S2 S3 S4]. Thus we expect S2 always to be the metrical head, and exhibit the associated tonal characteristics¹⁵.

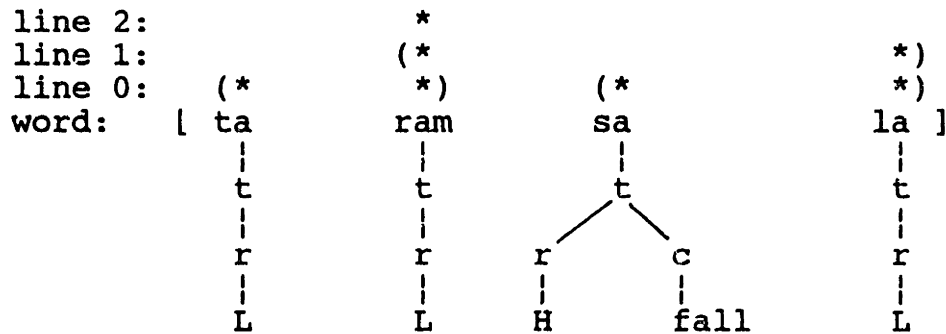
2.2.3.1 Quadrisyllabic Noun Derivations

Consider some sample derivations:

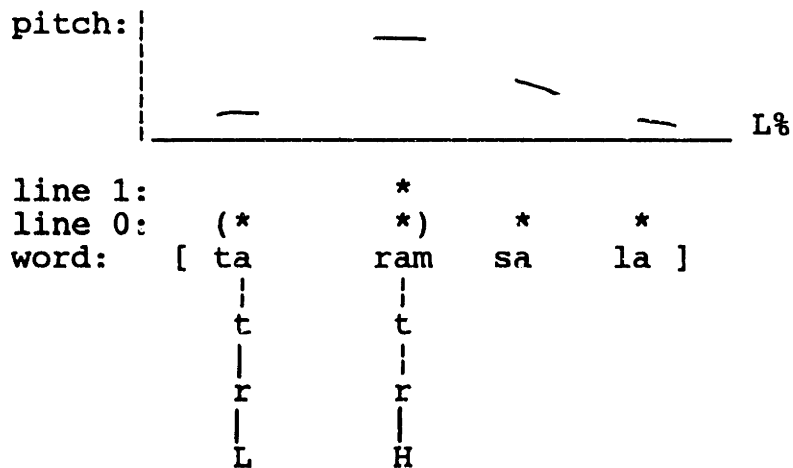
¹⁵ Note that there is a large class of quadrisyllabic compound nouns in RST formed of two disyllabic noun components. A typical example is: /nemdu pepdan/ (नमदुपेपदान) 'airplane' + 'station' = 'airport'. In these, it is clear that the second disyllabic compound element undergoes pitch range compression based on its position. Thus H register tones occurring in the rightmost disyllabic component are nearly neutralized with L register tones in that component. This interesting phenomenon will not be further treated here, though it calls out for detailed phonetic analysis.

(11)

a. Dharamsala ཧྲ་རམ་ས་ལ་



In (11)a. above, we see the metrical representation of the place name 'Dharamsala', with the lexical tones of the syllables as indicated (based on the orthography). The metrical structure has been built in accordance with (7) above, and is shown prior to conflation.



Above we see the familiar results of application of (7)f.-i.

b. H. H. the Dalai Lama ཡིན་པའི་འཕགས་པ་

line 2: *
 line 1: (* *)
 line 0: (* *) (* *)
 word: [yi sii noo pu]
 | | | |
 t t t t
 | | | |
 r r r r
 | | | |
 L L L L

In (11)b. above, we see the metrical structure prior to conflation.

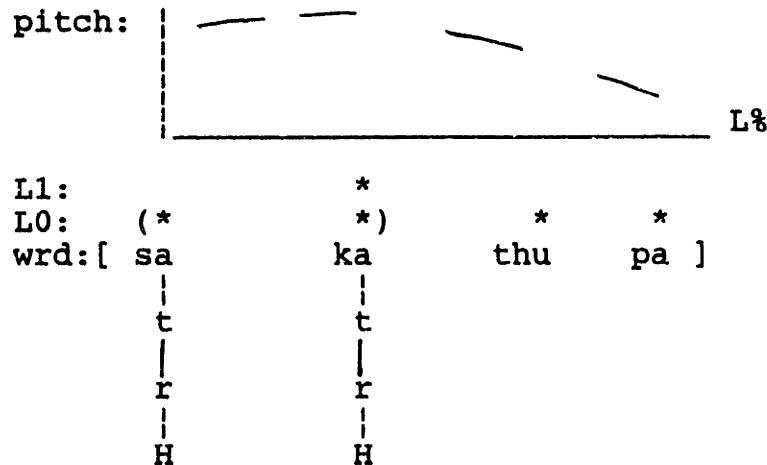
pitch: |
 ————— L%
 — — —

line 1: *
 line 0: (* *) * *
 word: [yi sii noo pu]
 | |
 t t
 | |
 r r
 | |
 L L

c. Buddha འུ་རྒྱལ་པོ་

L2: *
 L1: (* *)
 L0: (* *) (* *)
 wrd: [sa ka thu pa]
 | | | |
 t t t t
 / | | |
 r r r r
 | | | |
 H H L H
 fall

The stage shown in (11)c. above is prior to conflation.



2.3 Stress Patterns of Nouns

In addition to tone, RST has lexical stress. The stress system is based on a complex interplay of considerations of tone, syllable weight, and syllable position. In this section, we will show both that the metrical structure for stress marking is non-isomorphic with the abstract metrical structure motivated in the tonal derivations (process (7) MS1, above), but further that stress assignment is crucially dependent on certain aspects of the structure built by MS1, providing interesting confirmation for the metrical analysis of tone. The model we will initially assume is diagrammed below:

(12) Tone/Stress Process Ordering

- a. Input: juxtaposed syllables, bearing lexical tone
- b. Metrical structure 1 ((7)a.-f, MS1)
- c. Tone realignment ((7)g.-i., MS1)
- d. Metrical structure 2 (stress marking, MS2)
- e. Phonetic output (including interpolation
across toneless syllables)

In this section we will develop (12)d., 'Metrical Structure 2' (stress metrics, henceforth MS2), and consider its relation to MS1 (abstract metrics upon which tone alignment is dependent, cf. (7) above).

2.3.1 Relation of Stress Metrics to Abstract Tonal Metrics

The first consideration in relating MS2 to MS1 is the observation that only metrified syllables (resulting from (7)a.-f., MS1) are eligible for lexical stress. This is an interesting confirmation of the abstract accentual approach to tonal alignment pursued in the sections above, because if we took some non-metrical approach to tone assignment and alignment (replacing MS1), we would have to recreate much of the constituent structure anyway, in the discussion of stress derivation.

In practice, the requirement that only unmetrified syllables are eligible for stress limits the possible head positions to the initial two syllables of a noun. This

raises the question of whether it would be possible to simply take over the heads deriving from application of MS1 directly, resulting in complete isomorphism between the two systems. This will not be possible: MS2 is a related but distinct process, as we will see below.

2.3.2 Basic Units of Stressability

The stress system takes tone, syllable weight, and syllable position into account. The 'c' (contour) tonal subnode is not considered in stress calculations, which view only the 'r' (register) tonal subcomponent. The basic hierarchy of units upon which the stress system operates is given in the following (descending) hierarchy of eligibility for stress:

(13) Stressability Hierarchy

| Position | Syllable weight | Register |
|----------|-----------------|----------|
| 1 | Heavy | H |
| 2 | Heavy | L |
| 3 | Light | H |
| 4 | Light | L |

In the category 'Heavy' we include all syllables that have a coda of any sort ($V:, V_l, V_r, V_m, V_n, V_\eta, V_p, V?, V_k$). In the category 'Light' are short nucleus, codaless syllables (V).

We next observe that not only does the stress derivation respect the constituents left by MS1 (in the sense that the stress head must fall within the initial binary foot left by the operation of (7)), but also that the

tonal properties partly determined by MS1 ((7)g.-i.) are included in the hierarchy (13). This again indicates that MS2 is a process subsequent to MS1's operation.

The stressability hierarchy (13) is exactly analogous to the more familiar stressability hierarchies in languages that determine eligibility for line 0 projection based on the simple distinction light/heavy, where heavy may be either CVV/CVC, or just CVV. Furthermore, just as the distinction between heavy and light in the simpler stressability hierarchies is expressed in metrical representation by the initial choice of eligibility for line 0 projection, in the case of the more refined stressability hierarchy (13), the distinctions are expressed in terms of: non-projection to line 0, projection to line 0, and head marking on line 1. It is an interesting confirmation of bracketed grid theory as developed in HV that exactly the mechanisms required to express the elaborated hierarchy in (13) are indeed available.

The observation that stress eligibility is restricted to metrified syllables subsequent to MS1 predicts that in those trisyllabic words that, due to their morphological structure, have a prespecified line 1 asterisk, the stress head must be identical with the MS1 head, and this prediction is correct - stress must fall initially in these words (cf. (9)c.-d. above). Keeping in mind that in all other cases stress is effectively restricted to the initial two syllables (by the prohibition on stress falling on

syllables left unmetrified by (7)), we get the following assortment of head positions:

(14) Head Positions on Word Initial Binary Foot

| S1 | S2 |
|----------|----------|
| * | |
| Long, H | Long, H |
| | * |
| Long, L | Long, H |
| | * |
| Short, H | Long, H |
| | * |
| Short, L | Long, H |
| * | |
| Long, H | Short, H |
| * | |
| Long, L | Short, H |
| * | |
| Short, H | Short, H |
| | * |
| Short, L | Short, H |

In table (13), we see all possible tone/syllable arrangements in word-initial syllables. Stress is indicated by an asterisk above an entry. Note that two entries in the stress hierarchy (14) are absent from S2 position: Long L and Short L. This systematic absence is due to the fact that stress operates on the output of MS1 (7), and by (7)i. ([+stiff] insertion), L register is disallowed in S2, where S2 is a head (in the sense of MS1).

2.3.3 Parameters of Noun Stress (MS2)

Since the head from MS1 is always S2 in the cases above (the only exceptions to this generalization are the trisyllabic words with exceptional morphological structure, see discussion above), we note that there are a number of cases in (14) of non-isomorphism between MS1 heads and MS2 heads (i.e., wherever the S1 entry has an asterisk in (14)). This means that a separate stress plane is required for MS2. We cannot wipe out the stress plane from MS1, because we will use some of its constituent structure (we need to know which positions are unmetrified). The process and parameter settings are given below:

(15) Noun Stress (MS2)

- a. Unmetrified syllables from MS1 do not project on line 0.
- b. Short, L register syllables do not project on line 0.
- c. Short, H register syllables that are MS1 heads do not project on line 0¹⁶.
- d. Long, H register syllables are line 1 heads.

¹⁶ This is subject to the requirement that every word has a stress head: if the non-projection of a short, H register syllable, in conjunction with the stronger requirement that short, L register syllables do not project on line 0, would leave no projected syllables on line 0, then in this case only short, H register syllables project on line 0. This is a requirement analogous to conditions on extrametricality, where it is usually the case that an entire domain cannot be marked extrametrical.

e. Apply (7)b.-e. (MS1)

- (- Line 0 parameter settings are (+BND,+HT,right).
- Construct constituents on line 0.
- Mark heads of line 0 constituents on line 1.
- Line 1 parameter settings are (-BND,+HT,left).
- Construct line 1 constituents.
- Mark heads of line 1 constituents on line 2.)

Obviously there is a great deal of formal overlap between MS1 and MS2, and this provides further confirmation of the metrical approach to the tonal phenomenon (MS1). MS2 above can be thought of as constructing a metrical plane orthogonal to that constructed by MS1 for tonal adjustment and realization, by analogy with Tiberian Hebrew.

2.3.4 Noun Stress Derivations

(16) Sample Stress Derivations

(Note: where line 2 and line 3 unique heads coincide, only lines 0-2 will be shown)

a. 'Tibet + person, 'Tibetan'

type: long, L short, H

| | | |
|---------|--------|----|
| line 2: | * | |
| line 1: | (*) | |
| line 0: | (*) | . |
| word: | phö̀ö̀ | mi |
| | | |
| | t | t |
| | | |
| | r | r |
| | | |
| | L | L |

b. banner + iron, 'iron banner fixture'

| | | |
|---------|---------|---------|
| type: | long, H | long, H |
| line 2: | * | |
| line 1: | (*) | (*) |
| line 0: | (*) | (*) |
| word: | thuu | caa |
| | | |
| | t | t |
| | | |
| | r | r c |
| | | |
| | H | H fall |

c. cotton + robe, 'cotton robe'

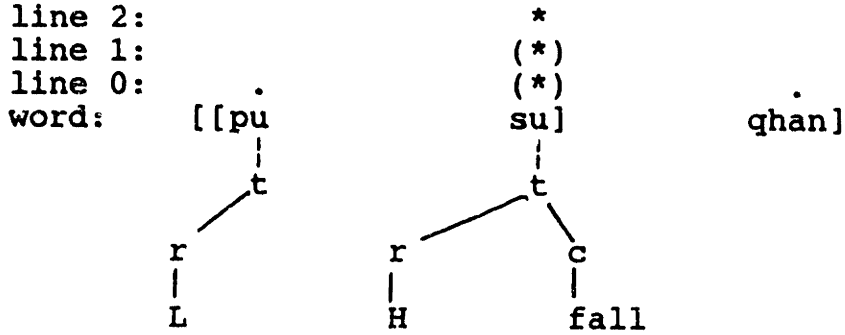
| | | |
|---------|---------|---------|
| type: | long, L | long, H |
| line 2: | | * |
| line 1: | (.) | (*) |
| line 0: | (*) | (*) |
| word: | ree | see |
| | | |
| | t | t |
| | | |
| | r | r |
| | | |
| | L | H |

d. mother + great, 'mother' (honorific)

| | | |
|---------|---------|---------|
| type: | long, L | long, H |
| line 2: | | * |
| line 1: | (.) | (*) |
| line 0: | (*) | (*) |
| word: | yum | chéé |
| | | |
| | t | t |
| | | |
| | r | r |
| | | |
| | L | H |

e. child + care + house, nursery

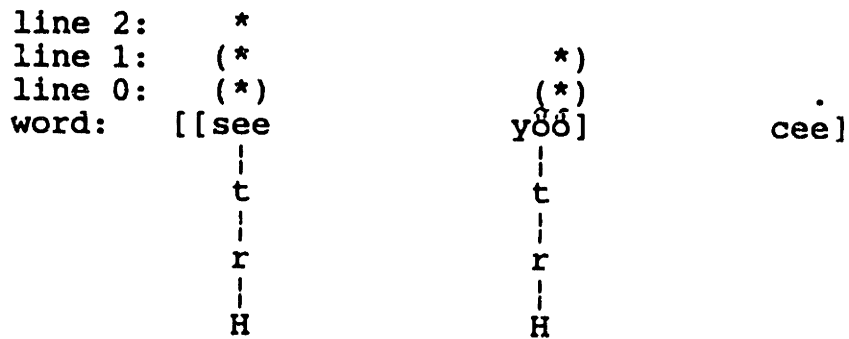
type: short, L short, H



In (16)e. above, we see the effect of short, L non-projection: because only [short, H] is left, the [short, H] must project (though by (15)c. it ordinarily would not) because otherwise the word will be left with no projection on line 0.

f. knowledge + possessor, 'intellectual'

type: long, H long, H



g. water + grape, 'grape'

type: short, H

line 2: *

line 1: (*)

line 0: (*)

word: [[chu] qu túu]

```

      |
      t
     / \
    r   c
    |   |
    H   fall
  
```

In (16)g., above, we see a case where the morphological structure upon which MS1 is based has determined that only the initial syllable is metrified, hence only it is eligible for stress, despite the fact that it is a [short, H], which normally does not project.

h. mother + great, 'mother' (honorific)

type: long, H . .

line 1: *

line 0: (*) * *

word: [[yum] chee mo]

```

      |
      t
      |
      r
      |
      H
     /
    c
    |
   rise
  
```

i. Dharamsala

```

type:      short, L      long, H

line 2:           *
line 1:  (.         *)
line 0:  (*         *)
word:  [ ta      ram      sa      la ]
        |         |         |         |
        t         t         .         .
        |         |         |         |
        r         r         |         |
        |         |         |         |
        L         H         |         |

```

j. H. H. the Dalai Lama

```

type:      long, L long, H

line 2:           *
line 1:  (.         *)
line 0:  (*         *)
word:  [ yii     sii     noo     pu ]
        |         |         |         |
        t         t         .         .
        |         |         |         |
        r         r         |         |
        |         |         |         |
        L         H         |         |

```

k. Buddha

```

type:

L2:      *
L1:  (*)
L0:  (*   *)
wrds:[ sa      ka      thu      pa ]
        |         |         |         |
        t         t         .         .
        |         |         |         |
        r         r         |         |
        |         |         |         |
        H         H         |         |

```

In (16)k., above, we see a case of [short, H] which is the MS1 metrical head not projecting, which throws the stress to the initial syllable by default. Note that although the

initial syllable is also [short, H], it is not covered by (15)c., because it was not a head from MS1.


In summary, we see that there are two metrical processes operating in the derivation of phonological properties of nouns: MS1, which provides the basis for certain processes of tone assignment and simplification, and MS2, the stress derivation, which relies on certain structure built by MS1, though it marks independent heads.

2.4 Head Shifting Determiners in NP Syntax

There are several determiners, notably /ciq/ (ཅིག) 'one', used as an indefinite article, and /ti/ (འདི) 'this' that have the effect of shifting the abstract MS1 head. This forces a recalculation of the tonal properties of a disyllabic N^0 to which the determiner may be prosodically attached. This is a very interesting effect, because it bears on the interaction of MS1 and MS2, as we shall see in the examples below. (We take the position here that /ciq/ and /ti/ are clitics, in the sense that they are contributed by the syntax but are prosodically dependent.)

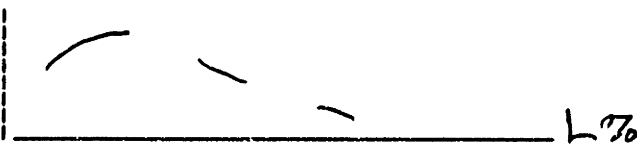
Consider the disyllabic noun 'Tibetan', shown in its post-MS1 and post-MS2 form below:

(17) Co-present Tone and Stress Metrics (MS1 and MS2)

| | | | |
|---------|---|----|-----|
| pitch: |  | | |
| line 1: | | * | MS1 |
| line 0: | (* | *) | MS1 |
| word: | phö | mi | |
| line 0: | (* | . | MS2 |
| line 1: | (* | | MS2 |
| line 2: | * | | MS2 |
| | | | |
| | t | t | |
| | | | |
| | r | r | |
| | | | |
| | L | H | |

In (17), the MS1 metrical structure, upon which tone is calculated, is in the grid above the word. The MS2 metrical structure, reflecting stress, is shown by the grid below the word. Tone and pitch are as shown, in citation form. Now consider what happens when the clitic /ciq/ (indefinite article) is applied (determiners are to the right in syntax):

(18) 'Tibetan' with determiner /ciq/

| | | | |
|--------|---|----|-----|
| pitch: |  | | |
| word: | phö | mi | ciq |

We observe in (18) that a high register rise now occurs on the initial syllable, quite distinct from its tonal shape

in bare N0 form above. In addition, the following syllables seem to have been tonally neutralized. We wish to consider whether this effect can be subsumed under any of the machinery motivated in the discussion of MS1 and MS2 above.

We will assume that there is a class of determiners, to which /ciq/ belongs, that can affect tone/accent. The exact property of this class is specified as follows:

(19) Properties of Head-shifting Determiners

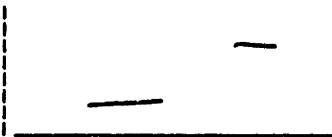
- Head-shifting determiners erase MS1 metrical structure and place a line 1 asterisk on the syllable that is MS2 head.
- Head-shifting determiners do (themselves) project to line 0 for either MS1 or MS2.

This may appear cryptic, but we will see that with this abstract formulation an otherwise arbitrary phenomenon falls into line without further stipulation. We will assume that by erasing MS1 metrical structure, (19) forces MS1 to reapply, now to the syntactic domain. Consider the effect of placing a new MS1 head: any syllable that is an MS1 head is eligible for full tonal specification, including contours and rise insertion (if low register). Furthermore, MS1 always allows the possibility of unmetrified post-head syllables, though that normally does not occur in disyllables (cf. discussion of MS1 above).

The regenerated MS1 metrical structure will appear as follows (recall that /ciq/ itself does not project to line 0):

(20) 'Tibetan' (a Tibetan person)

a. Standard compound form

pitch: 

| | | | | |
|---------|----|----|--|-----|
| line 1: | | * | | MS1 |
| line 0: | (* | *) | | MS1 |

word: phöö mi

| | | | | |
|---------|-----|---|--|-----|
| line 0: | (*) | . | | MS2 |
| line 1: | (*) | | | MS2 |
| line 2: | * | | | MS2 |

| | |
|---|---|
| | |
| t | t |
| | |
| r | r |
| | |
| L | H |

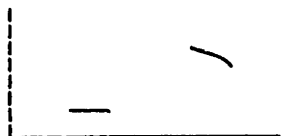
shown above to follow the initial application of MS1). Note that MS2 does not reapply.

Now consider a word with the same basic tonal pattern, but a different MS2 stress head location:

(22) 'goatherd' ɪˈθɜːd

a. 'goatherd' in isolation, without determiner

pitch:



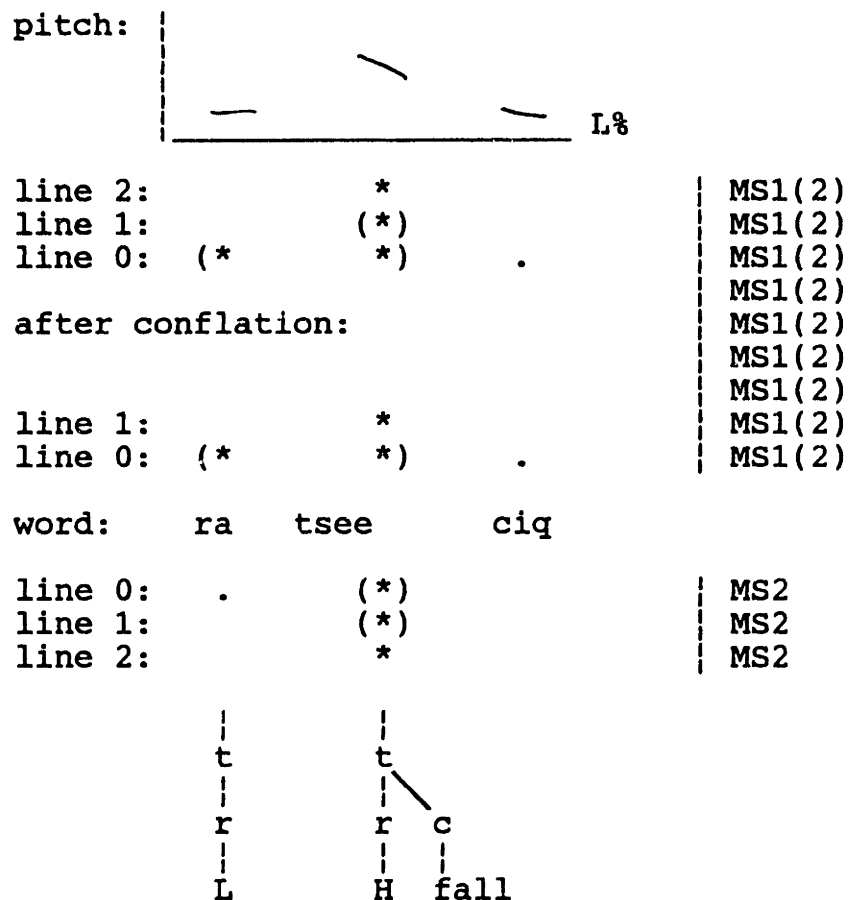
| | | | | |
|---------|----|---|--|-----|
| line 1: | | * | | MS1 |
| line 0: | (* | * | | MS1 |

word: ra tsee

| | | | | |
|---------|---|-----|--|-----|
| line 0: | . | (*) | | MS2 |
| line 1: | | (*) | | MS2 |
| line 2: | | * | | MS2 |

| | | |
|---|---|------|
| | | |
| t | t | |
| | | |
| r | r | c |
| | | |
| L | H | fall |

b. 'a goatherd', with reapplication of MS1:

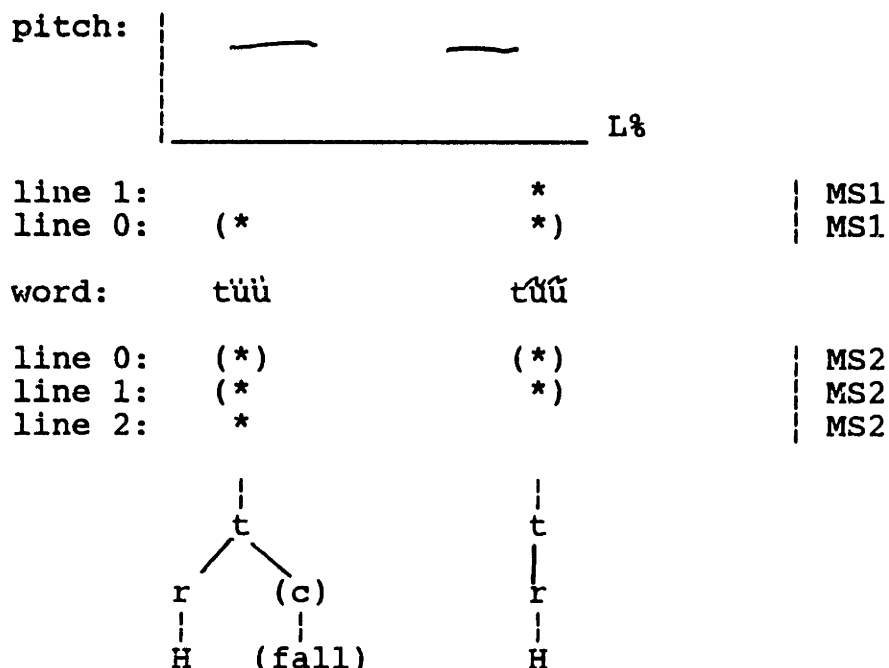


We see from the example above that in the case of 'goatherd', because the previous MS1 head happens to coincide with the MS2 head, the effect of the head-shifting clitic is vacuous. The reapplication of MS1 has the same result as the original application, as expected.

Let us consider one more interesting case of head-shift:

(23) 'shirt'

a. 'shirt' in isolation, without determiner:



We see above the standard citation form of /tuutuu/ 'shirt'. Note that the 'c' tonal subnode of the first syllable is shown in parentheses. This is because it is not realized (contours are obligatorily deleted in non-head syllables). However, there is a contour in the lexical specification of /tüü/ (the first syllable), and to draw attention to that fact that a contour was originally present and has been deleted by operation of MS1 in the derivation, we have included a parentesized 'c' tonal subnode in (23). Now consider the form with a determiner clitic:

b. 'shirt', with reapplication of MS1:

| | | | |
|---------|-----|-----|--------|
| pitch: | | | |
| line 1: | * | | MS1(2) |
| line 0: | (*) | * . | MS1(2) |
| word: | tüü | tũũ | ciq |
| line 0: | (*) | (*) | MS2 |
| line 1: | (* | *) | MS2 |
| line 2: | * | | MS2 |
| | | | |
| | t | | |
| | | | |
| | r | | |
| | | | |
| | H | | |

We see that the inclusion of /ciq/ in the domain has placed an accent for MS1 on the MS2 head, the initial syllable, and MS1 has reapplied (MS1(2)).

The form above gives evidence that the process is indeed a reapplication of MS1, rather than an initial application of MS1 over the entire domain (including the determiner). This is because the initial syllable, /tũũ/, appears without its lexical contour. Had MS1 applied to the whole domain (including head-shifting /ciq/) for the first time (in other words, had there been no prior application of MS1 to the bare form /tũũ/), we would expect the lexical falling contour to surface in (23)b., above. The fact that it does not surface is evidence that MS1 applied initially to /tũũ/, causing /tũũ/ to lose its contour (it was not

the head in the initial MS1 application). Thereafter, when the determiner is included in the domain of reapplication, the initial syllable becomes the head, and the subsequent syllables are detoned and interpolated over, all as expected, but the head syllable cannot recover the contour, because it has been lost on the initial application of MS1.

2.3.1 Verbal Complex Tone

2.3.1.1 The Verb Stem

The verb stems are all monosyllabic, and all stems are either long vowels or single vowels with a coda obstruent (simple V stems do not occur in RST). Tonologically, they divide into H and L register, as we saw with noun monosyllables. Within the registers, there are level and falling contours. The possible surface tonal shapes for verb stems are shown below:

(1)

| Register | Contour | Syllable type |
|----------|---------|--------------------------------|
| H | level | V:, Vl, Vr, Vm, Vn, Vŋ, Vp |
| H | fall | V:, Vm, Vn, Vŋ, Vp |
| L | rise | V:, Vl, Vr, Vm, Vn, Vŋ, Vp, Vk |
| L | fall | V:, Vm, Vn, Vŋ, Vp, Vk |

The heading 'syllable type' in (2) refers to surface realization, not orthography. The orthography of verb stems in relation to inflection and surface realization is a very complex area that will take us into matters far from our immediate phonological concerns. Suffice it for now to state that (1) represents the basic correlations, though there are scattered exceptions. We see from (1) that the only predictable feature of contour based on coda type is that [Vl] and [Vr], regardless of register, are non-falling. Furthermore, we note that [Vk] has been lost on the surface

in the RST high register verbs (unlike nouns), becoming [V:].

2.3.1.2 Metrical Dependencies

Verb stems never stand alone. The tonal shapes listed in (1) are determined on the basis of the behavior of the stems when suffixed by the infinitival affix -pa (). This brings us to the first metrical dependency in verb stem tone classification: the infinitival affix is metrically weak. Certain other affixes are strong. This classification is metrically characterized as follows:

(2) Metrical Properties of Affix Classes

- A strong affix projects on line 0.
- A weak affix does not project on line 0.

We can therefore classify 'pa', the infinitival affix, as a non-strong affix (about which nothing else need be said).

Let us now apply MS1, the metrical system motivated above for noun tone adjustment, to a minimal verbal complex: [verb stem + infinitival affix].

(3) Simple Infinitival Derivations

a.

stem type: H register, level

line 2: *

line 1: (*)

line 0: (*) .
yaa pa

'to lend' ɛʔɬɪ'ɬ'

b.

stem type: H register, falling

line 2: *
 line 1: (*)
 line 0: (*) .
 loo pa 'to read' ལོ་ཤིག་པ་

c.

stem type: L register, rising

line 2: *
 line 1: (*)
 line 0: (*) .
 tar pa 'to grind' བར་རྩ་པ་

d.

stem type: L register, falling

line 2: *
 line 1: (*)
 line 0: (*) .
 nöö pa 'to buy' རྩོམ་པ་

Since the infinitival suffix does not project on line 0, the infinitival system is transparently simple. Nevertheless, there are two interesting features that invite comment.

First, we see that according to MS1, only a metrical head can have a full tonal geometry (register and contour). The stems in each of the representations above are metrical heads, and therefore they are licensed to carry both register and contour contrasts, and they do so.

Second, consider (3)c. Here the stem type is given as [L register, rising]. But we recall from the discussion of nouns above that there is a process of Rise Insertion that applies to domain initial, L-register metrical heads. In (3)c. we have exactly that situation, so we will take the rise exhibited on the surface in (3)c. to be the result of Rise Insertion. Again, subsequent [+stiff] insertion on metrical heads ((7)g. (MS1)) will neutralize the register distinction in the verbal complex, however since rises can only be derived from low register stems, the contrast between H and L register stems is preserved, in a sense. Therefore, the symmetric tonal system below will be taken as the underlying classification:

(4) Underlying verb stem contours

| Register | Contour | Syllable types |
|----------|---------|--------------------------------|
| H | level | V:, V1, Vr, Vm, Vn, Vη, Vp |
| H | fall | V:, Vm, Vn, Vη, Vp |
| L | level | V:, V1, Vr, Vm, Vn, Vη, Vp, Vk |
| L | fall | V:, Vm, Vn, Vη, Vp, Vk |

2.3.1.3 Verb Stem Alternations: Level/Fall

There is a basic alternation in the contours of the two falling tone classes (high and low register) in the stem table (4) above. Before certain affixes, the falls are realized as level, in both high and low register. This is a salient feature of the RST verbal complex, and we would wish to capture it in a principled fashion, on the phonological

level if possible, with the minimum appeal to morphology, syntax, or lexical marking and special classification. In fact, a very simple description is available, based on the MS1 metrical process, independently motivated for nouns.

Let us consider, initially, the alternation in tonal form of falling contour verbs in two situations. In the narrative past tense, signalled by a particular affix/auxiliary verb combination, the falling contour stems exhibit the surface falling contour that we saw in the infinitive above. In the future tense, on the other hand, they exhibit level tone. This situation is outlined below:

(5) Verb stem tone alternations

| Verb | Gloss | Type | +paree (Narr.Past) | +qiree (Future) |
|------|--------|---------|--------------------|-----------------|
| loo | 'read' | H, fall | H, fall | H, level |
| nöö | 'buy' | L, fall | L, fall | L, level |

In 'paree' above, 'pa' is a verbal affix indicating past tense, while 'ree' is an auxiliary verb. Likewise, in 'qiree', 'gi' is a verbal affix and 'ree' is the same auxiliary verb. We will classify 'gi' as a strong affix, and we will adopt the following basic convention:

(6) Verbal Projection Categories

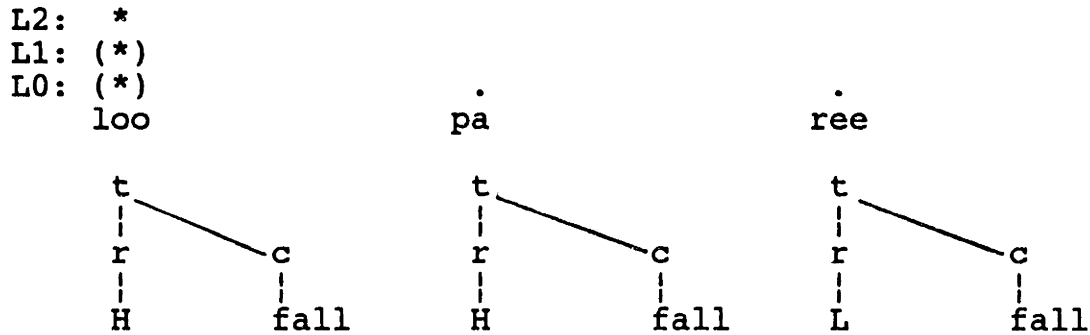
Of the verbal complex elements (1), only verb stems and strong affixes (marked in the lexicon) project on line 0 for MS1.

Then we apply MS1 to the verbal complex.

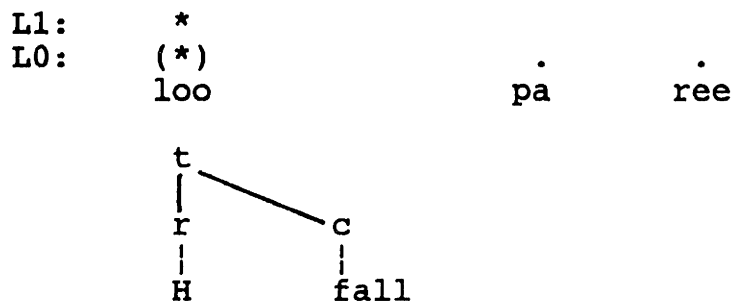
2.3.1.4 Verbal Complex Tone Derivations

(7) Verbal Complex Derivations

a. 'read' + pa + ree (narrative past)

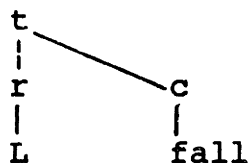


In (7)a. above, we see the projection of the verbal stem and the non-projection of the other elements. Applying MS1, we get the head marked (trivially) as shown. The post-head elements, 'pa' and 'ree' are obviously unmetrified (since they never projected), hence their tonal specifications are delinked in accordance with (7)h. (MS1), resulting in the representation below (after conflation):

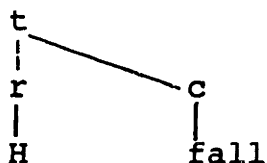


c. 'buy' + pa + ree (narrative past)

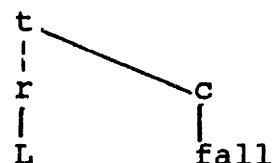
L2: *
L1: (*)
L0: (*)
nöö



pa



ree



In the tonal adjustment process, the result of which is depicted below, we observe that Rise Insertion takes precedence over previously existing contour specification, resulting in a fall (licensed by head status of its syllable) being replaced by a rise (Rise Insertion):



L1: *
L0: (*)
nöö

pa

ree

t

r

c

H

rise

entry says nothing about any of this. Affixes are labelled 'strong', or left unmarked. The strong affixes, by virtue of their marking, will cause a perturbation in the metrical system that entails, by the independently motivated provisions of MS1, the type of alternation that occurs.

(8) Some Strong Affixes

- qaa '(go) for purpose (of V)'
- roo 'please V'
- tuu 'when V'
- suu(riqaa) 'while V'
- Ntoo 'just finished V'

(9) Some Weak Affixes

- tsaa 'because, since V'
- ni 'after V'
- cee 'having V'
- qoo 'before V'
- pa(taqa) 'just finished V'

2.3.1.6 The Honorific Auxiliary Verb [nan]

For many verb stems, the verbal complex has an honorific form with the auxiliary verb /nan/ (ጥጥፍ)¹. This auxiliary verb occurs to the right of the verb stem. Examples are given below in the narrative past form:

¹ Certain verbs have an independent honorific stem form. These will not be discussed here.

(10)

| | | |
|-------|------------------|------------------|
| /sòò/ | /soo naŋ pə ree/ | 'made (food)' |
| /yaa/ | /yaa naŋ pə ree/ | 'lent, borrowed' |
| /pep/ | /pep naŋ pə ree/ | 'dismounted' |
| /tǎǎ/ | /tǎǎ naŋ pə ree/ | 'sent' |

/naŋ/ lexically bears high level tone, which appears as such on the surface. Furthermore, the class of verb stems that show level/falling tonal contour alternations (see discussion above) appear with falling contour preceding /naŋ/. We thus have an unexpected situation: on the one hand, /naŋ/ can be considered a strong affixal element in the sense that it occurs to the right of the verb stem, it retains its high level tone, and it is phonologically the most prominent element in the verbal complex. On the other hand, however, the alternating verb stems take the falling contour form preceding /naŋ/, where level contour would be expected, if /naŋ/ is truly a strong affixal element.

This situation is modeled in metrical analysis by allowing /naŋ/ to usurp the verb stem's place in the metrical verbal complex constituent. If /naŋ/ is regarded as a substitute verb stem, its behavior is no longer anomalous. First, verb stems are expected to retain their full lexical tone, as /naŋ/ does. Second, elements following the stem are effectively detoned, as we saw above, and this effect occurs to the right of /naŋ/, as with a standard verb stem in a verbal complex without /naŋ/. Third, the view of /naŋ/ as a

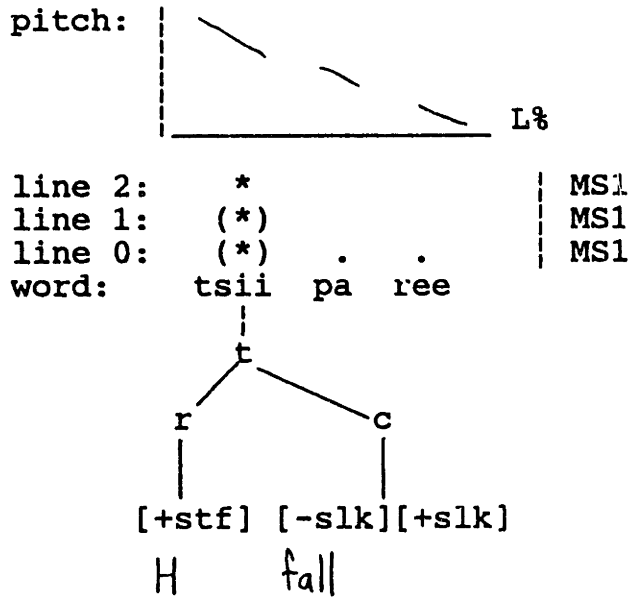
prosodic substitute for the verb stem in the prosodic verbal complex explains its prominence for stress, because, as we will see below, the verb stem is exceptionlessly the most prominently stressed verbal complex element. Finally, the contour property of the actual verb stems that precede /naŋ/ is explained: the loss of contour does not occur because the verb stems and /naŋ/ are not in a single metrical constituent at the lowest level of analysis.

There are various ways to think about the exact formal characterization of this observation (that /naŋ/ usurps the metrical position of the verb stem). One simple approach is to regard /naŋ/ as subcategorized for a line 0 left metrical bracket. The lexical specification of metrical heads is well-motivated, as shown by the use of this mechanism in the majority of stress systems surveyed by Halle and Vergnaud 1987. Heads that are so marked must be respected by the metrical processes based on generalized parameter settings. The necessity for an analogous mechanism with respect to lexical marking of brackets, in addition to heads, is demonstrated by Halle 1990. We are adopting this formal possibility here, so that /naŋ/ will be said to be lexically subcategorized for a left metrical constituent bracket. This will entail the relevant representational distinction in the derivation through the normal operation of MS1, as will be shown below. The lexical marking approach is further justified by the generally exceptional lexical status of

/naŋ/ (it is honorific; there is no other verb with its placement and function; it has some semantic content, with a sense of 'give'; etc.) Some sample derivations will clarify the analysis. The first two derivations below are non-honorific, for comparison. Note that since conflation of the MS1 lines 0-2 will not matter to the outcome of these cases, the representations are shown prior to application of conflation, for clarity:

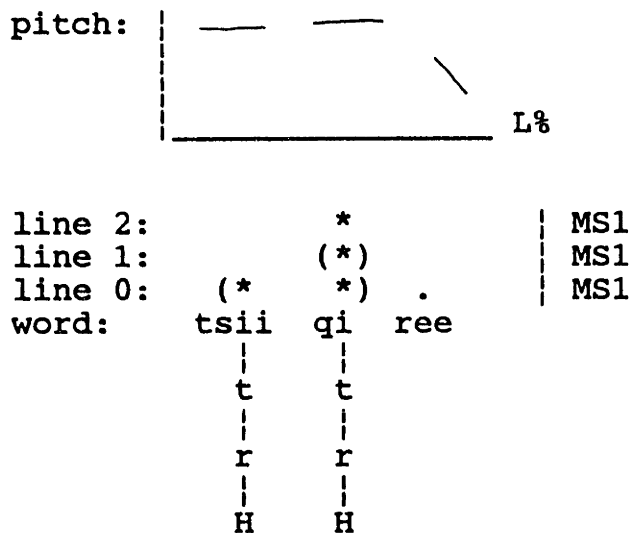
(11)

a. 'to count' non-honorific, narrative past (weak)

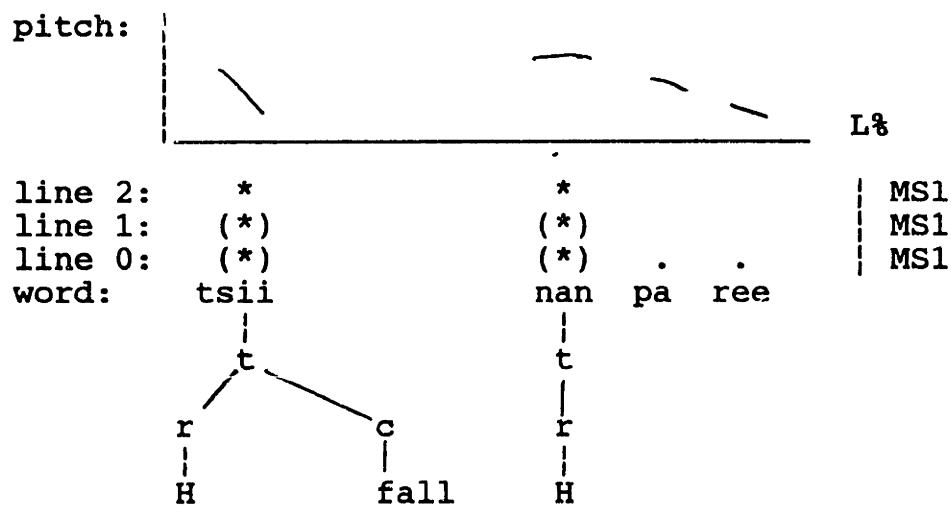


Above we see the narrative past verbal complex with verb stem /tsii/, 'to count'. Since the narrative past affix /pa/ is weak in the sense discussed in the section on affix strength above, it does not project to line 0, and MS1 makes the verb stem the metrical and tonal head. It thus surfaces with its lexical high register falling contour intact.

b. non-honorific, present progressive (strong)



c. honorific, narrative past (weak)



We observe that in (11)c. above the verb stem /tsii/ and /nan/ form two separate metrical constituents on lines 1 and 0, and that therefore the tonal adjustments of MS1 do not apply, and /tsii/ retains its lexical high falling contour. The representation in (11)c. is thus distinct from that of (11)a., where there is no separate line 0 constituent induced by the lexical left bracket carried by

The tonal configuration of Ca is just what would be expected if /caa loo/ had been independently fed to MS1. The lexical (low) falling contour on /caa/ has been lost, and /loo/ is high register ([+stiff]), and, by virtue of its status as a metrical head, retains its contour.

In summary, the metrical representation of tonal processes in RST affords us a simple and natural mechanism for representing the exceptional properties of the honorific verb.

2.3.1.7 Negation in the Verbal Complex

The negative element /ma/ (\overline{a}) in the verbal complex occurs pre-verbally, in accordance with the following schema:

(13) Verbal Complex Template

Neg - Verb stem - Affixes/particles - Aux

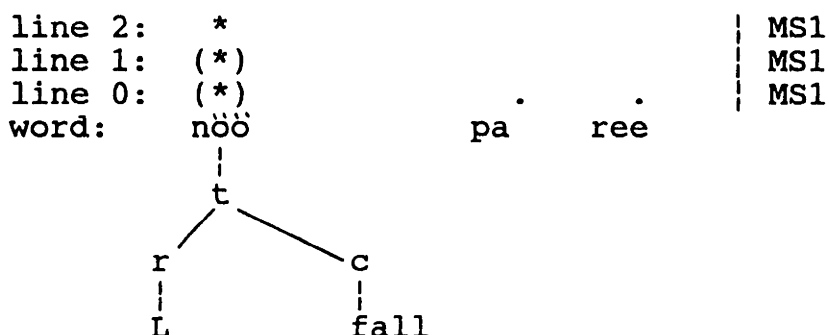
/ma/ has certain interesting tonal and metrical properties, all of which can be accounted for naturally in the general framework adopted here, and which in turn shed further light on certain details of this framework.

The first observation is that /ma/ surfaces with the same register specification (low or high) as the verb stem (immediately to its right). We note further that the verb

stem in negative verbal complexes always surfaces with high register. However, a falling contour that may have been present as part of the verb stem's lexical representation is retained. This complex of properties is almost what we expect if MS1 applies to the template (13), with the proviso that a process of register copy must crucially precede the assignment of [+stiff] register to the MS1 head. A sample derivation is given below (the first derivation is non-negative, for comparison):

(14)

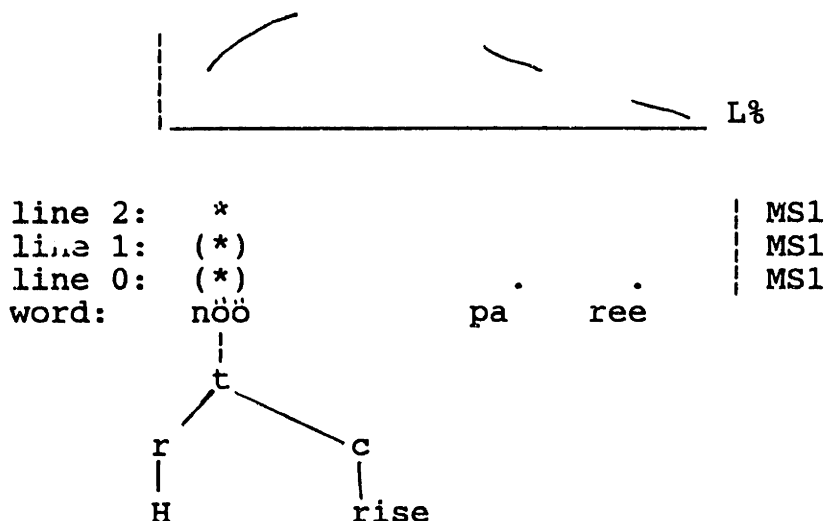
a. 'bought' non-honorific, narrative past (weak)



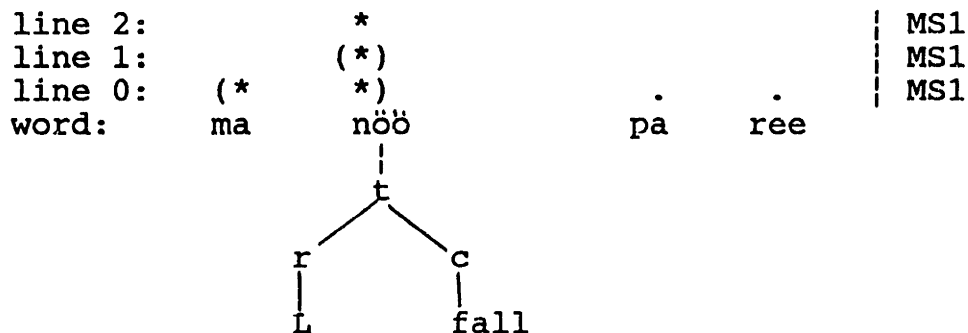
In (14)a. above, we see the initial MS1 structure. The tonal specifications for /nöö/, the verb stem, are as provided by the lexicon. When the tonal adjustment provisions of MS1 come into play, Rise Insertion will change the falling contour of /nöö/ in (14)a. to a rise, and [+stiff] insertion will change the register, resulting in a high register rise on the surface:

b.

pitch: |



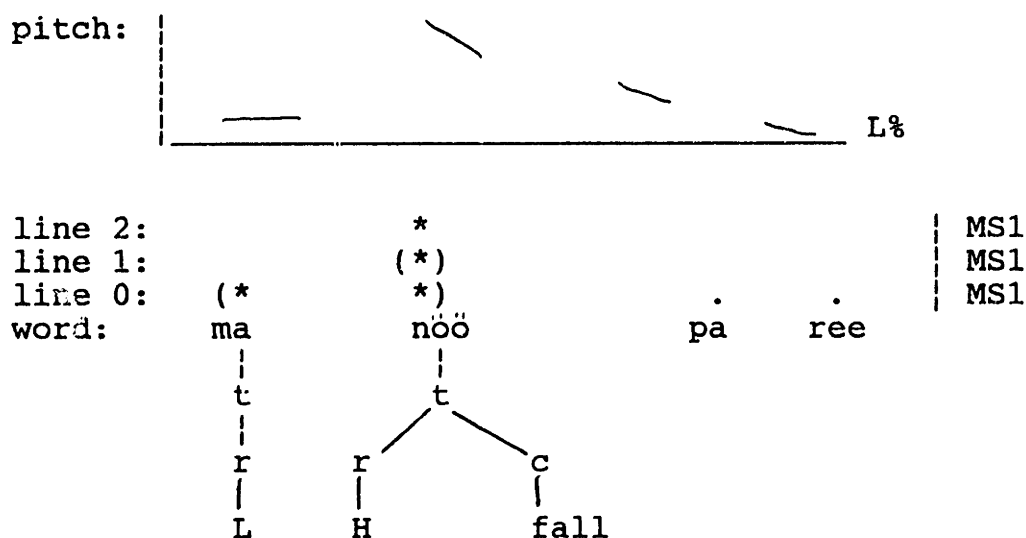
c.



In (14)c. above, we see the beginning of a derivation with the negative /ma/ element. There are two possibilities for the underlying register specification of /ma/. One possibility is that /ma/ is underlyingly unspecified for register. Another is that /ma/ is underlyingly low register ([-stiff]) or high register ([+stiff]). For the moment, assume that /ma/ underlyingly has no register tonal subnode. Then the first step of tonal adjustment in negative verbal complexes, following the purely metrical construction of MS1 constituents above is to copy the 'r' tonal subnode of the verb stem to /ma/. After the copy is made, the 'r'

specification of the verb stem is changed to [+stiff], by the usual provisions of MS1 tonal adjustments, and the falling contour of the verb stem in (14)c. is retained (it is the MS1 head). The representation below depicts the result of the two operations (register copy, then [+stiff] insertion on the verb stem):³

d.



Note that in (14)d. above, Rise Insertion is blocked, as the relevant environment does not exist.

The case of negation is of interest not only because we once again see MS1 operating in a syntactic and morphological domain far from the evidence that originally motivated it (disyllabic compound nouns), but also because it gives evidence for the synchronically active and

³ It will be observed that a spread operation could do the job of the copy operation here. There is no particular evidence favoring one approach over the other. In the case of spread, the mechanics must simply guarantee that subsequent [+stiff] insertion on the verb stem does not affect the register of /ma/ after spread occurs.

independent character of [+stiff] insertion as a process, given that an independently motivated copy process crucially precedes it.

2.3.2 Verbal Complex Stress

Like the nouns, the verbal complex has a stress marking system that is distinct from the tonal metrics. The verbal complex stress works according to the following parameter settings:

2.3.2.1 Verbal Complex Stress Parameters

(11) Verbal Complex Stress

- a. Certain weak affixes do not project to line 0.
- b. All other syllables project to line 0.
- c. A line 1 asterisk is placed on the verb stem.
- d. Line 0 parameter settings are:
(+BND,+HT,right,left to right).
- e. Construct line 0 constituents.
- f. Place heads of line 0 constituents on line 1.
- g. Line 1 parameter settings are: (-BND,+HT,left).
- h. Construct line 1 constituents.
- i. Place head of line 1 constituent on line 2.

Note that essentially everything in the verbal word projects for stress (except certain marked affixes). Unlike the noun system, where the MS2 noun stress marking respects the constituents left by MS1 in the sense that unmetrified positions are not eligible for stress, in the verbal complex

essentially every element is eligible for stress and the constituents left by MS1 are irrelevant. For the basic inflectional system, only 'pa', the narrative past tense affix, does not project on line zero. Sample derivations are given below (tone is not shown in the derivations because it is not relevant to verbal stress):

Another difference between the noun stress system (MS2) and the verbal stress system formalized above is that conflation does not apply in the derivation of verbal stress, and secondary stresses are salient.

(12) Verbal Complex Stress Derivations (MS3)

(all derivations use verb stem 'yaa',
[H register, level], 'to lend')

a. 1st P., past

line 2: *
line 1: (* *)
line 0: (*) . (*)
V complex: yaa pa yíí

b. 2nd P., past, interrogative⁴

line 2: *

line 1: (* *)

line 0: (*) (*)

V complex: yaa pee

c. 3rd. P., direct past, interrogative

line 2: *

line 1: (* *)

line 0: (*) (*) *

V complex: yaa son nee

d. 3rd. P., direct past

line 2: *

line 1: (* *)

line 0: (*) (*)

V complex: yaa son

e. 3rd. P., narrative past, interrogative

line 2: *

line 1: (* *)

line 0: (*) . (*) *

V complex: yaa pa re pee

f. 3rd. P., narrative past

line 2: *

line 1: (* *)

line 0: (*) . (*)

V complex: yaa pa yíí

g. 1st. P., present perfect

line 2: *

line 1: (* *)

line 0: (*) (*)

V complex: yaa yòò

⁴ Note that questions in the first person and statements in the second person use the third person forms.

h. 2nd. P., present perfect, interrogative⁵

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa yo pee

i. 3rd. P., direct present perfect, interrogative

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa tu qee

j. 3rd. P., direct present perfect

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa saa

k. 3rd. P., indirect present perfect interrogative

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*)
 V complex: yaa yöö re pee

l. 3rd. P., indirect present perfect

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa yo ree

m. 1st. P., present

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa qi yöö'

n. 2nd. P., present interrogative

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*)
 V complex: yaa qi yo pee

⁵ The particles 'yoo' (ཡོ་) and 'ree' (རེ་), though underlyingly long, are shortened if they occur in a position with stress below line 1.

o. 3rd. P., direct present interrogative

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*)
 V complex: yaa qi tu qee

p. 3rd. P., direct present

line 2: *
 line 1: (* *)
 line 0: (*) (*)
 V complex: yaa qii

q. 3rd. P., indirect present interrogative

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*) (*)
 V complex: yaa qi yoo re pee

r. 3rd. P., indirect present

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*)
 V complex: yaa qi yo ree

s. 1st. P., future

line 2: *
 line 1: (* *)
 line 0: (*) (*) (*)
 V complex: yaa qi yii

t. 2nd. P., future interrogative

line 2: *
 line 1: (* *)
 line 0: (*) (*) (*)
 V complex: yaa qi pee

u. 3rd. P., future interrogative

line 2: *
 line 1: (* * *)
 line 0: (*) (*) (*)
 V complex: yaa qi re pee

v. 3rd. P., future

```
line 2:      *
line 1:      (*      *)
line 0:      (*) (*  *)
V complex:  yaa qi ree
```

Chapter 3

Lexical and Phrasal Stress in Beijing Mandarin

3.1 Introduction

The Beijing dialect of Mandarin Chinese (henceforth BMC) is considered to have stress as well as tone (Hockett 1947, Rygaloff 1956, Xu 1960, Kratochvil 1969 1974, Ho 1976, Hoa 1983). It therefore represents a potentially fruitful domain of inquiry into questions of the interaction of these phonological entities. The following discussion presents a novel analysis of BMC stress, showing that the stress properties of lexical items are the result of interactions between syllable tone and parametrized metrical constituent structure. It is argued that the four syllable-based tones of BMC form a hierarchy of stressability that is reflected in metrical structure, in a fashion analogous to stressability hierarchies based on syllable weight in certain languages (e.g. Hindi, Sierra Miwok, and many others). The metrical formalism employed for this analysis is that of Halle & Vergnaud 1987. Although familiarity with Halle and Vergnaud 1987 is helpful in following the mechanics of the derivations, the essential character of the analysis should be accessible to those familiar with alternative systems of metrical prominence labelling. Finally, it must be understood that the facts as analyzed here are representative of the Beijing city dialect of

Chinese only. While other varieties of Mandarin may also exhibit lexical stress, any similarity between the stress system presented here for BMC and the facts of other dialects is fortuitous.

A note on the data and associated notation is in order. All data that is marked with the sign [sm] results from the author's collaboration with two native speakers of BMC. However, the bulk of the data discussed here is from Hoa 1983, a seminal and exhaustive study of BMC stress, and no special marking is given for data from that source. The stress as marked over examples represents the judgements resulting from the author's collaboration with two native speakers of BMC. Where these judgements differ from the stress patterns reported by Hoa 1983 in important or relevant ways, the discrepancy will be noted in accompanying discussion. All data is given in Pinyin standard transcription.

3.2 The Tone/Stress Correlation

3.2.1 The Tones of Beijing Mandarin

Mandarin Chinese has four lexical tones, the phonetic specifications of which are shown below (cf. Shih 1987):

(1)

| Number | 1st target | 2nd target | 3rd target | Shape |
|--------|------------|------------|------------|----------|
| Tone 1 | H | H | H | level |
| Tone 2 | M | M(-) | H | rise |
| Tone 3 | M | L | M | dip-rise |
| Tone 4 | H | H+ | L | fall |

Note that the specifications in table (1) above would not be acceptable as phonological representations (due to the symbols 'M' and -/+). Our purpose here is simply to give a concrete picture of the tones' phonetic shapes. In table (1) above, the 1st target is an indication of relative pitch level at the initial rime margin (the tone bearing unit of BMC will be taken to be the syllable rime), the 2nd target is a relative pitch specification for approximately the mid-point of the syllable, and the 3rd target is a specification for the final margin.

3.2.2 The Relative Strength of Tones

The four tones of BMC, given in terms of phonetic targets in (1) above, form a hierarchy for stress which can be expressed by the following descending scale, based on the 'Number' column in (1):

(2) Stressability Hierarchy

4 > 1 > 2 > 3

The relative strength of tones 4 and 1 is indeterminate, but each of them is stronger than 2 or 3. This indeterminacy is one reason for appealing to abstract metrical structure in BMC stress derivations, as will be seen further below. The effect of (2) can be seen with a few simple examples, where lexical main stress is indicated with an asterisk (the numbers appearing just above syllables refers to the tonal categories in table (1)):

(3)

| Word | Gloss |
|-----------------------|------------------|
| * 1 2 gong ping | 'fair' |
| * 4 2 xing fu | 'fortunate' |
| * 1 3 si xiang | 'thought' |
| * 2 3 ya chi | 'teeth' |
| * 4 3 chi bang | 'wings' |
| * 3 4 nu li | 'hard-working' |
| * 3 2 jian cha | 'to investigate' |
| * 2 1 guo jia | 'country' |

For the simple cases given above, the stressability hierarchy (2) can be directly applied to derive the main stress. But there are numerous more complex examples that show that the stressability hierarchy as presented in (2) is only one component of the BMC stress system. (2) interacts with a battery of metrical parameter settings that apply to construct a bracketed grid as the fundamental prosodic representation of a word (zero level category in the X-bar system).

3.2.3 Formalization of Stress

The statement (2) is adequate for some disyllabic items, but in considering longer strings we require a more general formulation. The basic system is formalized in (4) below. (4) derives a bracketed grid representation of BMC stress, employing the metrical formalism of Halle and Vergnaud 1987. In (4) below, we use the term 'tone n', $1 \leq n \leq 4$, as short-hand for 'syllable bearing lexical tone n'. This notation ('tone n') is sometimes abbreviated in discussion below to 'Tn'.

(4) Mandarin Lexical Stress Marking (MSR)

- a. All tones project to line 0¹.
- b. Mark tones 1 and 4 as heads on line 1.
- c. Line 0 parameters are (-BND,+HT,left).
- d. Construct line 0 constituents.
- e. Mark heads of line 0 constituents.
- f. Line 1 parameters are (-BND,+HT,right).
- g. Mark head of line 1 constituent.
- h. Tone 3 line 0 asterisk deletes when adjacent to tones 1, 2, or 4.

The result of (4) is a bracketed metrical grid. After application of (4), the primary stressed syllable in a lexical item is that syllable which projects as a head on line 2, and relations of non-primary stress are indicated by head placement on successively lower lines. Provision (4)h. captures the fact that tone 3 is the weakest tone of the stressability hierarchy. When the line 0 asterisk of a syllable bearing tone 3 deletes, that syllable can no longer participate in a metrical constituent on any line, and the constituent boundary is adjusted in accordance with the parameter settings (4).

In the derivations given below, we will see that while (4) represents the basic system, certain minor processes

¹ This means that tone bearing syllable, regardless of the tone quality it bears, is represented by a line 0 asterisk in the initial grid configuration.

remain to be discussed and incorporated. Let us now consider how the system works in some typical cases (examples will be shown of the grids after application of (4)²:

3.2.4 Derivations

(5) Derivations

a. 2 syllable words

i.

| | | |
|---------|---------|--------|
| line 2: | | * |
| line 1: | | (*) |
| line 0: | . | (*) |
| tone: | 3 | 1 |
| word: | ya zhou | 'Asia' |

ii.

| | | |
|---------|----------|---------------|
| line 2: | | * |
| line 1: | (* | *) |
| line 0: | (* | *) |
| tone: | 1 | 1 |
| word: | bei guan | 'pessimistic' |

iii.

| | | |
|---------|--------|-----|
| line 2: | | * |
| line 1: | (* |) |
| line 0: | (* |) . |
| tone: | 2 | 3 |
| word: | ya chi | |

² In particular the metrical grids are shown after the application of (4)h., which provides for the destressing of syllables bearing tone 3. These are represented by a period (.) after destressing and constituent adjustment.

iv.

```

line 2:          *
line 1:          (*)
line 0:          (*  *)
tone:           2  2
word:          xue xi  'to study'

```

Note that in (5)a.iv., the left-headedness of the item is given 'for free' as it were by the metrical settings in (4). Were we to attempt to mark the stress of items like iv. above directly from the tone hierarchy, we would be required to stipulate that tone 2 on the left edge is stronger than tone 2 on the right edge.

Likewise, returning for a moment to example ii. above, 'bei guan', we would have to stipulate that for tone 1, just the opposite is true: tone 1 on the right edge is stronger than tone 1 on the left edge. A similar statement would be required for tone 4, as we will see in detailed discussion below. All these generalizations fall out from the metrical settings in (4), which also handle the more pedestrian cases (disyllables with distinct tones) uniformly³.

v.

In the three words below, we see the domain of potential application of a minor stress process in BMC. This process

³ Note that the sequence tone 3 tone 3 is forbidden by an independent sandhi process which changes the initial tone 3 to tone 2 in such a sequence. The stress mechanism in (4) is assumed to apply subsequent to the operation of this sandhi process.

results in derivation of greater stress on the syllable containing a high front tense vowel [i] in words that exhibit only tone 4 (high falling tone, the strongest of the hierarchy in (2)). The rule is that in a word containing only tone 4, any syllable that does not contain [i] must be metrically subordinate to any syllable that does contain [i], if any such exists. If no syllable containing [i] exists, the stress system works as in all cases above, likewise, if both syllables contain [i], the stress system works as above (i.e., the rightmost syllable will take the greatest stress). The derivations proceed by deletion of a line 1 asterisk above a syllable that does not contain [i] when it is adjacent to one that does (all bearing tone 4). In the first example below, the exceptional strength of a tone 4 syllable with [i] is vacuous in its effect on the derivation, in the sense that the main stress would be derived correctly without deletion of the line 1 asterisk above /kang/. In the second example, we observe that the asterisk above the final syllable, which lacks [i], is deleted, leaving the main stress to fall on the initial syllable. Finally, in the last example, we again see that no special action is required (i.e. the deletion of the line 1 asterisk above /zha/ is vacuous), because in the absence of [i] in either syllable, (4) operates as expected:

| | | |
|---------|---------|--------------|
| line 2: | | * |
| line 1: | . | (*) |
| line 0: | (*) | (*) |
| tone: | 4 | 4 |
| word: | kang yi | 'to protest' |

line 2: *
 line 1: (*) .
 line 0: (*) (*)
 tone: 4 4
 word: ji hua 'to plan'

line 2: *
 line 1: (*) (*)
 line 0: (*) (*)
 tone: 4 4
 word: zha dan 'bomb'

b. 3 syllable words

i.

line 2: *
 line 1: (*) (*)
 line 0: . (*) (*)
 tone: 3 4 1
 word: Ma ke si 'Marx'

ii.

line 2: *
 line 1: (*) (*)
 line 0: (*) (*) .
 tone: 4 1 3
 word: lei xing ti 'quasi-stellar object'

iii.

line 2: *
 line 1: (*)
 line 0: (*) * *
 tone: 2 2 2
 word: chen pi mei 'preserved prune'

In 'chen pi mei', the syllables 'pi' and 'mei' have approximately equal stress, as indicated by the grid.

iv.

```

line 2:           *
line 1:         (* *)
line 0:         (*)(* *)
tone:           2 4 2
word:           Ji bu ti      Djibouti

```

Note that in (5).b.iv., above, the initial syllable bears greater secondary stress than the final syllable, with the same tone (T2). This is indicated by the line 1 head markings.

v.

```

line 2:           *
line 1:         (* * *)
line 0:         (*)(*)(*)
tone:           4 4 4
word:           Nuo man di    Normandy

```

This example is interesting because it illustrates a common effect of metrical structure on tone in BMC (up to now we have only considered the effect of tone on metrical structure). The tonal realization of v. 'Normandy' substitutes a level tone for the falling tone on /man/:

```

Tone shape:      \   \   \      ->  \   -   \
Tone number:     4   4   4      ->  4       4
Word:           Nuo man di      Nuo man di

```

This effect can be accounted for by a process of line 1 asterisk deletion, for line 1 asterisks that are surrounded on both sides by line 1 asterisks, as in the case above. The

process will be named 'Medial Deletion', and is formalized below, where the superscripted 'n' means 'any number of':

Medial Deletion (MD)

line 1: (* *ⁿ *) -> (* .ⁿ *)

We will then assume that syllables to which MD have applied are supplied with a level contour tone which overrides their lexically specified tone. Therefore, although such syllables retain their line 0 projection, they differ from other line 0 projected syllables in that MD has applied to them.

c. 4 syllable words

i.

| | | | | | |
|---------|----|----|-----|-----|----------|
| line 2: | | | | * | |
| line 1: | (* | . | * | *) | |
| line 0: | (* | *) | (*) | (*) | |
| tone: | 4 | 2 | 4 | 1 | |
| word: | Bu | da | pei | si | Budapest |

ii.

| | | | | | |
|---------|-----|-----|----|-----|----------|
| line 2: | | | | * | |
| line 1: | . | (* | . | *) | |
| line 0: | . | (* | *) | (*) | |
| tone: | 3 | 1 | 2 | 4 | |
| word: | Tan | san | ni | ya | Tanzania |

iii.

| | | | | | |
|---------|----|-----|-----|----|--|
| line 2: | | | | * | |
| line 1: | (* | . | * | .) | |
| line 0: | (* | *) | (* | *) | |
| tone: | 1 | 2 | 4 | 2 | |
| word: | Wu | lan | nuo | wa | |

d. 5 syllable words

i.

```

line 2:                *
line 1:      (* *      *)
line 0:      (*) (* * .)(*)
tone:        1  4  2  3  4
word:        Ai sai e bi ya  Ethiopia

```

ii.

```

line 2:                *
line 1:      (* . . . *)
line 0:      (* * . *) (*)
tone:        2  2  3  2  4
word:        Mao li tan i ya  Mauritania

```

e. 6 syllable words

i.

```

line 2:                *
line 1:      (* * . * * *)
line 0:      (*) (*) (*) (*) (*) (*)
tone:        4  1  2  1  1  1
word:        Bu lə di si la fa  Bratislava

```

ii.

```

line 2:                *
line 1:      . (* * . * *)
line 0:      . (*) (*) . (*) (*)
tone:        3  4  1  3  4  1
word:        Ya di si ya bei ba  Addis-Abebbba

```

3.3 Alternative Approaches

It may be argued, in view of the close dependence of stress on tone in BMC, that no stress marking of any kind is

required. There are two possible variations on this theme, each of which is explored in a separate section below.

3.3.1 Universal Phonetics

The first version of the 'no stress' argument would hold that since stress is predictable from tone on the basis of the stressability hierarchy (2), and since the tone markings are independently required in lexical entries, the stress markings, whether lexically encoded or derived as in (4) are simply restating the tone specifications. We know immediately that this view is false, as there are two strictly metrical components to the stress system (i.e. stress effects depending on position in the word rather than rank in the stressability hierarchy). These are exemplified in the derivations above, and they are summarized as follows:

(6) Metrical effects and their relation to the stressability hierarchy

a. When identical tones exist in a word (whether adjacent or not), relative stress among them is determined by metrical constituency.

b. The stressability hierarchy is 'violated' in the relation between the two strongest tones (1 and 4) in simple nouns. That is, for example, in disyllabic items like /mian bao/

'bread' or /ji rou/ 'muscle', the stress always falls rightmost in the item⁴.

The more serious problem with this argument, however, is conceptual. Native speakers of BMC have judgements of stress. These judgements must be accounted for. If it could be shown that it is exceptionlessly required, across tone languages, for a tone of a certain shape to give rise to a particular stress judgement relative to another tone of a different shape, then the stress mechanism for Mandarin could perhaps be eliminated, or at least simplified. This (imaginary) principle might state that somehow higher tones are intrinsically more salient than lower tones, for example, and that the stress judgements of Mandarin speakers are simply mechanical reflections of this universal phonetic fact. But until such a principle can be precisely formulated, in a fashion that correctly deals with the shapes and qualities of all tones in all tone languages, and with all relative stress judgements in tone languages, the statement (4) is language particular and must be stated in the grammar of BMC. In fact, the precise formulation of BMC tonally-dependent stress may serve as an initial guide on the road to determining the universal phonetic facts.

⁴ One might be tempted to say that the stressability hierarchy ought to be expressed as: (4 = 1) > 2 > 3 in view of these facts. We will see in the discussion of compounds below that a complete statement of relative strength can be motivated and that it is useful.

3.3.2 Independence of Stress and Tone ?

Another approach to BMC stress would divorce it from tonal quality altogether, and view stress as an arbitrary lexically marked property of words. This is essentially the approach adopted in Hoa 1983, and it leads one to miss significant generalizations. The appendix to Hoa 1983 includes lists of disyllabic lexical items, one list for each of the two possible stress patterns (trochees and iambs). This arrangement, and the discussion in the text, imply that the stress qualities of such items are arbitrary, i.e. that there is nothing in the segmental or tonal content of such item that could affect the stress assignment (although S-W is thought to represent the 'marked' pattern).

There are 15 possible arrangements of the four BMC tones on a disyllabic lexical item (one arrangement, 3 3, is independently excluded). Thirteen of these fifteen arrangements are exemplified, all but one with multiple examples, in the appendix to Hoa 1983. Below is a list of these nine combinations. The 'conforming' specification means the percent of examples of this type whose inclusion in either the list of trochees or iambs can be predicted by (4) above:

(11) Comparison with Hoa 1983 Stress Markings

| Tones | Predicted | Conforming |
|-------|------------|-------------------|
| 1 1 | - * | 100% |
| 2 2 | * - | 100% |
| 4 4 | - * * - | 100% ⁵ |
| 1 2 | * - | 100% |
| 3 2 | - * | 100% |
| 2 1 | - * | 100% |
| 3 1 | - * | 100% |
| 2 4 | - * | 80% |
| 4 1 | - * | 80% |
| 1 3 | * - | 75% |
| 2 3 | * - | 75% |
| 3 4 | - * | 50% |
| 4 3 | * - | 0% |

It can be observed that only two categories have less than 75% conformity with the predictions of the MSR. For all lexical items that show a conflict between Hoa's designation and the predictions of the MSR, whatever their tonal

⁵ See the discussion above concerning concerning deletion depending on the presence of [i] in disyllabic words containing only tone 4. In fact, the native speakers consulted for this study have stress judgements in accordance with the operation of the MSR (4) as given, without the embellishment of deletion based on metrical subordination to [i] described above. The process described in the text for disyllabic words containing only tone 4, then, applies to the idiolects of Hoa 1983 and her informants.

category, repeated checks with two BMC native speakers have yielded judgements in conformity with the MSR's prediction. The reasons for the remaining discrepancies are thus obscure, perhaps depending on morphological or idiolectal factors. In many cases phonological near minimal pairs can be found, for which Hoa gives differing stress assignments, and for both members of which my informants have strong consistent intuitions in favor of what the MSR predicts. An example from the troubling 50% agreeing [3 4] category is the pair: /guli/ 'to encourage', which Hoa gives as S-W, and /nuli/ 'hard, diligent (work)', which Hao gives as W-S. My informants have the clear, consistent intuition that both are W S, agreeing with the prediction of the MSR. It is very hard to see what the distinction might be between these, and how such distinctions could be learned or maintained. This interesting question must be left for future research. In any case, based on my own informant work, I believe the claim that stress in BMC has no relation to tone can justifiably be said to be missing a very significant generalization.

3.4 Compound Noun Stress

In this section we will examine the stress properties of compound nouns, taking the parameter settings given in (4) above (the MSR) as the fundamental statement of BMC stress. Compound nouns are characterized by possession of internal syntactic structure. We assume that while some of

the shorter lexical items in the derivations above (5) have internal morphological structure, they are all 0-level categories in terms of X-bar ranking. A compound noun is a joining of two 0-level categories, resulting in N^0 , but it is not necessarily the case that both components are nouns.

3.4.1 Compound Noun Stress Parameters

Compound nouns are head-final in BMC. The simplest types have a [modifier [head]] structure, with main stress generally falling on the head. The model of stress derivation we will initially assume, for these simple compounds, is shown below:

(1) Compound noun stress parameter settings (CNSR)

- a. MSR applies to head and modifier separately (resulting in independent line 2 heads on each).
- b. Line 2 parameter settings are (+HT,-BND,right).
- c. Build constituents on line 2.
- d. Mark heads of line 2 constituents on line 3.

Examples of the operation of (1) are given below in (2). Wherever a syllable is marked with tone 0, this indicates that the syllable is lexically marked as non-tone-bearing, and it is not eligible for any degree of stress. In addition, the bracketing given indicates the division between modifier and head (the major constituent break), and the MSR applies separately to lines 0-2 on either side of

the break so indicated. Morphological structure that may exist within the modifier or head is not shown.

3.4.2 NP[modifier [head]]_{NP} Derivations

3.4.2.1 Three-syllable Compounds

a.

```

L3:          *
L2:      (*  *)
L1:      (*) (*)
L0:      (*) (*) .
tone:      1   4   3
word:  [jin [jie zhi]] 'gold ring'

```

b.

```

L3:          *
L2:      (*  *)
L1:      (*) (*)
L0:      (*) (*) .
tone:      3   4   0
word:  [lao [gan bu]] 'old cadre'

```

Note in b., above, that when the modifier is monosyllabic, we will assume that the tone 3 line 0 asterisk deletion clause of the MSR cannot apply to it prior to compounding, as this would result in a word (output from the earlier, non-compound cycle) that bears no stress. We will assume that there is a general convention of metrical or prosodic theory that states that a word must have a main stress. This means that /lao/ above enters the compounding process with a stress (tone 3 destressing could not apply since no stronger tones are adjacent to tone 3 in /lao/ in isolation on the cycle prior to compounding). Tone 3 destressing will eventually apply to the compound as a whole, and therefore

is to be seen as a very late process applying after all compounding and compound stress rules have operated. The final representation of b. above is:

```

L3:           *
L2:           (*)
L1:           (*)
L0:           . (*) .
tone:         3   4   0
word:        [lao [gan bu]] 'old cadre'

```

c.

```

L3:           *
L2:         (*) (*)
L1:         (*) (*) (*)
L0:         (*) (*) (*)
tone:         4   1   1
word:        [da [gong ji]] 'big rooster'

```

d.

```

L3:           *
L2:         (*) (*)
L1:         (*) (*) (*)
L0:         (*) . (*)
tone:         4   0   2
word:        [yun dong [yuan]] 'athlete'

```

e.

```

L3:           *
L2:         (*) (*)
L1:         (*) (*) (*)
L0:         (*) (*) (*)
tone:         4   2   1
word:        [shao nian [gong]] 'youth palace'

```

f.

```

L3:           *
L2:         (*) (*)
L1:         (*) (*) (*)
L0:         (*) . (*)
tone:         1   0   3
word:        [chuang hu [zhi]] 'window paper'

```

f. above provides evidence that tone 3 destressing cannot apply on the output of the compound cycle when the syllable bearing tone 3 has acquired the main stress through prior operation of the MSR and CNSR.

In certain compounds, the possibility of stress clash exists.

Consider:

(2)

| | | | |
|-------|------|------|-------------------------|
| L3: | | | * |
| L2: | | (* | *) |
| L1: | (* | *) | (*) |
| L0: | (*) | (*) | (*) |
| tone: | 1 | 4 | 4 |
| word: | [xin | zang | [bing]] 'heart disease' |

The derivation above, in accordance with the CNSR, shows secondary stress falling on 'zang'. However, all informants agree that secondary stress in [[xin zang] bing] falls on 'xin', the first syllable. This is a stress clash effect. We will assume that a stress clash involving adjacent line 2 asterisks is resolved by deletion of the line 1 head which underlies the syllable bearing lesser stress. In (3), this deletion will result in a line 2 asterisk over /xin/ because /zang/ will no longer have a line 1 head to bear the line 2 asterisk for /xin zang/. The modified configuration is shown below:

(4)

```

L3:           *
L2:    (*      *)
L1:    (*) .  (*)
L0:    (*) *  (*)
tone:   1   4   4
word:  [xin zang [bing]] 'heart disease'

```

We will have occasion to treat stress clash in greater detail below.

3.4.3 Four-syllable Compounds and Stress Clash

We turn now to four-syllable compounds of essentially the same syntactic character as the three-syllable cases discussed above: [modifier [head]]. These cases have to be treated separately, however, because, as Hoa 1983 notes, there are often two distinct stress patterns observed. The notation used by Hoa is '2 X X 1' for the pattern of primary stress falling on the last syllable, with secondary stress on the initial syllable, and 'X 2 X 1' for the pattern where secondary stress falls on the second syllable. In many cases, either pattern can be observed. NPs of this type have the syntactic bracketing [[S1 S2] [S3 S4]], where S_n means 'syllable number n'. Examples of this syntactic type include:

(5) [[S1 S2] [S3 S4]] Compounds

- a. [[Bei jing] [fan dian]] 'Beijing Hotel'
[[Beijing] [hotel]]
- b. [[ren zao] [wei xing]] 'satellite'
[[manmade] [satellite]]
- c. [[zuan jing] [ping tai]] 'drilling platform'
[[drill] [platform]]

Let us consider the derivations of these by CNSR:

(6)

| | | | | |
|---------|-------------------------|-----|-----|-----|
| L3: | | | | * |
| L2: | | (* | | *) |
| L1: | (* | *) | (* | *) |
| L0: | (*) | (*) | (*) | (*) |
| tone: | 3 | 1 | 4 | 4 |
| phrase: | [[Bei jing] [fan dian]] | | | |

The derivation above underlies the pattern 'X 2 X 1'. The final stage in (6) would be tone 3 destressing. The existence of the alternative pattern, '2 X X 1', indicates that there is a condition of 'optional clash', that is, a clash which the speaker optionally takes action to avoid. The condition for optional clash is:

(7) Optional Clash Condition

- An optional clash (i.e. a clash that may but need not be repaired) exists when a line 1 head is adjacent to a line 2 head across the major compound constituent boundary.

In the case above, the line 2 head involved is 'jing' and the line 1 head involved is 'fan'. They are adjacent, and the speaker optionally adjusts by leftward movement of the L2 head. The leftward movement is necessitated by

deletion of the line 1 asterisk dominating the weaker tone (in the stressability hierarchy). The result is shown below:

(8)

```

L3:                *
L2:      (*        *)
L1:      (*)  .   (*  *)
L0:      (*)  *   (*) (*)
tone:      3   1   4   4
phrase:    [[Bei jing] [fan dian]]

```

The derivation above underlies the pattern '2 X X 1'. Obviously the optional clash repair has to precede tone 3 destressing, to allow a landing site for the displaced line 2 head. We will stipulate that tone 3 destressing cannot apply to a syllable whose column has been augmented by clash movement of a line 2 head.

We may feel that the type of clash and the choice of moved element are atypical in the case above. This intuition is valid, if we bear in mind that the whole stress system of BMC is atypical in being partially dependent on tone. This tonal dependence comes out in the area of clash avoidance, as well. The full statement of (optional) clash resolution is:

(9) Optional Clash Repair

- When an optional clash exists, the line 1 asterisk dominating the syllable bearing the weaker tone of the two syllables involved in the clash (according to the stressability hierarchy) is deleted.

The adjustments resulting from (9) follow automatically from the parameter settings and general conventions of metrical constituent theory. We note again that Optional Clash Repair (9), if it is to take place, must precede tone 3 destressing (so that a landing site is available), which is always the last rule of compound formation. To assure ourselves that (9) is crucially tone dependent, consider:

(10) [sm]

| | | | | |
|---------|-----------|-------------|--------------------------------|-----|
| L3: | | | * | |
| L2: | | (* | *) | |
| L1: | (* | *) | (* | *) |
| L0: | (*) | (*) | (*) | (*) |
| tone: | 3 | 4 | 1 | 4 |
| phrase: | [[Bei da] | [shu dian]] | 'Beijing University bookstore' | |

(10) represents the optional clash configuration, much as in (6) above. The optional clash existing between 'da' and 'shu' need not be resolved, and if it is not the resulting pattern would be notated as 'X 2 X 1' in Hoa's terms. Now consider how optional clash resolution is carried out in this case, as opposed to (6):

(11)

| | | | | |
|---------|-----------|-------------|----|-----|
| L3: | | | * | |
| L2: | | (* | *) | |
| L1: | | (*) | . | (*) |
| L0: | (* | *) | * | (*) |
| tone: | 3 | 4 | 1 | 4 |
| phrase: | [[Bei da] | [shu dian]] | | |

We observe that since /shu/ bears the weaker tone according to the stressability hierarchy ($1 < 4$), its line 1 asterisk

must be deleted. But in this case, unlike (8), the deletion does not lead to head movement on a higher line. Hence /Bei/ is not augmented by movement, as it was in (6), and this difference is quite striking to the ear, as metrically prominent syllables tend to have more fully realized tone. The result is that Hoa's 2 X X 1 pattern effectively does not exist for this compound, only X 2 X 1, with '2' more or less prominent relative to the second 'X', depending on whether or not Optional Clash Repair has applied.

Let us consider another compound:

(12)

| | | | | |
|---------|------------|-----|-------------|-----|
| L3: | | | | * |
| L2: | | (* | | *) |
| L1: | (* | *) | (* | *) |
| L0: | (*) | (*) | (*) | (*) |
| tone: | 2 | 4 | 4 | 1 |
| phrase: | [[ren zao] | | [wei xing]] | |

The derivation above underlies Hoa's 'X 2 X 1' pattern. We observe that an 'optional clash' exists involving 'zao' and 'wei', across the syntactic boundary. Here, since the tones on the two syllables involved are identical, either line 1 asterisk can potentially be deleted in optional clash resolution. If the line 1 asterisk dominating /wei/ is chosen for deletion, the result is just a slight variant on the metrical pattern in (12), since no head movement on a higher line can result. If, alternatively, the line 1 asterisk dominating /zao/ is deleted, the resulting head movement will give:

(13)

```

L3:                *
L2:          (*      *)
L1:          (*)  .  (*)  (*)
L0:          (*)  *  (*)  (*)
tone:         2   4   4   1
phrase:  [[ren zao] [wei xing]]

```

The derivation above underlies Hoa's 2 X X 1 pattern for this item.

(14)

```

L3:                *
L2:          (*      *)
L1:          (*  *)  ( ?
L0:          (*)  (*)  (*  *)
tone:         4   3   2   2
phrase:  [[zuan jing] [ping tai]]

```

In the derivation above, there is a clash between two line 2 heads across the internal compound boundary. This clash effect is more serious than the optional clashes so far considered. The standard repair mechanism can apply (deletion of the line 1 asterisk of the weaker tone, in this case tone 3 on /jing/), resulting in (prior to tone 3 destressing on the compound):

(15)

```

L3:                *
L2:          (*      *)
L1:          (*)  .  (*)
L0:          (*)  (*)  (*  *)
tone:         4   3   2   2
phrase:  [[zuan jing] [ping tai]]

```

In Hoa's terms, this would be notated as: '2 X 1 X'. We note that Hoa does not include this item under her list of '2 X 1 X' patterns, and places it with the alternating types such as 'Beijing fandian', above. Here the results of our native speaker consultations differ from Hoa's, and the issue must be left unresolved. Note however that to the extent an alternative pronunciation has been observed by Hoa, it is probably partially based on an independent mechanism, operating on /zuan jing/. /zuan jing/ in isolation is a transitive verb phrase consisting of 'to drill' and the direct object noun 'well'. In Chinese there is a syntactically based prosodic rule that has effects resembling the familiar Nuclear Stress Rule for English. However, this rule makes crucial reference to syntactic configuration, and it has the power to override the tonally based stress mechanism that operates within X^0 . This source of variation must be independently recognized by any analysis. Examples from Hoa 1983 (but note that she does not explicitly recognize this as a syntactically based process) include:

(15)

| | | |
|-----------|---|---------------|
| | * | |
| 1 | 3 | |
| 'he shui' | | 'drink water' |

| | | |
|-----------|---|------------|
| | * | |
| 4 | 2 | |
| 'pao cha' | | 'brew tea' |

| | | |
|------------|---|---------------------|
| | * | |
| 4 | 2 | |
| 'guo qiao' | | 'pass (the) bridge' |

The examples in (15) are all 'violations' of the tone-based stressability hierarchy, because a 'weaker' tone gets the main stress. These results are very clearly based on the near universal prosodic rule that in the configuration (S) V O (or (S) O V), the object is main-stressed. Now the effect Hoa is observing may be a resistance to clash repair when it would violate the principle that stresses a direct object in preference to a verb (even in a compound). It may be that Hoa's informants have this stronger version of the constraint than our informants⁶.

There is another possible compound internal bracketing, which has effects on the stress pattern. This is [S1 S2 S3 [S4]]. Examples include:

(18)

[[dong xiao mai] di] 'winter millet field'
 [[suan niu nai] ping] 'yogurt bottle'
 [[tuo la ji] zhan] 'tractor station'

⁶ Another possibility is that since apparently for Hoa in /ping tai/, the main stress is on /tai/, no clash existed to begin with. Our informants do not agree on this point, and the issue must be left unresolved.

The derivations are as follows:

(19)

```

L3:                                *
L2:                                (* *)
L1:          (*          *)      (*)
L0:          (*)  (*)  (*)  (*)
tone:         1    3    4    4
phrase: [[dong xiao mai] di] 'winter millet field'

```

In (19), we see an example of what we will now call 'Obligatory Clash' (we saw a preliminary example of Obligatory Clash in (14) above). This means a clash that must be repaired, as opposed to the 'Optional Clash' discussed above. The condition for Obligatory Clash is stated below:

(20) Obligatory Clash Condition

- An obligatory clash (i.e. a clash that must be repaired) exists when a line 2 head is adjacent to another line 2 head.

Note that in (20), unlike (7) (the Optional Clash Condition), we need not state that the clash is only defined across a compound constituent boundary, as distinct line 2 heads can only be adjacent across a compound boundary. In obligatory clash, just as in optional clash, the repair mechanism deletes the tonally weaker line 1 asterisk, which will result in head replacement:

(22)

```

L3:             *
L2:      (*      *)
L1:      (*      *      *)
L0:      (*      *) *      (*)
tone:      1      3      4      4
phrase: [[dong xiao mai] di] 'winter millet field'

```

We note that obligatory clash exists in (2) /xin zang bing/ above, and it is mandatorily repaired, in the same way.

A example where no clash exists is:

(23)

```

L3:             *
L2:      (*      *)
L1:      (*      *      *)
L0:      (*      *      *) (*)
tone:      1      2      3      2
phrase: [[suan niu nai] ping]

```

Here there is no clash of any sort. The pattern can only be, using Hoa's notation, '2 X X 1'. This example shows that the obligatoriness of the '2 X X 1' surface pattern in words of the morphological pattern discussed here can have two distinct sources: either there was a clash, which was mandatorily repaired, or there was not a clash, and the single underlying pattern is '2 X X 1', which simply appears as such on the surface.

Finally, consider:

(24)

```
L3:                *
L2:                (*  *)
L1:                (*  *  *)  (*)
L0:                (*) (*) (*)  (*)
tone:              1  1  1  4
phrase:  [[tuo la ji] zhan]
```

Here there is an obligatory clash between the heads on /ji/ and /zhan/. If repair is as in the other cases above, we expect that the line 2 head will be replaced to /la/. In fact, the head is unexpectedly replaced on /tuo/, furthest to the left. However, recall the rule of MD, motivated above. If we order MD before stress clash deletion, the head replacement must be to /tuo/ rather than /la/, as /la/ has lost the ability to bear a head⁷:

Medial Deletion:

```
L3:                *
L2:                (*  *)
L1:                (*  .  *)  (*)
L0:                (*) * (*)  (*)
tone:              1  1  1  4
phrase:  [[tuo la ji] zhan]
```

Obligatory Clash Repair (deletion and head replacement):

```
L3:                *
L2:                (*  *)
L1:                (*  .  .  *)  (*)
L0:                (*) * *  (*)
tone:              1  1  1  4
phrase:  [[tuo la ji] zhan]
```

⁷ We are assuming that Medial Deletion is cyclic, otherwise in (24) it would delete both line 1 medial asterisks.

3.5 An Alternative Treatment (Yip 1989)

Yip 1989 has treated the four syllable compounds discussed above. We will present her approach and then consider it in relation to the treatment suggested above⁸.

The analysis of Yip 1989 uses two basic elements, a rule and a domain. The domain is the phonological phrase. Yip holds that in Mandarin a phonological phrase can be up to four syllables long. In a typical 4-syllable compound, it is generally possible either to treat the entire compound as a single P-phrase, or to break the compound into two distinct P-phrases, at the major constituent boundary. So, for example:

(25) P-phrasing (Yip 1989)

- a. [p-phr. Beijingfandian]p-phr.
- b. [p-phr. Beijing]p-phr. [p-phr. fandian]p-phr.

A stress grid is build over a P-phr. by an "edge-anchored grid rule that lays down grid marks on the two edges of the domain". This results in the following grids for (25)a. and b.:

⁸ It should be noted that the analysis of Yip 1989 was part of a short oral presentation, and could hardly be expected to be comprehensive. This fact does not, of course, diminish the force of these objections.

(26) P-phrase grids

*
*
* * * *

a. [p-phr. Beijingfandian]p-phr.

* * * *
* * * *

b. [p-phr. Beijing]p-phr. [p-phr. fandian]p-phr.

The effect of the grid rule is basically the same in cases a. and b. above: the left edge is weak relative to the right edge. The different patterns are accounted for by the distinction in domain size.

Another rule labels the right sister of two neighboring P-phrases strong, so (26)b. will emerge as:

(27)

* * * *
* * * *

b. [p-phr. Beijing]p-phr. [p-phr. fandian]p-phr.

The right strong rule does not apply to (26)a., as there are not two distinct P-phr.s to be related.

Yip 1989 treats the [S1 S2 S3] S4] bracketing (which, recall has only 2 X X 1 stress patterning, see (18) above) by showing that if the compound is analyzed as two P-phrases, a clash will occur that forces leftward movement of

the asterisk on S3 to the nearest column that can support it on its own level, thus:

(28)

a. initial input:

[dong xiao mai]_{p-phr} [di]_{p-phr} [winter millet] [field]

b. edge-anchored grid rule on P-phrases:

```

          *
        *
       *
      *
     *
    *
   *
  *
 *
[ dong xiao mai ]p-phr [di ]p-phr

```

c. right sister P-phr. strong rule:

```

          *
        *
       *
      *
     *
    *
   *
  *
 *
[ dong xiao mai ]p-phr [di ]p-phr

```

Yip then suggests that a stress clash exists between 'mai' and 'di', so that the head of 'di's' column moves leftward until it finds a column that can directly support it, namely 'dong':

d.

```

          *
        *
       *
      *
     *
    *
   *
  *
 *
[ dong xiao mai ]p-phr [di ]p-phr

```

While this is in many ways an attractive formulation, we do not adopt it in its entirety, for two reasons. The first reason is theory internal. There is no way to characterize the 'edge-anchored grid rule' in our framework, with precisely the desired properties. This is in part a principled failure, resulting from the restrictiveness of the Halle and Vergnaud 1987 theory. In addition, such a process, could it be stated, would take no account of tonal quality, which, as we have attempted to demonstrate above, is an important factor in the BMC stress system.

The second reason is that in Yip's approach the P-phrase has no independent status. It is used to explain the different stress patterns, but the only evidence presented for its existence is the existence of different stress patterns.

Perhaps it is not unnatural, however, to make a phonological division of a compound at the major syntactic boundary. In the analysis presented here, the phonological phrasing follows from the existence of clash, optional or obligatory. That is, in addition to Optional Clash Repair (9) and Obligatory Clash Repair (21), which manipulate the heads and grid constituent structure of the affected compounds, there is always the strategy of separating the two clashing components in time, resulting in a slight pause. If Optional Clash exists, this separation (effectively, break-up into two phonological phrases,

characterized by slight pause insertion between phrases) is optional. Thus, in case of optional clash, the speaker may choose:

(29)

- a. to ignore the clash
- b. to resolve the clash within a phonological phrase by asterisk movement
- c. to resolve the clash by time separation of the offending elements, resulting in distinct phonological phrases

In case of Obligatory Clash, the options are only (29)b. and (29)c. above, that is, there is no option to ignore the clash.

We believe that this characterization of the existence and properties of the phonological phrase avoids the potential circularity encountered in Yip 1989's formulation.

For the [S1 S2 S3] S4] bracketing, which results in Obligatory Clash, this means that asterisk movement gives Hoa's [2 X X 1] pattern, but there ought to be another pattern, where the optional clash is resolved by phonological phrasing (pause insertion at constituent boundary). And indeed, contra Hoa 1983, this option does exist.

3.6 Conclusion

In summary, we have provided evidence that stress in Beijing Mandarin Chinese is partially dependent on syllable tone, and is derived by means of metrical parameter settings. BMC is an interesting case in metrical theory, because while languages that reflect syllable weight in the metrical structure are common, languages that reflect tone quality in metrical structure are quite uncommon, though the reverse is sometimes found (cf. chapter 2 and Sietsema 1989).

Chapter 4

On the Representation of Intonation in Linguistic Theory

4.1 Introduction

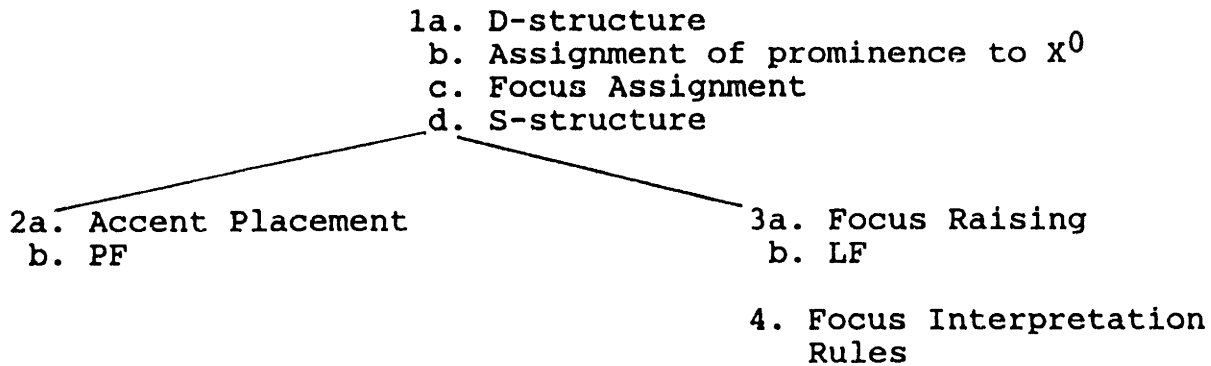
There has been a great deal of research on the form and function of English intonation, resulting in a variety of models with somewhat different formal characteristics (Bruce 1982, Pierrehumbert 1980, Ladd 1983, Bolinger 1986, Garding 1979, Bing 1979, etc.). Here we will propose a maximally simple model of English intonation, and consider its adequacy on empirical and conceptual grounds, comparing the new formulation with its rivals where such comparison is found to be illuminating.

4.1.1 The Focus Model

Any model of intonation has to assume some kind of prominence or focus marking in an utterance that is input to the intonational component. For present purposes (presentation of a new intonation model), the exact character of the focus input to intonation is not crucial. All that is required is a theory that distinguishes certain essential properties of focus. No doubt there are many subtle and interesting issues in the theory of focus that remain to be illuminated, but these distinctions are not crucial for this discussion of the intonation model. The

general linguistic model we will assume initially is based on that motivated by Rochemont 1986 (henceforth MSR). A gross outline of the model is shown below (MSR p. 36):

(1) Model of Grammar



4.1.1.1 Focus Structure, Accent Placement, Interpretation

Initially, we will present Rochemont's conception of focus derivation, which is based on free assignment of prominence, focus labelling based on prominence specification, and subsequent percolation of focus labelling, up the syntactic tree as s-structure. We will then make a minimal modification to his scheme, replacing the free assignment of prominence with free assignment of focus, and doing away with the percolation conventions. But because so much of the mechanics and the spirit of Rochemont's model is retained, it is important that it be understood on its own terms.

Component 1. in (1) is the familiar root of the 'T' model of the REST (Chomsky 1980). In Rochemont's

conceptualization, an abstract feature specification [+prominent] is freely assigned after the application of Move-alpha in the derivation from D-structure to S-structure. The assignment of another abstract feature specification [+focus], is dependent on the location of [+prominent] in a sentence, in accordance with the following rules of percolation¹:

(2) Focus Assignment ((1)1.c)

- a. If α is prominent, then α is [+focus].
- b. i. If α is [+focus] and α is X^0 , then X^n is [+focus].
- ii. If α is [+focus] and α is an argument of X^0 contained in X^n , then X^0 is [+focus].
- iii. If X^0 is [+focus] and α is an adjunct of X^0 , then α is [+focus].

Free prominence assignment and focus percolation, as specified above, act together to derive a focus-annotated S-structure (Focus Structure in Selkirk 1984). Moving on to PF, Rochemont provides a reformulation of the familiar Nuclear Stress Rule to formalize Accent Placement ((1)2a.):

(3) Nuclear Stress Rule:

- Assign an accent to the rightmost lexical category (in a [+focus] constituent in S).

¹ Although these rules will be set aside in favor of maximally free assignment of [+focus] immediately below, they are introduced here as part of our larger presentation of Rochemont's model, most of which is crucially assumed for our theory of the focus/intonation relation.

The exact phonological character of the 'accent' is left unspecified, and we will have more to say on this in the discussion of the intonation model below.

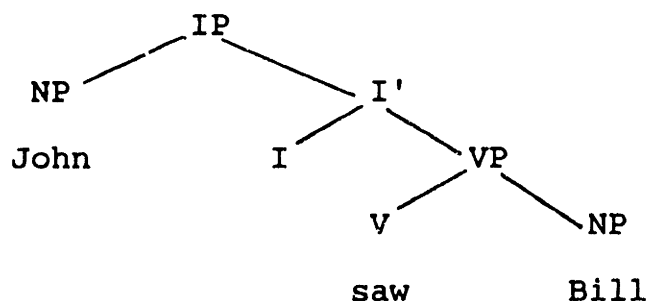
At LF, a rule of Focus Raising applies to uniquely distinguish the [+focus] constituents of a sentence, and to capture the analogy, observed by Chomsky (1976), between focused NPs and certain anaphoric expressions. Focus Raising is, then, a typical CP adjunction operation, leaving a trace. The exact characterization and justification of Focus Raising at LF will not concern us here. The intonation model to be presented below will not be affected by the presence or absence of this operation in the larger theory of focus. It is mentioned here for completeness.

In MSR, a battery of discourse relevance conditions are held to interpret the LF representation of an utterance, including the focus marking. These are the Focus Interpretation Rules ((1)4. above). Although it appears that Rochemont has taken a very long step towards capturing the correct formulation of discourse relevance conditions, the exact formulation of these rules will not concern us here. It is sufficient for our purposes to assume the existence of discourse relevance conditions that act as a check on the legitimacy of [+focus] specifications, and to assume that these can in principle be fully and accurately specified.

4.1.1.2 A Focus Structure Derivation

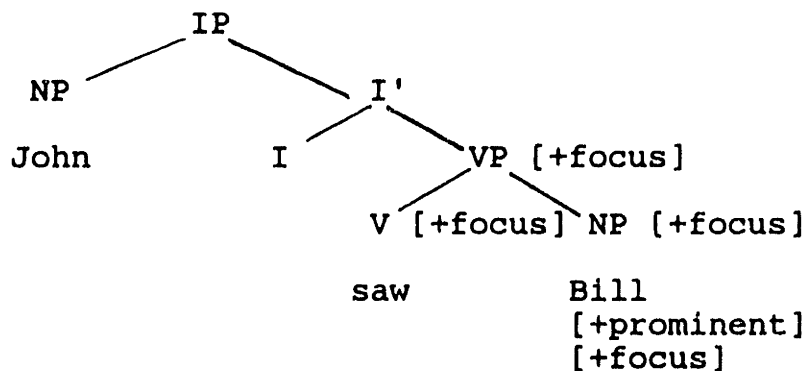
An example will serve to clarify the operation of the key components of the MSR model. Consider the sentence:

(3)



In the context of a question like: "What did John do?", an assignment of [+prominent] to 'Bill' will yield [+focus] on 'Bill' ((2)a.), and then [+focus] on 'saw' ((2)b.ii), and then [+focus] on the whole VP ((2)b.i) (in fact the percolation of [+focus] must continue up to IP, a fact we will return to shortly). The result of the basic percolation up to VP as just described is:

(4)

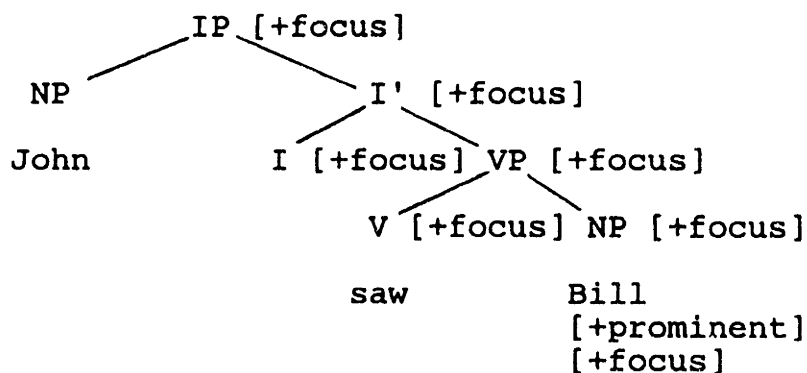


Since 'John' was present in the context question, we do not wish [+focus] to percolate to 'John', and this is correctly blocked by (2).

An assignment of [+prominent] to 'John' in the given context (recall that prominence assignment is free) will result in a focus structure that is ruled out by the discourse relevance conditions. It is ruled out because an item immediately recoverable from the discourse cannot be marked [+focus], and 'John' is immediately recoverable from the discourse question given: "What did John do?" (cf. MSR for details).

In discourse initiating contexts or situations, such as the question "what happened?" or the situation of an individual walking into a room and initiating a new discourse we will wish to specify the entire utterance as [+focus], since 'John' will not be immediately recoverable from the context, at least not any more readily recoverable than 'Bill'. In this case the result of assigning [+prominent] to 'Bill' will be the same as before, but now we show the full percolation to IP:

(5)



A problem thus arises in distinguishing (4) and (5) (recall that in (4) the percolation will continue from VP up to the full IP, identically with (5)). A full consideration of the technical details of this problem (not explicitly discussed by Rochemont), which will implicate the exact formulation of Focus Raising and Focus Reconstruction (another LF focus operation provided by MSR's model) at LF, would take us too far afield here. Our concern is to provide a minimally plausible focus model as an annotation system for the syntactic input to intonation. We will therefore adopt, in place of [+prominent] assignment and Focus Assignment (2) the simpler proposal of Horvath 1981, that [+focus] assignment is free, and, now back to the MSR model, that filtering of this free assignment of [+focus] is accomplished by the discourse relevance conditions directly. It is clear that they have the power to perform this filtering, and that therefore the explicit percolation mechanism of (2) is unnecessary. Furthermore, we note that in general the single (free) assignment of [+prominent] that

guarantees the correct focus structure as input to PF is usually duplicated by the assignment of accent at PF by the Nuclear Stress Rule (3). So, for example, in the case of 'John hit Bill' in the context 'What did John do?', the assignment of [+prom] that guarantees the correct percolation of [+focus] (shown in (4) is on 'Bill' and given the [[_{VP} hit Bill]_{+focus}] structure, the PF Nuclear Stress Rule will, independently place an accent on 'Bill'. In general we would wish to avoid this redundancy if possible, and with free assignment of [+focus] this will indeed be possible.²

To summarize, the discourse relevance conditions of the model we are adopting will filter all assignments of [+focus] except the ones indicated below, in the given contexts:

(6)

| Context | Legal [+focus] Assignment |
|-------------------|--|
| What happened? | [_{IP} John saw Bill] _{IP} [+focus] |
| What did John do? | John [_{VP} saw Bill] _{VP} [+focus] |

² Free assignment of [+focus] raises certain questions, e.g. whether resulting embedded [+focus] marking is to have any status in the theory. For simplicity, we will rule out embedded [+focus] with a filter:

*[... [...]_{+focus} ...]_{+focus}.

4.1.1.3 The Formulation of Accent Placement

The Nuclear Stress Rule as given in (3) is unacceptably vague as a PF operation. This is understandable, given Rochemont's primary concern with the syntax and semantics of focus, but since intonation is a PF process, we would wish a more precise specification for the prominence structure that is to be input to it. We will therefore adopt, in place of the MSR Nuclear Stress Rule, the formalization of Nuclear Stress as an operation constructing bracketed metrical grids as given by Halle and Vergnaud 1987. Again, we emphasize that while perhaps the details of Halle and Vergnaud's account could be revised or extended in interesting ways, the model they propose is sufficient for our present purpose, which is to adopt a precise model of abstract metrical prominence at PF, upon which the intonation model can be based.

The example treated by Halle and Vergnaud (henceforth HV) is the sentence: "Jesus preached to the people of Judea", for which the Nuclear Stress falls on "Judea", the rightmost element. In terms of the focus model we are assuming here, a phonological prominence on "Judea" is compatible with any of the following context/focus-structure pairings:

(7)

Context / Focus structureWhat happened? / [J. preached to the people of Judea]_{+F}What did J. do? / J. [preached to the people of Judea]_{+F}Who did J. preach to?/J. preached to [the people of Judea]_{+F}Which people did J. preach to? / J. preached to the people
[of Judea]_{+F}**(8) Nuclear Stress Rule (HV)**

- a. Parameter settings on line N (N>=3) are (-BND,+HT,right).
- b. Interpret boundaries of syntactic constituents composed of two or more stressed words as metrical boundaries.
- c. Locate the heads of line N constituents on line N+1.

HV point out that:

The Nuclear Stress Rule [8] systematically skips over constituents where one of the elements is a stressless word (clitic), since such clitics have no effect on the stress contour of the phrase. Thus, of the four nested constituents in the prepositional phrase in [9],

[9] [to [the [people [of Judea]]]]

the boundaries of only one, shown in [10], are interpreted as metrical boundaries.

[10] to the [people of Judea]

We now quote the full HV discussion of the workings of the NSR. We assume, in terms of the model presented here, that their example sentence 'Jesus preached to the people of Judea' is focus annotated as [IP ...]IP[+focus], in a discourse-initiating context.

The procedure sketched above will fail in cases where the phrases ... form part of a larger constituent, as, for example, in [Jesus [preached [to the people [of Judea]]]], for as shown in [11] it will not reflect the fact that 'Jesus' has more stress than 'preached' and 'preached' has more stress than 'people'.

[11]

```

      .           .           .           *
-----
    (.           .           .           * )
      -----
      .   ( .           .           * )
            -----
      *           *           (*           *)
[Jesus [preached [to [the [people [of [Judea]]]]]]

```

In order to reflect the stress contours of such phrases correctly, we postulate that stress rules are subject to the Stress Equalization Convention [12].

[12] Stress Equalization Convention

When two or more constituents are conjoined into a single higher-level constituent, the asterisk columns of the heads of the constituents are equalized by adding asterisks to the lesser column(s).

We illustrate the effects of [12] in [13], where the asterisks added by [12] are enclosed in braces { }.

[13]

```

line 6  .           .           .           *
-----
line 5  {*}         .           .           *
      -----
line 4  {*}         {*}         .           *
            -----
line 3  *           *           *           *
      (Jesus (preached to the (people of Judea)))

```

Given the procedure just outlined, the number of lines in the metrical grid is equal to $3+n$, where n is the number of metrically interpreted constituents of the most deeply embedded word. Since there is no upper bound on the depth of embedding, there is also no upper bound on the number of stress distinctions. Phonetically, however, speakers make fewer distinctions than are provided by this procedure. To reflect this fact, we propose the Grid Simplification Convention [14].

[14] Grid Simplification Convention

Lines in the metrical grid may be eliminated beginning with the line immediately above the word layer and proceeding upward without skipping. This process may stop at any line but must leave at least the top line in the grid intact.

In the example sentence used above, the entire sentence was assumed to be marked [+focus], in a discourse-initiating context. It may, however, be the case that only a portion of a sentence is marked [+focus] (by virtue of the discourse relevance conditions), the remainder taking [-focus] as a default, as specified by MSR. An example is:

(15)

What did Joe do in the library ?

[Joe]_[-F] [studied his books]_[+F] [in the library]_[-F]

The Nuclear Stress Rule will operate on any domain, whether is it labeled [+focus] or [-focus]. These are separate applications. The boundaries of [+focus] constituents are intonation phrase boundaries. The intonational phrases so formed are not necessarily indicated

by pause, rather they are signaled by pitch range expansion (though pause is of course possible). Likewise, the onset of a [-focus] phrase has the intonational reflex of pitch range compression. Note that this formulation is not stating that focus constituents are the only basis for assigning intonation phrasing. However, the intonational phrasing whose phonetic reflex is pitch range compression and expansion is always based on focus constituency. We will see the phonetic effects of the phrasing in the discussion of the intonational model below. For the moment, we observe that the final metrical structure over (15) will have constituents resulting from three separate applications of the NSR, and therefore three separate intonational phrases:

(16)

| | | | |
|------|---------------------------------------|--|---|
| 1. 4 | | * | |
| 1. 3 | (*) | (*) | (*) |
| | (John) _{IP1} _[-F] | (studied his (books)) _{IP2} _[+F] | (in the library) _{IP3} _[-F] |

Note crucially that the metrical structure in (16) is independent across IP boundaries. That is, there is nothing in the metrical structure that says anything about the metrical relation of 'library' to 'books', for example. However, the intonational phrase boundaries and the value of the [focus] feature are visible to the intonational component, which adjust pitch range accordingly. We will see

how this works for a variety of examples in the presentation of the intonation model below.

In summary, we are essentially adopting the general focus model of Rochemont 1984, combined with the PF metrical prominence assignment machinery of Halle and Vergnaud 1987. A representation like that in (16) will be input to the intonation component, which will respect the relative metrical prominences, in a fashion to be made precise below.

4.2 The Intonation Model

4.2.1 Laryngeal Generation

In discussing the intonation model, we will, for expository convenience, adopt the symbol IG (intonation grid) to stand for a Cartesian coordinate system with frequency (or perhaps the psychologically determined concept 'pitch') on the y-axis and time on the x-axis. The symbol IP will stand for an intonation domain. For the present, we can think of an intonational domain (IP) as coextensive with a metrical domain based on focus marking. The details are not important for this preliminary discussion.

Intonation is a function of fundamental frequency generated at the larynx (with some non-laryngeal perturbations which will be abstracted away from at present) across time. The larynx has intrinsic limitations governing the form of fundamental frequency contours it can generate.

An example of these physiological limitations is the pitch range: presumably every speaker has a maximum and minimum frequency that can be generated. Another example is regularity: as with other organic functions, a general pattern can be achieved, but precise repetition of a contour or target is impossible. There are a variety of limitations of this type. Bearing in mind these intrinsic limitations on the possibilities of frequency generation, the simplest specification of intonational contour over ID is the following:

(17) Basic Intonation Rule (BIR)

Generate A

(A has the form of the graph of a function which is continuous in those portions of ID independently (segmentally) specified for voicing)

Since (17) is a 'command' to the larynx, the resulting graph will be filtered by the intrinsic, physiological properties of the laryngeal generation. Any contour not so filtered, and not filtered by other mechanisms to be described, is licensed as the intonational contour over some English ID. The claim here, then, is that (17) is the fundamental component of English intonation.

The Basic Intonation Rule is a statement of the primary function performed by FG, the fundamental frequency generation 'black box' introduced in chapter 1. As always, one may think of statements like the BIR as either rules

governing the construction of representations, or as filters on the output of a component (in this case FG). The BIR happens to be maximally simple as a statement of function, but the resulting forms are limited by FG's required metrical constituent structure input, to be discussed immediately below.

4.2.2 Relation to Metrical Structure

We now turn to the theoretical grounding of the observation that the Basic Intonation Rule (BIR) respects metrical structure. Consider the metrical representation of a simple discourse-initiating utterance:

(18)

```

line 5           *
line 4   (*           *)
line 3   (*)  (*)   (*)
          [Joe studied his books][+focus]

```

It is generally the case that an item with greater metrical prominence, such as 'books' in (18), will exhibit some kind of greater pitch movement, whether raising or lowering. To make this intuition specific, we will first give a procedure for drawing a reference line across IG (intonational grid) that contains the intonational contour resulting from an application of the BIR to an intonational phrase (IP):

(19) Reference Line Rule (RLR)

A line is drawn from point /A/ to point /B/ across IG in IP.

/A/ is chosen as follows:

- i. the F0 offset of the first stressed word in IP,
if the first stressed word in IP does not carry
the main stress of IP,

otherwise,
- ii. the F0 onset point of the first stressed word in
IP, if the first stressed word in IP carries the
main stress of IP.

/B/ is chosen as follows:

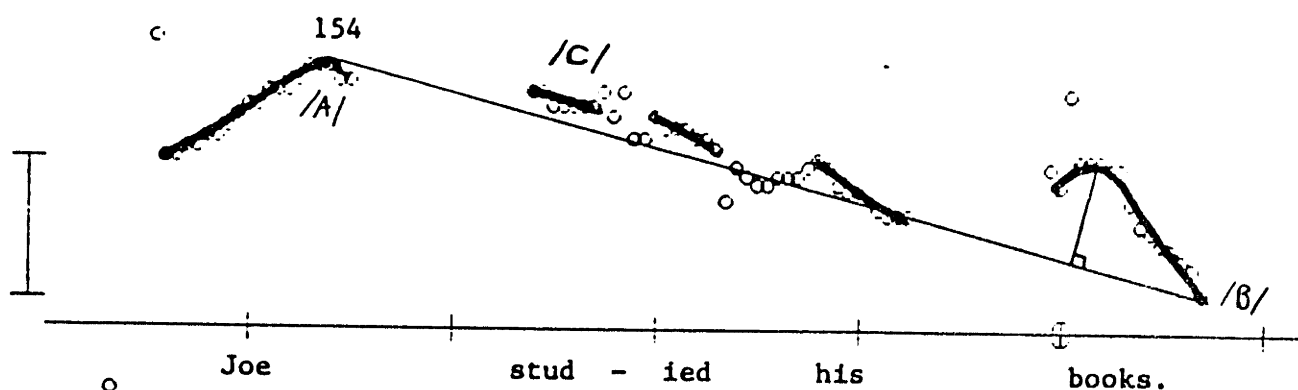
- iii. the F0 onset point of the last stressed word in
IP, if the last stressed word in IP does not carry
the main stress of IP,

otherwise,
- iv. the F0 offset point of the last stressed word in
IP, if the last stressed word in IP carries the
main stress of IP.

A 'stressed word' is a word that has a line 3 representation
in IP.

Let us consider an example of the use of the RLR to
draw a reference line across the IG resulting from an
application of the BIR in the IP (18):

(20)



(20) is an actual measured intonational contour across the intonation domain given in (18)³. Clearly, since it is a contour produced by a human being, it falls within the range of contours permitted by the BIR. As for drawing the reference line, we must choose point /A/ as specified by (19)i., since 'Joe' does not carry the main stress of its IP (as shown in (18)). Point /B/ is specified by (19)iv. Then let us draw a line between points /A/ and /B/. This line will be referred to as the reference line in IG over IP.

The Excursion Filter then applies to the representation (20), to guarantee that the results of application of the BIR respect relative metrical prominence⁴:

³ All F0 tracings in this section are from O'Shaughnessy 1976, unless otherwise specified.

⁴ The excursion filter as formulated here is intended to derive the properties of pitch that are called 'pitch accents' in many other theories (cf. Selkirk 1984, for example). Pitch excursions may also be associated with boundary situations in the surface phrase marker ('boundary tones' in other theories, cf. Pierrehumbert 1980, for example). To the extent that the 'boundary tones' co-occur with independently-motivated categories in linguistic theory, the Excursion Filter given here should be extended by an additional clause that specifies a minimal distance that the contour must maintain from the reference line under pressure from those categories.

(23) Excursion Filter

*IG over IP, if an item represented (maximally) on line i in the metrical grid of IP has a greater distance in F0 from the reference line in IG than an item represented (maximally) on line j in the grid, where $j > i$ ⁵.

We see from (20) that the reference line begins at the end of 'Joe'. Therefore the Excursion Filter does not apply to 'Joe', because no reference line exists across this item, and the Excursion Filter is defined over the reference line.

We then observe that (20) is in conformity with the Excursion Filter, in that 'books' has a greater F0 distance from the line than any other (metrically less prominent) item, as specified by the grid (18). If, for example, there had been a pitch excursion up to point (C) in (20), over 'studied', the Excursion Filter would have ruled out the contour as incompatible with the given focus and metrical structure. An excursion over 'studied' is not, of course, ruled out absolutely, as it can be generated by the BIR. It is simply that this excursion would reflect a different focus and metrical structure, with emphasis or focus on 'studied'. The Excursion Filter is therefore a statement of a function internal to the fundamental frequency generator (FG) 'black box' introduced in chapter 1: the method of

⁵ The question naturally arises here as to what the units of distance measurement are to be. We regard the exact specification of this as an open question, though in practice it will matter very little for presentation of the system, as the examples just below will show.

reconciling the freedom of the BIR with the required metrical constituent input representation.

Another interesting case involves modals. Certain modals and auxiliaries carry the main stress within a [+focus] constituent just as in cases of polarity focus. We will treat an example involving 'might'. Consider the utterances:

(24) What will Joe do?

Joe [might study his books][+focus]

The result of applying the NSR to (24) without taking the special status of modals into account would be:

(25)

| | | | |
|--------|-----------|-----------------------------|----|
| line 6 | | | * |
| line 5 | (* | | *) |
| line 4 | * | (* | *) |
| line 3 | [Joe][-F] | [might study his books][+F] | |

We see that 'books' is more prominent than 'might'. In fact, 'might' will usually be the most prominent element in the sentence. We will therefore posit a rule that extraposes the rightmost asterisk on the line 5 constituent built by the NSR (or, in general, the line just below that in which the asterisk dominating the modal would be lost by normal operation of the NSR), just in case there is a modal on the

left periphery of the focus constituent, as in (25). This gives:

(26)

line 6 *

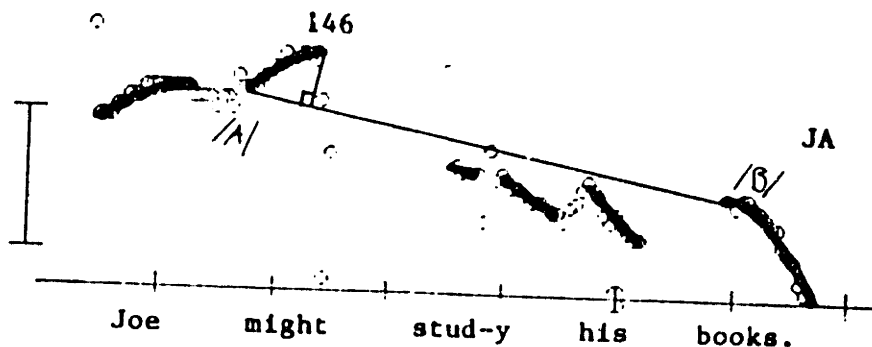
line 5 (* *)

line 4 * (* *)

line 3 [Joe]_[-F] [might study his books]_[+F]

The IG corresponding to (26) is:

(27)

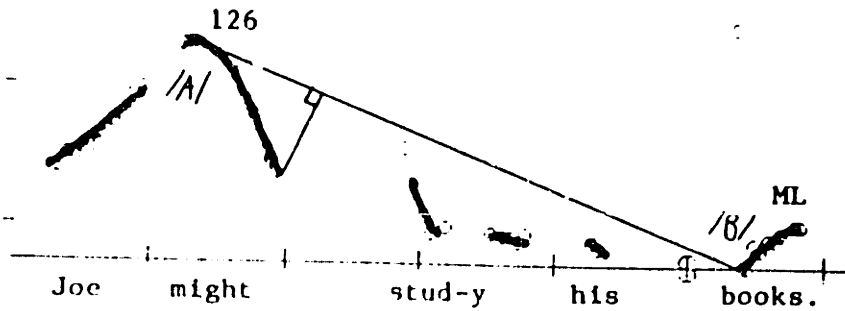


In (27) we see the reference line drawn from the F0 onset of 'might' to the F0 onset of 'books', as specified by the RLR, based on the relative stresses of 'Joe' and 'might'. As before, where no reference line is defined (in this case over 'books') the excursion filter does not apply. Clearly, 'might' has greater distance from the line than any other item over which the reference line is defined, and the contour is therefore acceptable, under the focus conditions specified.

A central claim of the BIR is that directionality in contour formation (high vs. low, rising vs. falling) is not

significant at the level of well-formedness checking. Let us consider another renditions of the IP corresponding to (26), with the pitch directions on the two most prominent items reversed:

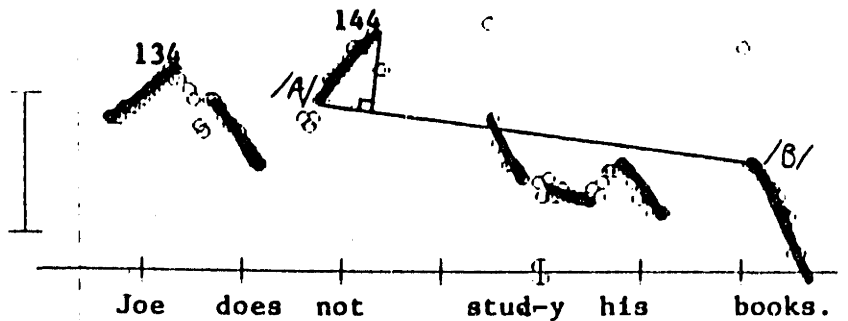
(28)



Here the prominence on 'might' is signalled by the a falling tone, while 'books' has a small rise. The reference line is drawn exactly as in (27), and the utterance is well-formed in terms of the excursion filter, just as in (27).

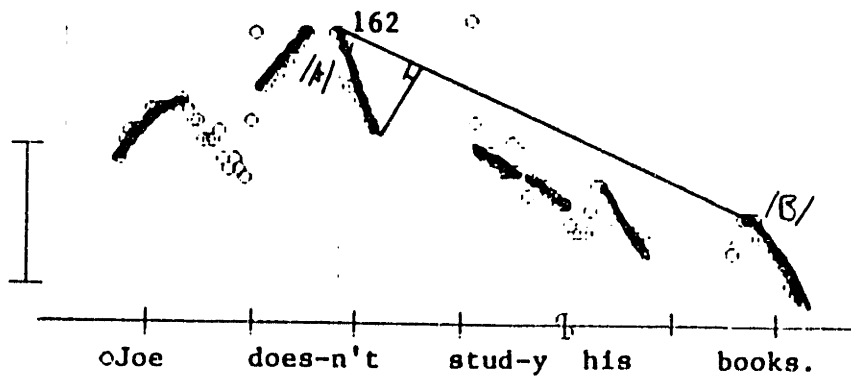
Negative polarity focus works in a fashion similar to what has been shown for modals:

(29)



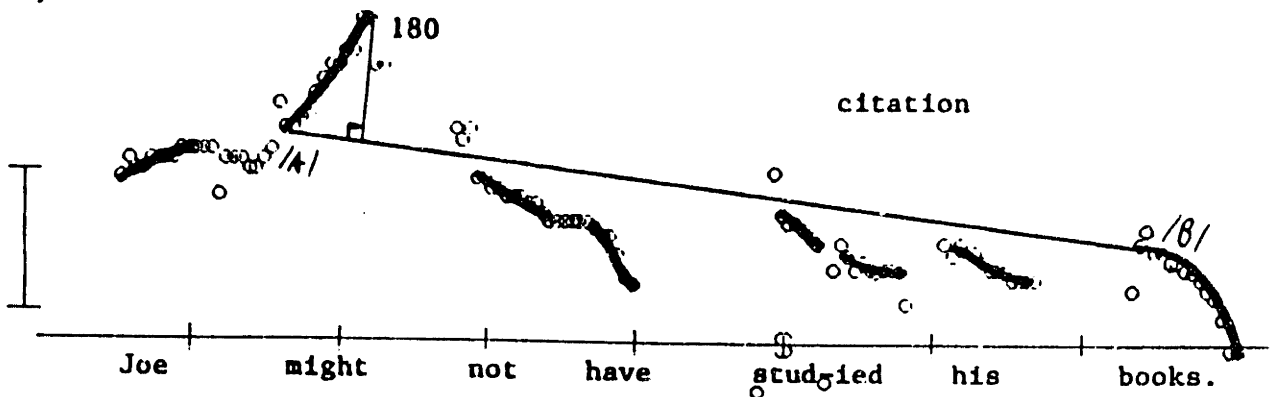
A similar case:

(30)

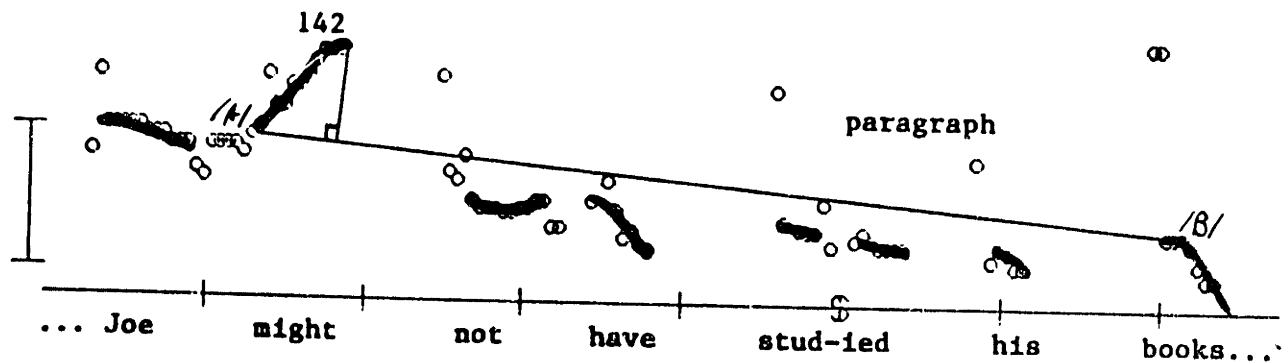


Two examples of modal focus preempting negative polarity focus:

(31)



(32)



Modals are not always on the left periphery of a [+focus] constituent, and thus do not always trigger reversal of the NSR's top line headship. Consider:

(32)

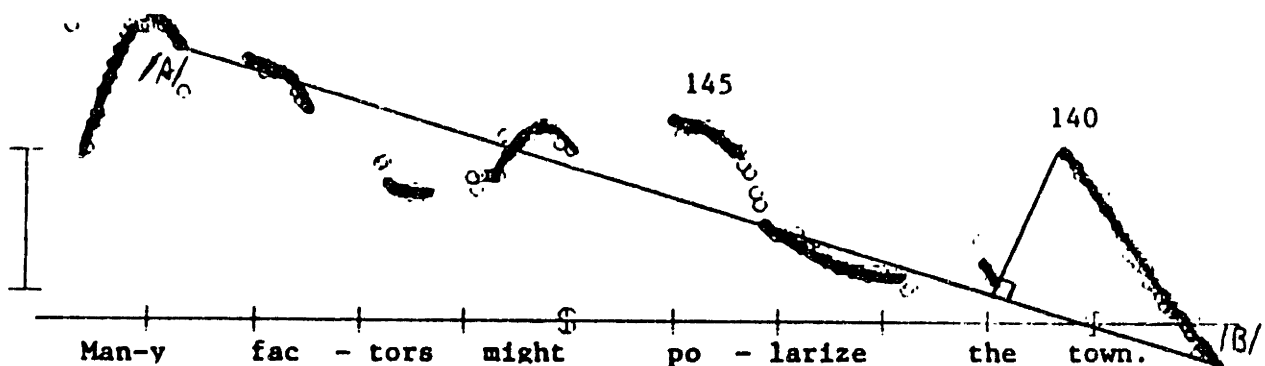
1.6 *
 1.5 (* *)
 1.4 * (* *)
 1.3 * * (* *)

[Many factors might polarize the town][+focus]

In (33), since 'might' was not on the periphery of the focus constituent, it does not trigger headship reversal.

The IG for (33), with reference line drawn, is:

(34)



In (34) we see that the main stressed 'town' has the greatest distance from the reference line.

Headship reversal in the application of the NSR (top line becomes left-headed rather than right-headed) can also be triggered by syntactic means, as in (35):

(35)

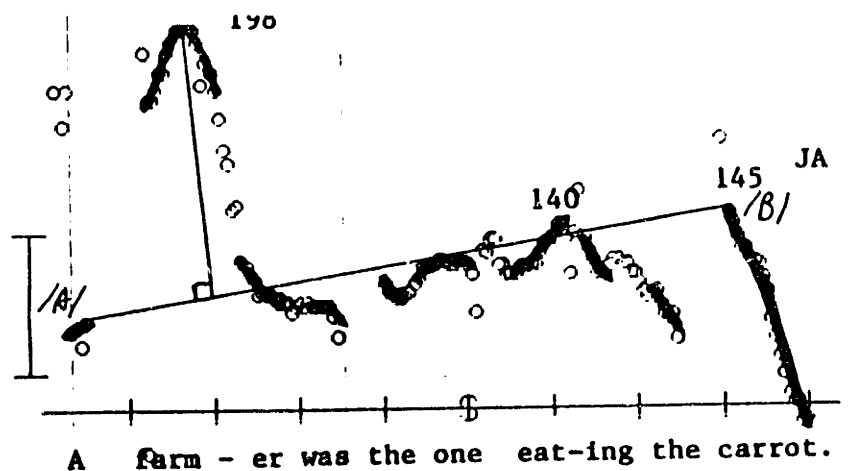
1.5 *
 1.4 (* *)
 1.3 * (* *)

[A farmer was the one eating the carrot][+focus]

Two IGs (from different speakers) corresponding to the metrical grid (35), with reference lines drawn, are shown

below. The method of drawing the reference line in these examples differs from that used above, in that the choice for /A/ is the offset of an unstressed word preceding the main stressed word ('farmer'), rather than the onset of that word. This then is an additional condition in (19).ii: that when an unstressed word exists to the left of the main stressed word in the intonational phrase, its offset point is chosen as /A/.

(36)



(37)

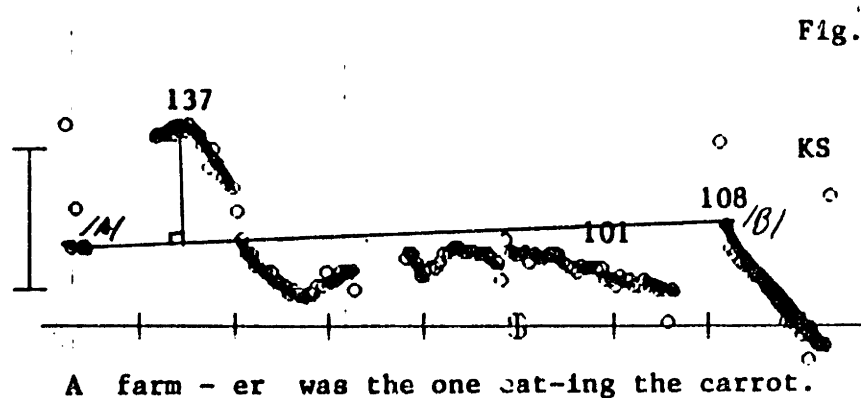


Fig.

Finally, let us consider the famous 'yes-no question' rising terminal intonation. In a yes-no question, we will assume that the NSR applies just as in statements. An example in a discourse-initiating context is:

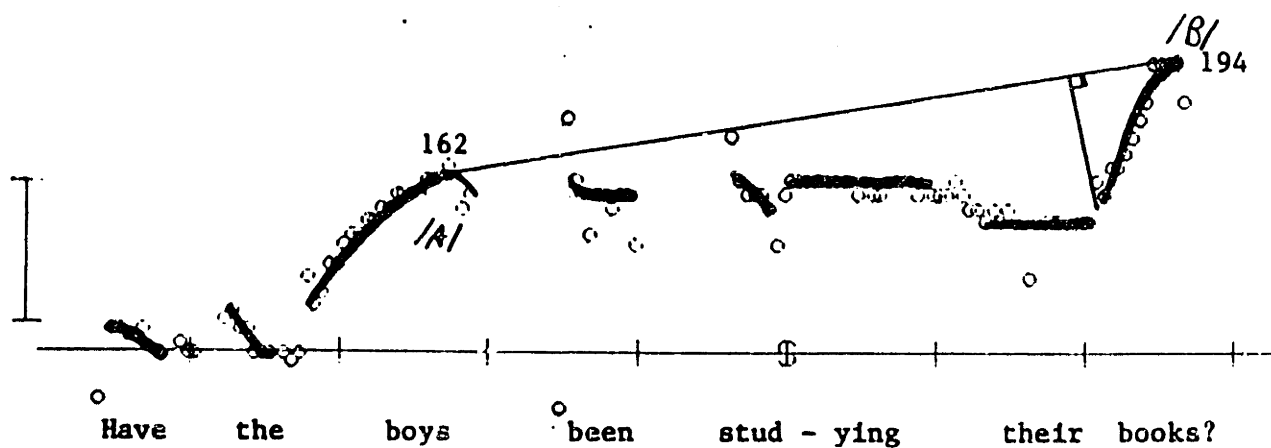
(38)

| | | | |
|-----|----|----|----|
| 1.5 | | | * |
| 1.4 | (* | | *) |
| 1.3 | * | (* | *) |

[Have the boys been studying their books][+F]
[IP1]

We now position the utterance within an IG, and draw the reference line according to the Reference Line Rule (19):

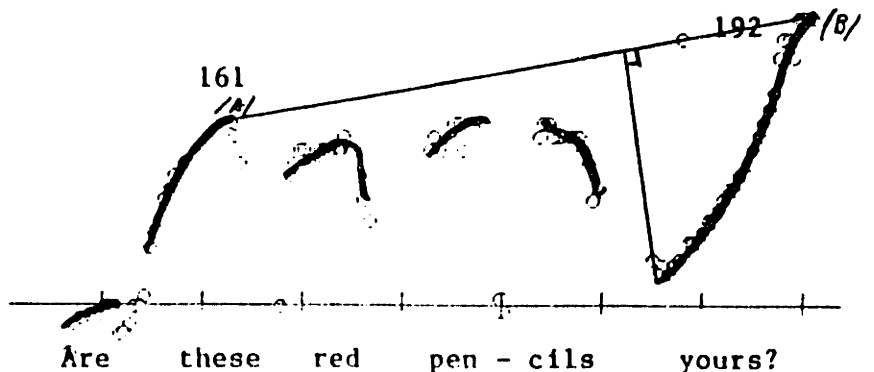
(39)



choose /A/ its offset, as specified by the RLR. Since 'books' is the main stress within IP1, we choose /B/ at its offset. Then we observe that the IG (39) is well-formed, because the main-stressed 'books' has a greater excursion from the reference line than any other item within IP1 to which the Excursion Filter may apply.

A similar example is:

(40)



In summary, we have seen that a model incorporating only the Basic Intonation Rule, which allows maximal freedom in English intonation, and the Reference Line Rule and Excursion Filter, which constrain that freedom to respect metrical structure, gives an account of the well-formedness of a variety of metrical and syntactically diverse utterance types. The model presented here raises many serious conceptual issues, however. In particular, there are outstanding questions of intonational meaning, global vs. local models of intonation, and a host of others. Many of these conceptual questions will be taken up in section 4. below.

4.3 Issues of Duration

The theory of English intonation given above presupposes that the relative duration of words in intonational phrases input to the BIR have been predetermined. The relative durations are assigned based on focus and metrical structure, as discussed in chapter 1. In

this sense, the model reflects the conclusions of the survey of experimental results on the phonetic correlates of stress carried out by Beckman 1986. The general conclusion from experimental research in this area is that duration is the primary acoustic cue to stress in English. This fact would be captured in the present theory by ordering the operation of the duration generator (DG), introduced in chapter 1, prior to the operation of the fundamental frequency generator (FG). Then the segmental durations fixed by DG on the basis of its various inputs (cf. chapter 1) form part of the segmental input to FG (recall from chapter 1 that this is a required input to FG). This is a slight elaboration of the model introduced in chapter 1, as it allows both phonological and phonetic information (determined by DG) to be relevant as the segmental input to FG.

Since duration of each item is fixed or given by a separate prior component (DG), the possibilities for intonational realization on each word are correspondingly limited. The larynx must have time to execute complicated level adjustments, which means that the less segmental material is available (and the amount of material is independently given, in this model), the less varied the intonational realization of a given word can be.

The difference between speech and singing, then, depends on the preassigned durational properties of the intonational phrase. In singing more time may be available

for execution of extreme pitch changes and sustained levels. Nevertheless, any laryngeally generable contour is allowed by the BIR over segmental stretches generated by DG, therefore there need not be a rigid distinction between speech and singing. Further discussion of this topic, however, will be too great a digression, and we therefore put it aside.

4.4 Comparison with a Discrete-level Theory

Perhaps the most interesting, successful, and fully-developed model of English intonation is that of Pierrehumbert 1980, which synthesizes and extends ideas from Liberman 1975, Bruce 1977, and Garding and Bruce 1981 (the Swedish model). Refinements of the Pierrehumbert 1980 model are given in Liberman and Pierrehumbert 1984, and Pierrehumbert and Beckman 1986, 1988. In the discussion that follows, we will concentrate primarily on Pierrehumbert 1980, both for expositional simplicity and because we are attempting a fundamental critique, and the essential character of the model has not been altered in subsequent work. We discuss the discrete-level theory in such detail because this approach arrogates to itself the use of abstract phonological features, rather than contenting itself with purely descriptive entities such as "a relatively high pitch followed by a jump down" (cf. Bolinger 1986, 1989). The description in the last sentence of Bolinger's Profile A is an example of a descriptive entity which is not motivated by considerations of the universal typology of tone systems (apart from the description of intonation itself), and which therefore lacks the potential for category confusion that arises in the discrete-level theory.

4.4.1 Overview of Pierrehumbert 1980

To facilitate the comparative discussion below, it will be helpful to survey the relevant features of Pierrehumbert 1980, henceforth JBP. JBP proposes a model of English intonation that incorporates the following theoretical mechanisms:

(1)

- a. A tonal inventory, consisting of primitive H and L, standing for High and Low.
- b. The following diacritics distinguishing different types of H and L, according to their alignment with syntactic and prosodic structure:
 - '*' - superfixed to T (a tone) indicates a pitch accent, i.e. a tone aligned with a metrically strong syllable
 - '%' - suffixed to T indicates that T is aligned with the boundary of an intermediate phrase or an utterance
 - '+' - suffixed to T indicates that T and the following tone together form a single complex pitch accent
 - '-' - superfixed to T indicates that T is that member of a bi-tonal accent which is not directly aligned to a metrically strong element, or indicates that T is a 'phrase accent' (this aspect of the theory has changed in irrelevant ways in more recent work)
- c. A finite-state automaton indicating the possible linear combinations of a. and b.

- d. 9 primary rules of tonal evaluation for H and L annotated as in b. and configured as in c. These rules are expressed as mathematical formulae which assign an numerical (essentially F0) value to a tone based on the type and numerical value of tones to its left. The rules incorporate conventions governing downstep, declination, and upstep, whose (purported) existence in English is analogized to the case of African tonal phenomena. Some of the tonal evaluation rules will be discussed below where relevant.
- e. A baseline indicating the lowest possible pitch value across an utterance (this has changed somewhat in more recent work and it is not important here).
- f. A convention allowing rightward spread of H tone, under certain conditions not relevant to this discussion.
- g. Conventions governing interpolation between tones which have been evaluated (set to numerical values) by d. (above). While these conventions are not formalized in a general fashion, they state that the interpolation between L and L is monotonic (a straight line), while the interpolation between H and H 'sags' or 'dips', i.e. is not monotonic.

In addition, the model requires access to a metrical grid or some other representation of relative prominence for the elements in an intonational phrase. In practice, the major rules among (1)d. for evaluating tone incorporate a term 'prominence', which is simply a real number that affects tone scaling. This number is determined outside the intonational system.

4.4.2 The Tonal Inventory of English

The claim in (1) that we will focus on is the simplest and in some sense the least controversial, (1) a., the claim that English 'has' a tonal inventory, consisting of primitive H and L. What this means in practice is that:

- (2) Any well-formed intonational contour of English can be described in terms of H and L, with associated diacritics as in (1)b.

For (2) to be a meaningful claim in linguistic theory, it must be shown that the description claimed by (2) to exist for any utterance correlates with other, theoretically well-motivated entities in the grammar of English.

4.4.2.1 PF-external Considerations

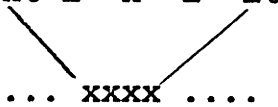
4.4.2.1.1 Syntax

One possible area of correlation of the abstract decomposition of an intonational contour by means of (1)a. and (1)b. is in syntactic structure. A common approach to the study of intonation is an attempt to correlate contour features with syntactic structure. Unfortunately, the use of the term 'syntax' is generally severely compromised in these discussions. A standard example is so-called 'question intonation' in English and many other languages, in which a final rise in the intonation contour tends to correlate with a questioning function (yes/no). The theoretical status of a

syntactic term 'question' is obscure, but in any case it has been shown by Liberman 1975 and others that the 'final rise' can occur on non-questions ('My name is Mark Liberman', with rising intonation, to a dental receptionist on arrival for an appointment) and that a question need not exhibit a final rise. I believe that all other pairings of intonational form (described under any of the popular terminologies) and syntactic structure would succumb to rigorous search for counterexamples.

4.4.2.1.2 Semantics

An apparently more promising area of extra-PF correlations to abstract intonational structure is in the area of semantics. Here there are volumes of literature on the supposed different meanings of different contour types. In this domain it would indeed be significant to be able to say something (for English) like:

(3) H% L* H* L- L%

 ... XXXX ... means the speaker is annoyed

(The 'x's in (3) are intended to informally indicate that some arbitrary text is aligned with the given tune).

The attractiveness of (3) is the implied analogy with lexical meaning in core areas of phonology, where we are able to say, for Mandarin Chinese:

(4) $\begin{array}{c} H \ L \\ \diagdown \ / \\ ma \end{array}$ means 'to scold' $\begin{array}{c} L \ H \\ \diagdown \ / \\ ma \end{array}$ means 'sesame'

Abstracting away from problems in the theory of meaning and translation, there is a certain rigor in the claim (4), upon which phonology is based. There are two obvious problems with any claim such as (3). First, the empirical facts are virtually certain to contradict the claim, and to assert otherwise is usually just a failure of imagination. That is, we can almost certainly find a situation under which the tune in (10) expresses some other attitude (and, of course, we can almost certainly find another tune that could express annoyance in some context (this kind of criticism for this tune is given by JBP herself, p. 61)). Second, the concept 'annoyed' is essentially pragmatic. By that we mean that it lacks correlates in the core modules of linguistic theory and it cannot be precisely specified or talked about in linguistic theory. The idea that one could develop an intonational lexicon incorporating entries like (3) is extremely attractive, but even the most comprehensive analysis of intonational meaning yet attempted, Bolinger 1989, synthesizing and extending much previous work, does not, and, we claim here, could not in principle, go beyond

anecdotal discussion, and postulation of certain probabilistic tendencies.

Nevertheless, it may be argued, the phonological specification of a 'tune' in terms of the primitives provided by (1)a. and b. provides a very useful language of description. After all, it may generally be the case that (3) is a good characterization of an admittedly not-watertight correspondence, a correspondence that lacks status in the core modules of grammar, but that is nevertheless worth talking about. There may be other such loose correspondences, such as a 'contradiction contour' (cf. the penetrating criticism of this supposed example of rigorous intonational meaning in Bolinger 19) and others. And given that any text may substitute for 'xxxx' in (3), why should we be deprived (as the theory proposed in section 1 would have it) of the simple and abstract descriptive language that (1)a. and b. provides? The answer is that we should not be so deprived, even accepting the theory proposed in section 1. This is a question related to current concerns in syntactic theory. Chomsky (1987 class lectures) has suggested that of the four configurations of [+/- language specific] and [+/- construction specific] in discussions of syntactic phenomena, only [-construction specific] phenomena have an existence in linguistic theory (UG). That is to say, for example, there is no role for a construction-specific term 'passive' in the theory of

universal grammar, rather there are various possible configurations of syntactic elements that are licensed by various principles and parameters, none of which makes reference to a term 'passive'. However, it is obvious that the term 'passive' has a certain utility in facilitating discussion. It is this status that can be claimed for the abstract tone decompositions of JBP (that is, in the discussion of extra-PF correlates; PF-internal considerations come up in the next section). The term 'contradiction contour' thus corresponds, in its theoretical status, to the term 'passive', and the tonal description of that contour, in terms of (1) corresponds to a phrase-structure rewrite rule for passive, or a transformation 'Passive'. Neither of these elements to which we analogize has status in current syntactic theory; rightly so, we believe. But in accepting (1) as a language of semi-formal description, one has to be careful.

Consider in this regard the theory of Liberman 1975, which allowed four level tones (rather than JBP's two) as the English primitive tonal inventory (in practice, H (high), M (mid) and L (low) were primarily used for analysis). Liberman provides an analysis of the 'vocative chant' (a tune for which, like the 'contradiction contour' we would claim the same theoretical status as the term 'passive'), that is, a 'calling contour' usually associated with the text of a name. This contour is pre-theoretically

described as "a fall from a peak on the nuclear stress partway to the baseline" (JBP). In the 4-tone analysis of Liberman 1975, it is abstractly characterized as: H M (high mid). In the 2-tone analysis of JBP it is characterized as: H*+L- H- L%, with a complex downstep rule (one of the group (8)d.) guaranteeing the relative levels of the first and second H tones. Now that we are reduced to arguing for abstract tonal decomposition of contours on the basis of simplicity and generality of semi-formal description of loose syntactic and semantic correlations, we have no reason to prefer the JBP description of the vocative chant over Liberman's formulation, in fact rather the opposite, one might think¹. Yet Liberman's theory suffers from the same weaknesses described above for any primitive tonal description of English intonation. As the hope of rigorous syntactic and semantic correlations to JBP's primitive descriptions recedes, the possible utility of these abstract specifications in terms of PF operations is increasingly highlighted, and to this issue we now turn.

¹ Note further that to the extent JBP makes use of the downstepping tonal primitive accent (H*+L) both to derive the 'mid' in the vocative contour and for semantically less salient downstepping effects, the principled separation of abstract phonological characterization of the 'intonational lexicon' from details of phonetic implementation is called into question.

4.4.2.2 PF-internal Considerations

We have seen that any attempt to establish correlations between the abstract contour decompositions provided by (1)a. and b. and extra-PF constructs in linguistic theory is awkward at best. Let us now consider the utility of the English primitive tonal inventory for the organization of processes in PF.

4.4.2.2.1 The Basic H/L Distinction

The central question here will be: what is the phonological justification for postulating two distinct primitive entities, H and L? JBP gives three justifications (to facilitate discussion, the comparison will be cast in terms of differences between H* and L* (high and low pitch accents). The generalizations to H and L in other configurations have the same character, but are slightly more complex in irrelevant ways):

(5)

- a. L* is always lower than H* in the same context.
- b. The phonetic value of L* decreases if its prominence is increased, while the phonetic value of H* increases if its prominence is increased.
- c. L* and H* are treated differently by interpolation rules. That is, the interpolation between L* and L* is monotonic (a straight line is drawn between them), while interpolation between H* and H* is non-monotonic (a sagging line, like a suspension bridge, is drawn between them).

The problem with evaluating (5) is that the existence of H and L is essentially presupposed (perhaps based on the supposed correlations with syntax and semantics discussed above). To begin the discussion we need to step back. We will consider a simple hypothetical case intended to show that the abstract tonal quality of pitch accents in a contour cannot be uniquely determined in JBP's system, and hence that the empirical foundations of JBP's theory are unsteady.

Consider an English intonation contour in a very simplified abstract graph space that allows 5 'units above the baseline' for variation of F0. The y-axis thus represents '(arbitrary) units above the baseline' (pitch or F0 scale) and the x-axis represents time. 'Prominence' is a term in the tone scaling equations (1)d, which will be relevant in analyzing contours. For the following discussion, we will allow 'prominence' to range over the integers 1 to 5, in accordance with the possibilities for its realization in the tone space (no explicit range for prominence is given by JBP). Consider an utterance in this hypothetical tone space:

(6)

| | | |
|---|-------|---------|
| 5 | | |
| 4 | | |
| 3 | | |
| 2 | | |
| 1 | t(i)* | t(i+1)* |
| | John | left |

In (6), we see a (hypothetical) utterance, with each word specified for a tonal pitch accent. The pitch accent status of the tones is indicated by the '*' notation. The pitch accents are given as abstract tones, suppressing the (putative) H/L distinction. Ignoring the interpolation between pitch accents for the moment, imagine that we are given an intonational observation (F0 measurements) in terms of (6) and we wish to use the tools provided by JBP to assign a tonal description. There are five different placements of $t(i+1)$ with respect to each of the five different placements of $t(i)$ on the vertical scale in (6), for a total of 25 unique arrangements. For any placement of $t(i)$ relative to $t(i+1)$, can we, without referring to any extra-theoretical intuition, uniquely determine the quality of $t(i)$ and $t(i+1)$, in terms of H and L?

Here we encounter a crucial property of JBP's tonal implementation system. In preparing to evaluate (assign numerical values to) the abstract (H and L) tones in an abstract intonational description, "the [numerical, F0 - S.M.] value of the first pitch accent in the phrase is a free choice, governed by pragmatic or expressive factors". All subsequent operation of the nine tonal evaluation rules will depend on this choice. If one accepts the presupposed H and L distinction, this is not overly serious, as the H and L distinction itself is referred to by the rules. But that

is a theory-internal factor. Here we are attempting the reverse process, that is, we are attempting to discover how many unique assignments of H and L to $t(i)$ and $t(i+1)$ are possible, given any arbitrary position of observed pitch accents on the pitch scale. If the assignment of H and L cannot be uniquely determined, the empirical basis of JBP's theory may be called into question.

The rule for evaluating H pitch accents is the following (one of the battery of nine tonal evaluation rules referred to in (1)d.):

(7) H* evaluation rule

$$/H^* i+1/ = /H^*i/ \frac{\text{Prominence (H}^* i+1)}{\text{Prominence (H}^*i)}$$

(7), given an initial value of the leftmost pitch accent in a string, uses the prominence value of the pitch accent to the left (i) of the current pitch accent ($i+1$) and the prominence value of the current ($i+1$) pitch accent itself to assign a tone value (in terms of F0 or units above the baseline) to the current ($i+1$) pitch accent. The rule for evaluating L pitch accents, which works similarly, is:

(8) L* evaluation rule

$$/L^* i+1/ = /L^*i/ \frac{\text{Prominence (L}^*i)}{\text{Prominence (L}^*i+1)}$$

We can view these formulas as a mechanism to help us determine the empirical limits on the H or L status of a given tone. Tables (9) and (10) below show that, by manipulating the prominence values (which are free to vary for expressive purposes) and the initial pitch accent value (which is free to vary for expressive purposes) we can achieve any height for any either tone, using either rule (7) or (8). We therefore conclude that, in the absence of any independent justification for an assignment of H or L status to a given observed pitch accent, the PF implementation system does not constrain the choice. Hence, given that the implementation system does not rely on the distinction, and given that there is no PF-external support for tonal descriptions in these terms, the existence of this distinction is called into serious question. Although the possible relations should be clear from the formulae given, all possible values are computed in (9) and (10):

(9) Assignments of initial values to $t(i)$, prominence of $t(i)$, and prominence of $t(i+1)$ that give any placement of $t(i+1)$ on the tone scale, using the H^* evaluation rule (7)

| $t(i)$ | $PRt(i)$ | $PRt(i+1)$ | $t(i+1)$,with (7) |
|---------|----------|------------|--------------------|
| 1 H^* | 1 | 1 | 1 H^* |
| 1 H^* | 1 | 2 | 2 H^* |
| 1 H^* | 1 | 3 | 3 H^* |
| 1 H^* | 1 | 4 | 4 H^* |
| 1 H^* | 1 | 5 | 5 H^* |
| 2 H^* | 2 | 1 | 1 H^* |
| 2 H^* | 1 | 1 | 2 H^* |
| 2 H^* | 2 | 3 | 3 H^* |
| 2 H^* | 1 | 2 | 4 H^* |
| 2 H^* | 2 | 5 | 5 H^* |
| 3 H^* | 3 | 1 | 1 H^* |
| 3 H^* | 3 | 2 | 2 H^* |
| 3 H^* | 1 | 1 | 3 H^* |
| 3 H^* | 3 | 4 | 4 H^* |
| 3 H^* | 2 | 5 | 5 H^* |
| 4 H^* | 4 | 1 | 1 H^* |
| 4 H^* | 2 | 1 | 2 H^* |
| 4 H^* | 4 | 3 | 3 H^* |
| 4 H^* | 1 | 1 | 4 H^* |
| 4 H^* | 4 | 5 | 5 H^* |
| 5 H^* | 5 | 1 | 1 H^* |
| 5 H^* | 5 | 2 | 2 H^* |
| 5 H^* | 5 | 5 | 3 H^* |
| 5 H^* | 5 | 4 | 4 H^* |
| 5 H^* | 1 | 1 | 5 H^* |

(10) Assignments of initial values to $t(i)$, prominence of $t(i)$, and prominence of $t(i+1)$ that give any placement of $t(i+1)$ on the tone scale, using the L^* evaluation rule (8)

| $t(i)$ | $PRt(i)$ | $PRt(i+1)$ | $t(i+1)$,with (8) |
|---------|----------|------------|--------------------|
| 1 L^* | 1 | 1 | 1 L^* |
| 1 L^* | 2 | 1 | 2 L^* |
| 1 L^* | 3 | 1 | 3 L^* |
| 1 L^* | 4 | 1 | 4 L^* |
| 1 L^* | 5 | 1 | 5 L^* |
| | | | |
| 2 L^* | 2 | 1 | 1 L^* |
| 2 L^* | 1 | 1 | 2 L^* |
| 2 L^* | 3 | 2 | 3 L^* |
| 2 L^* | 2 | 1 | 4 L^* |
| 2 L^* | 5 | 2 | 5 L^* |
| | | | |
| 3 L^* | 1 | 3 | 1 L^* |
| 3 L^* | 2 | 3 | 2 L^* |
| 3 L^* | 1 | 1 | 3 L^* |
| 3 L^* | 4 | 3 | 4 L^* |
| 3 L^* | 5 | 3 | 5 L^* |
| | | | |
| 4 L^* | 1 | 4 | 1 L^* |
| 4 L^* | 1 | 2 | 2 L^* |
| 4 L^* | 3 | 4 | 3 L^* |
| 4 L^* | 1 | 1 | 4 L^* |
| 4 L^* | 5 | 4 | 5 L^* |
| | | | |
| 5 L^* | 1 | 5 | 1 L^* |
| 5 L^* | 2 | 5 | 2 L^* |
| 5 L^* | 5 | 5 | 3 L^* |
| 5 L^* | 4 | 5 | 4 L^* |
| 5 L^* | 1 | 1 | 5 L^* |

If the degrees of prominence used above had rigorous independent justification, the situation would not perhaps be beyond salvage, because the indeterminism in labelling would be much reduced (though not eliminated). However, it is very difficult to see what empirical constraints could be brought to bear in establishing independently justified levels of expressive prominence. We therefore claim that the

use of expressive prominence in JBP is an ultimately misguided attempt to buttress an otherwise questionably motivated system of discrete level tone categories.

4.4.2.2.2 Differences in Interpolation

We see from the tables above that for the example given it is quite difficult to use the theory to diagnose the tone labels. But this may be seen as an ungenerous example. JBP notes that when identical pitch accents are sufficiently close in time, the interpolation between them is usually obscured. That is, for example, in a string of two nearly contiguous H* pitch accents, the interpolation is trivial, and doesn't have time to 'sag' as it presumably otherwise would. Thus, had we chosen an example like:

(11)

| | | |
|---|-------|-------------------|
| 5 | | |
| 4 | | |
| 3 | | |
| 2 | | |
| 1 | t(i)* | t(i+1)* |
| | John | went to the store |

with pitch accents farther removed in time, the interpolation between the two pitch accents would reveal the tone labels involved. Recall from (5) c. above that the interpolation rules are supposed to treat H* and L* differently, in that interpolation between H* and H* is non-

monotonic, while interpolation between L* and L* is monotonic. Although the exact rule for calculating the interpolation in a case like (11) is not given, we can make the following observations.

Either there will be sagging in the contour between $t(i)^*$ and $t(i+1)^*$ or there will not. If there is no sagging, we cannot determine whether this apparent monotonicity is due to the end tones being L* or whether it is due to the end tones being H*, with a sufficient number of intervening contiguous H*s to suspend the sagging that would otherwise occur. Since there is no independent characterization of what level of abstract metrical structure qualifies a lexical content word for a pitch accent, we cannot appeal to metrical structure to resolve this question.

If there is sagging, we cannot determine whether the sagging is due to simple interpolation between end-point H* targets, or due to intermediate interpolation between the initial H* and some intervening L* target(s). The reason this uncertainty exists is that there is no independent specification for the meaning of L. While H means 'how high above the baseline', L can sometimes be realized on the baseline. But to give the claim about the difference in interpolation character some content, JBP must hold that L can also be realized above the baseline, or else the difference in interpolation character would fall out from a

single scale system - in which, when the end points are both at a minimum, the interpolation between them must of course be monotonic. Without an understanding of what counts as an L*, we cannot resolve the above noted indeterminacy. What would count as a solution to this indeterminacy, and what would give powerful, interesting support to the notion of an L* categorially distinct from H*, would be mirror-image symmetry between the interpolation behavior of sequential H* accents (sagging interpolation) and L* accents (rising or puffing interpolation). Unfortunately, as JBP points out (p. 70), this symmetry does not exist, so one of the 3 major arguments for a categorial distinction between H* and L*, (12)c., is dubious and can in fact be thought of as an argument against any such distinction.

4.4.2.2.3 Downstep

The discussion so far has been limited to simple cases of sequentially arranged pitch accents. JBP's theory, however, is complex and covers a great deal more ground. One area to which considerable attention is devoted is downstep/catathesis/declination - the process whereby, on analogy to African tonal process, a H tone following a L tone will be realized at a lower level than it would be had no L tone preceded.

Leaving the African analogy aside, it is clear in discussing English that provision must be made for the optionality of these processes (cf. Umeda 1982). Even if we accept the existence of H/L labelled pitch accents, it should be clear that downstep need not apply to a series of these. We would claim, that is, that a contour consisting of a series of accents to which downstep had not applied (though JBP's rule context was met) will be acceptable and interpretable in some context, just as a contour consisting of a series of accents to which downstep has applied (in JBP's rule context) will be acceptable and interpretable in some context². Clearly then, this process is not an exceptionless feature of phonetic implementation. When we try to characterize the phenomenon in phonological terms, we will want to describe it in terms of correlations with well-motivated constructs of either phonology or syntax/semantics. We will have the same problems for the characterization of downstep that we encountered for the characterization of abstract tonal descriptions discussed in the sections above. That is, while we may use such pre-theoretic terms as 'normal declarative utterance' to

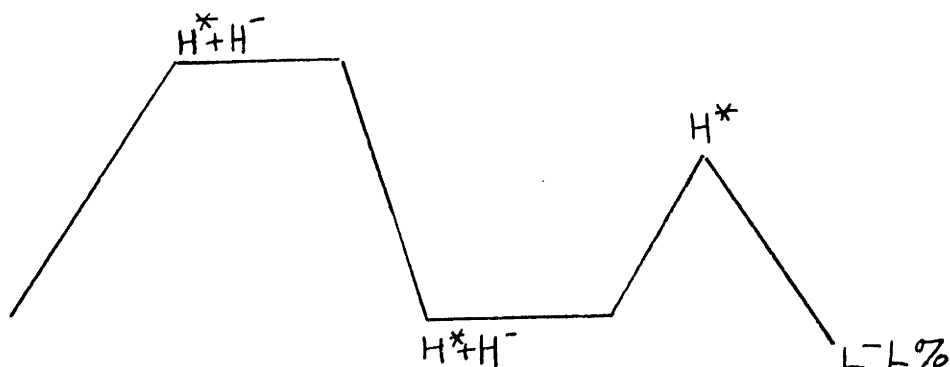
² This discussion is somewhat oversimplified, as the pitch accents that trigger downstep in JBP are abstract and are not intended ever to be surface realized as such. The claim that, in empirical terms, declination across a series of accent is strictly optional remains, and JBP would not dispute this (because here finite-state automaton allows for free generation of these on a par with any other choice of pitch accent quality). However, in her terms, this optionality is expressed by the choice of pitch accent itself, rather than the optionality of application of the downstep interpretation rule.

characterize the syntactic/semantic correlates of the downstep process, it is probable that this kind of statement has no theoretical status. Within the phonology, if the H/L units which condition the hypothetical downstep process, and in terms of which it is described, had independent status, the fact that the statement of downstep crucially relies on them would be impressive corroborating evidence for the H/L distinction. Since, however, H/L (arguably) cannot be independently justified, the formulation of downstep in terms of these primitives must be seen as an unmotivated (though ingenious) feature of JBP's presentation.

4.4.2.2.4 Impossible Contours

Since JBP's theory is more expensive and apparently more restrictive than the theory proposed in section 1 above, (though as we have seen, its greater restrictiveness is partly an illusion), we expect JBP to present impossible contours, that is, physiologically possible contours that do not conform to what the machinery in (1) can generate. Six such examples are discussed by JBP, and all are abstract, such as that shown in (12) below:

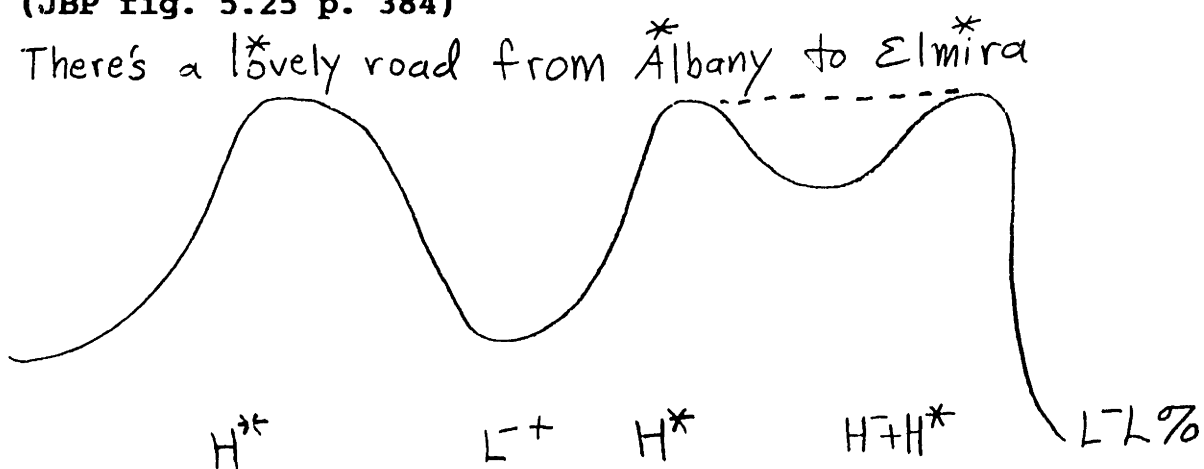
(12) An impossible intonation pattern in English ?
 (JBP fig. 5.20 p. 381)



These cases are very difficult to discuss. In the theory proposed here, they are ruled out in the form shown by the condition on laryngeal naturalness incorporated in the BIR (discussed in section 1). That is, the claim here is that the abstract impossible contours of the type (12) owe their impossibility to their continuity and angular precision, which we will claim is impossible for the larynx to achieve over ordinary speech-timed segments. Taking an example like (12), if we broke the line for non-voiced segments, introduced F_0 perturbations due to segmental effects, and acknowledged the indeterminacy of tone labeling discussed above, it is not clear in what sense the contour could still be said to be impossible.

An interesting abstract example of this type is shown below:

(13) An impossible intonation pattern in English ?
 (JBP fig. 5.25 p. 384)



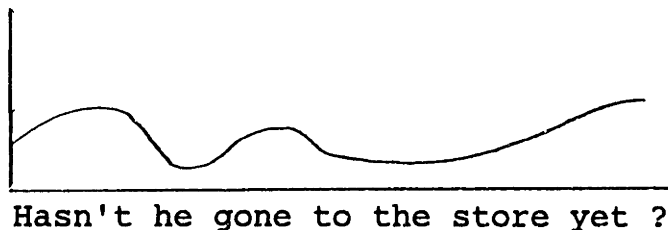
For (13) JBP gives a precise indication of the source of the trouble. Evidently the 'oddness' of (20) results from the sustained high level across 'to' between the pitch accents on 'Albany' and 'Elmira'. It must first be noted, in keeping with the basic approach of the theory being defended here, that referring to the dotted contour of (13) as 'odd' is a vague claim, that probably reflects, as is usual with this kind of claim for intonation, some kind of statistical corpus norm (usually thought to be inadmissible in generative linguistic theory) or simply a failure of contextual imagination. Be that as it may, we observe that no prominence values are given for (13). We may assume that 'Albany' and 'Elmira' are more prominent than 'to'. In the theory defended in section 1, any 'oddness' of the dotted contour in (13) would be related to its failure to conform to the Excursion Filter in the relation between the prominent 'Elmira' and the non-prominent 'to'.

4.5 Cross-dialect Intonation

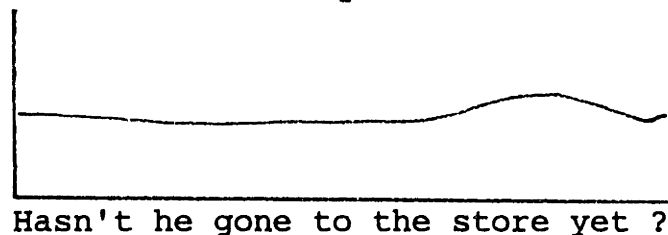
One potentially fruitful domain of inquiry for counter-examples to the claim of maximal laryngeal generation discussed in section 1 is cross-dialect intonation. A typical example would be the difference in question intonation between standard American English and the question intonation of certain British dialects. This difference is given pre-theoretically in (20) below:

(14)

a. American English question



b. British dialect question



Clearly the difference between the contours in (14) a. and b. can be semiformaly described, with distinct resulting descriptions, in any of the currently recognized models of English intonation.

The relevance of these cases for the claim being made here is that in (14) b., for example (we could frame this the opposite way, of course), we have an intonational utterance, easily generable by the BIR, that might be thought distinctly odd in the dialect of (14) a. This feeling of oddness is then a good candidate for a linguistic judgement setting limitations on the operation of the BIR in American English. Other examples of this type could be cited. The interest they have is in pointing out how careful one must be in evaluating judgement evidence in linguistic theory. The answer to the challenge posed by (14) is the prediction that the judgement of oddness of (14)b. in the context of dialect (14)a. relies on knowledge (not necessarily completely conscious) of the existence of dialect (14)b. Thus the (perhaps unconscious) judgement of oddness (if it could be elicited at all) must take the form: 'that sounds British'. This distinction is crucially different from the standard sorts of syntactic well-formedness judgements (although of course cross-dialect syntactic judgements will have this character as well). That is, one need not contrast (explicitly as in (14) above, or implicitly, as in 'knowledge' of British English) the Irish English utterance in (15)a. with the American English (15)b. in order to get a well-formedness judgement on (15)a. in a mono-dialectal speaker of American English:

(15)

- a. I'm after using the copier.
- b. I've finished using the copier.

The claim advanced here is that (14)b. will be neither impossible to interpret nor to associate with a context for speakers of dialect (15)a. as long as the implicit/explicit comparison that leads to the 'that sounds British' conclusion does not interfere in the testing. In addition, the contour of (14)b. will not be judged ill-formed under those conditions. Therefore, the dialectal 'knowledge of language' that underlies the judgements on (14) is of a stylistic character, while the dialectal knowledge that underlies judgements in (15) must be treated in the core modules of grammar. While this prediction may be subtle to deal with in practice, in principle it can be tested.

Note crucially that we are by no means claiming that any language, including dialects of English, lacks intonational rules, in the stylistic sense referred to above. We are claiming that American English lacks tonal rules in the sense familiar in core areas of phonological theory, and that where candidates for such rules exist (as in the area of association to metrical structure, treated in section 1 above), they must be stated in a vastly different form than current theories of intonation admit.

It is this free quality of American English intonation, captured in the model based on the BIR, that has so puzzled generations of comparative researchers. The claim here is that recognizing this freedom explicitly and working back from it only where substantial motivation forces such retreats provides a more insightful model of American English intonation.

4.6 Intonation and Gesture

Some researchers (notably Dwight Bolinger, 1986, and through the years) have observed the close parallels between intonation and gesture as expressive mechanisms. It appears that this comparison is apt, and very revealing. In particular, the use of concepts from core phonology to describe intonation (e.g. H/L tonal primitives, downstep rules, etc.) is analogous to the use of concepts of structural linguistics in the analysis of gesture and body motion communication carried out in kinesics. Analysis of intonation and body communication are doubtless valuable and revealing activities, and the borrowing of properly grounded linguistics concepts for another use is not necessarily to be deplored. But one has to be careful of the implicit claims that arise from this kind of appropriation. We claim here that intonational meaning has the same status as kinesic meaning, and that neither is a linguistically specifiable concept, in the sense of the core modules of UG.

Consider an example of kinesic analysis (the diacritics above the example sentences are kinesic morphemes):

... any normal child of six will recognize that, though they are made up of the same words, he has he has no difficulty in distinguishing between:

"She is a nice girl" and "She is a nice girl"?

Or to put it differently, although he cannot tell you exactly how it is done, any normal American informant will tell you that you have given quite a different message when you say "She is a nice girl?"# and "She is a nice girl."|| The difference between those two sentences takes place in about 3 to 5 milliseconds. the way in which the terminal pitch of the two sentences is handled makes the first what we might call a declarative statement and the second an interrogatory or doubt statement. ... When these linguistic particles are put together in a communicational frame, in actual speech one does a series of things with one's body. In speaking these sentences, I do not have very much choice about which movements I make. Each of these sentences, within its context, requires a very special set of movements. To review,

"She is a nice girl" is marked by a set of head movements which take place over the 'She', the 'nice', and the 'girl'. In this example I mark the sentence by lowering my head. I can just as easily do this with my eyelids, with my hand, or even with my entire body. These kinesic markers, as we have termed them, can be seen too in the contrast sentence

"She's a nice girl" in which I cross-reference with with the markers just as I can with drawl in my voice over the "She's", the 'nice', and the 'girl'. ... the "sweep" marker over the 'nice' indicates that I am not totally enthusiastic about the young lady. Comparably, the example, "She is a nice girl"|| contains a series in which I may knit my brows over 'nice' and make a slight lateral and upward movement over 'girl'. ... the meaning [varies] in a consistent manner with each significant vocalic or kinesic shift. (Birdwhistell 1970, p. 17)

Another example:

There is clearly a difference between the order of statement which I make about myself when I close my lids with no perceptible duration of holding at the point of closure, and when I close the lids at the same rate of speed, allowing about a quarter of a second duration of the closure. Or again, contrast these with the situation in which I close my eyelids much more slowly, leave them closed for a duration, or close them slowly and leave them closed for a hardly perceptible duration. (Birdwhistell 1970, p. 18)

Obviously there is nothing objectionable about such research efforts, and doubtless they have resulted in a rich fund of knowledge concerning the possibilities of human interaction behavior, and probably this fund of knowledge will come into practical use someday for building life-like dynamic mannequins. And if this research finds borrowed concepts or tools of linguistic research valuable, so much the better. It is, however, important for linguistic research, concerned with delimiting the core modules of grammar, not to borrow in the reverse direction and assume that any behavior that can be described by linguistic methods (of a given era) is part of the genetic language endowment. There is, of course, a real language system that happens to make use of gesture, namely sign language, and the subject matter of kinesics has the relation to sign language that the study of intonation has to spoken language.

In chapter 1 above, we present the precise relation between gesture and fundamental frequency generation defended in this approach.

4.7 Intonation and Linguistic Theory

The claim of the present proposal is that the maximally simple intonation model will consist of just the following:

(16) Components of the intonation model

- a. Metrical/focus structure
- b. Basic Intonation Rule
- c. Reference Line Rule
- d. Excursion Filter

It is clear that this theory does without various pieces of machinery that have been crucial to previous accounts, e.g. a primitive tonal inventory (Pierrehumbert 1980, Liberman 1975, Ladd 1883), starred tones (Pierrehumbert 1980, 1986), linked tones (Pierrehumbert 1980, 1986), special mathematical routines for catathesis and pitch range calculation (Pierrehumbert 1980, Beckman and Pierrehumbert 1987), unitary contours (Bolinger 1986, 1989), features on primitive tones (Ladd 1983) and much else. The claim of the present theory is that any intonation contour developed in accordance with (16) can be accepted and judged as well-formed by native speakers of American English.

This is in one sense a very strong claim and in another sense a very weak claim. It is a strong claim in the sense that very little theoretical machinery is postulated to account for a large amount of data that is thought to be fairly complex. It is a weak claim in that it makes no attempt to cover normative properties. This dual nature of the present theory is in accordance with recent conceptions of the proper form of linguistic theory (Chomsky 1986, Chomsky class lectures 1987, 1988, 1989), based on a longer tradition of autonomous modular interaction (Hale et al 1977, Chomsky 1981). In this conception, a given module generates forms maximally (modulo its internal constraints) and filtering may be performed by other components. It appears that intonation provides a paradigmatic example of this kind of functioning.

If we wished to characterize statistical properties of utterances, we would want to give greater structure to our intonational model. Likewise, if we wished to provide a vocabulary for stating functional correspondences between intonational properties and semantics or pragmatics, we would again wish to provide a richer mechanical system. However, even those theorists (e.g. Pierrehumbert 1980) who do provide richer theoretical structure have generally concluded that the search for invariant relations between either syntactic, semantic, or pragmatic properties of utterances is not viable, or at least that such

correspondences cannot be stated as part of the intonational model. This insight is implicit in the finite-state automaton organization of 'pitch accents' in Pierrehumbert 1980. Furthermore, this insight is compatible with recent syntactic theory, where constructions such as 'passive' no longer have primitive status, but are seen as convenient labels for the blind operation of principles and parameters. The claim here is that although an important step in the recognition of the difference between English intonation and the use of tone in lexical tone languages was taken by Pierrehumbert 1980 in her finite-state automaton model of pitch accent assignment, the difference is brought out more clearly and at lesser theoretical cost by the present proposal.

The model in (23) would probably not be the most suitable for driving a speech synthesizer, just as a computer-based natural language parsing or generating system will require more elaborate and more normative principles of syntax than those provided by the abstract well-formedness statements of current principles and parameters theory. But those kinds of functional considerations ought to have no bearing on the form of linguistic theory.

Appendix 1: Metrical Constituent Structure

The formal system of metrical representation of prominence relation adopted in this thesis is the bracketed grid theory of Halle and Vergnaud 1987. Below we present some essential elements of this theory, largely following and condensing the excellent summary in Sietsema 1989, chapter 1. The reader is referred there and to Halle and Vergnaud 1987 for numerous important details that are not treated below.

1. The elements of metrical structure

Metrical structure consists of regular groupings of phonological substance. The phonological substance concerned is usually chosen as either syllables (as in English) or morae (as in Manam). Whichever is unit is chosen, each such unit in a given lexical item is represented on line 0 of a metrical grid by an asterisk. Thus, the initial representation of a stress derivation in English lexical items would be something like (with syllable boundaries in the word marked by parentheses):

(1)

| | | |
|---------|--------------|-------------|
| line 0: | * * * * * | * * * * |
| | apalachicola | Mississippi |

while the initial representation of Manam lexical items (where the mora is the unit of line 0 representation, or projection) would be:

(2)

| | | |
|---------|--------|---------|
| line 0: | * ** | * * ** |
| | so ?ai | malabon |

Each asterisk on line 0 in (1) and (2) represents a unit of projecting phonological substance, and the prominence derivation must respect these, in that prominence (referred to as stress henceforth) must be chosen on all higher grid line from only the projected positions on line 0.

2. Metrical constituents

Given the possible prominence or stress-bearing elements in a lexical item which are projected on line 0, metrically significant groupings may be established. In syntax, words may be shown (by means of movement and extraction tests, selectional and subcategorizing restrictions, among other diagnostics) to have a greater affinity with certain of their neighbors in a string than with others. A similar concept is applied in metrical theory, where affinities among line 0 elements (asterisks) are postulated, based on native speaker intuitions of rhythm in word pronunciation. Not only do these groupings capture native speaker intuitions, they establish the potential elements for projection and representation on lines higher than line 0 in the grid, because, just as in syntax, a constituent is limited to a single unique head at higher

levels of structure. This means that, for each constituent marked out among the line 0 asterisks, there will be a single head projected onto line 1, and again any further stress derivational processes that apply thereafter are limited to the consideration of the heads so marked (the line 0 asterisks in each constituent which were not chosen as heads are not directly represented on line 1 and hence are invisible for operations on line 1 or any succeeding grid line).

For the English example above, the line 0 constituents and their heads projecting on line 1 are shown below:

(3)

| | | | | | | |
|---------|------|------|------|----|-------|--------|
| line 1: | * | * | * | | * | * |
| line 0: | (* | *) | (* | *) | (* | *) |
| | apal | achi | cola | | missi | ssippi |

The groupings of line 0 asterisks in (3) represent metrical constituent structure on line 0. We observe that there are two elements in each constituent, reflecting the native speaker intuition of alternating strength relation among adjacent syllables in these words. On line 1 in (3), we see that the syllable which is felt by native speakers to be 'stronger' in each pair of line 0 asterisks grouped as a constituent is marked on the next higher line, grid line 1. The asterisks appearing across line 1 are the heads of the line 0 constituents, 3 such heads in the first word, and 2 in the second, corresponding to the number of line 0

constituents (recall that every constituent must have exactly one head to represent it on the next higher line).

When the leftmost element of an inferior constituent is chosen for projection to the higher line (headship), we say that the inferior constituent is 'left-headed', and likewise for 'right-headed' constituents. It is an interesting fact about prominence and stress relations in phonology that these 'edge' elements are usually selected to be the heads on higher line. In (3), since the constituents had only two members, the head must be an edge element, in any case.

3. Well-formedness Statements for Grid Lines

There is a formal system for manipulating the asterisk and constituent boundaries introduced above, and this system has the advantage of constraining the possible operations. This system makes an important distinction between statements about well-formed bracketed grid representations in a given language and rules that construct these representations. Every line in the grid will have an associated statement of parameters defining well-formedness on that line. An example of such a parameter is:

- whether or not the head of the constituent is separated from its constituent boundaries by no more than one intervening element.

What do we mean by a head being separated from its constituent boundaries by no more than one intervening element? It is an interesting fact about stress that in

general metrical constituents are either alternating pairs, as we saw in (3) above, or long strings of indefinite size. Rarely are there constituents of exactly 3 elements (these can be derived where necessary by Halle and Vergnaud, of course, as we will see below), and there are never languages where it is necessary to mark off constituents of arbitrary but exact size, ascending prime numbers, for example. So there is a basic choice between pairs or long indefinite strings.

If the constituents are formed in a reflection of alternating rhythm, as in (3), we find that the head will not be separated from either of its constituent boundaries by more than one intervening element (and by zero intervening elements on the other side). Thus, the theory provides the binary parameter +/-BOUNDED, which when set '+' in the well-formedness statement pertaining to a given grid line (say line 0 in (3) above) expresses the restriction that any construction of constituents on this line must have the property that when heads are chosen (and as far as this parameter is concerned the heads could be chosen as any element within each constituent on the given line), those heads (on the next line up) will dominate asterisks on the inferior line which are no more than one asterisk away from a constituent boundary. This situation clearly pertains in (3), and the +BND settings would also be satisfied for (3) had the head been chosen as the rightmost element of each line 0 constituent.

Another parameter that must be set as part of the well-formedness statement for each grid line expresses:

- whether or not the head of the constituent is adjacent to one of the constituent boundaries.

In the binary constituents that we saw in (3), the head must always be adjacent to a boundary. However, if constituents are formed over indefinitely long strings (-BND, non-ternary), then the possibility exists that a head might be arbitrarily far from a constituent boundary. The parameter +/-HEAD-TERMINAL restricts the representations on a given line. If the setting for a given grid line is +HT, then the head of the constituents on this line must be adjacent to a constituent boundary, while if the setting is -HT, the head need not be adjacent to a constituent boundary.

Halle and Vergnaud 1987 provide the following graphical summary of the possibilities for setting these parameters:

| | |
|----------------------|--------------|
| (4) [-BND, +HT] | [+BND, +HT] |
| * (* * * * * *) | * (* *) |
| * (* * * * * *) | * (* *) |

Now varying HT to a minus specification, we can additionally derive:

(5) [-BND, -HT]

[+BND, -HT]

*

 (* * * * * *)

*

 (* * *)

In (5) we see that the heads of constituents are not adjacent to constituent boundaries, reflecting the -HT setting. However, the -BND representation in (5) is ruled out by the Recoverability Condition, which states:

(6) Recoverability Condition

Given the direction of government of the constituent heads in the grammar, the location of the metrical constituent boundaries must be unambiguously recoverable from the location of the heads, and conversely the location of the heads must be recoverable from that of the boundaries.

The phrase 'direction of government' in (6) refers to a further specification in the well-formedness statement of constituent structure for each grid line: whether the head of a +HT constituent is to be chosen as the leftmost asterisk or the rightmost asterisk in the constituent. This parameter would select between the first row in (4), which consists of left-headed constituents, and the second, which consists of right-headed constituents. (6) then states that if we are given this direction parameter ('right' or 'left') for a grid line, and we are given the line with constituent boundaries marked, we must be able to place the heads unambiguously, by reference only to the HT and BND parameter settings. Clearly this cannot be done for the first

representation in (5), and therefore this representation is rule out of the theory by the recoverability condition. In fact, representations such as this are never required for modeling stress or prominence relations in the phonology of any language.

A final parameter in the well-formedness statement for each grid line dictates whether constituents are built starting from the left edge of the line 0 asterisk or the right edge. This will give different possibilities for binary constituents, depending on whether there are an odd or an even number of line 0 asterisks in the word. Unbounded constituents are the same whether boundaries introduced from the right edge or the left.

4. Derivational Procedure

In stating the well-formedness conditions for each grid line, according to the small set of parameters introduced above, we have not said anything about actual derivation, which is done by rule. The rules are very general, since the representations they generate are filtered by the well-formedness parameters introduced above. The rule for boundary placement on a grid line n is thus:

- Construct constituent boundaries $\left\langle \begin{array}{l} \text{left to right} \\ \text{right to left} \end{array} \right\rangle$ on line n .

The rule for locating head can be equally simple:

- Locate the heads of the line n metrical constituents on line $n+1$.

Though these rules are very simple, it is important to distinguish them clearly from the well-formedness statements

which filter their output (or, alternatively, one may understand the rules as actually referring to the parameter settings in the well-formedness statements in their operation). As phonological rules, they may be ordered and interspersed among other phonological rules in a derivation.

5. Sample Derivation

Halle and Vergnaud provide the following sample derivation illustrating the parameter settings:

Consider such simple stress patterns as those of Latvian, where stress is word-initial, and French, where stress is word-final. ... We formalize this as follows.

[7]

- a. Line 0 constituents are [+HT, -BND, left_{Latv}/right_{Fr}].
- b. Construct constituent boundaries on line 0.
- c. Locate the heads of line 0 constituents on line 1.

In [8] we illustrate the effects of [7] graphically.

[8]

- | | |
|--------------------------------|--|
| a. * (* * *) Latvija | b. * (* * * * * *) originalité |
|--------------------------------|--|

(Halle and Vergnaud, p. 12)

It should be remembered that this example is exceedingly simple, and that there are many extremely complex possibilities for derivational systems based on this theory, the reader is referred to Halle and Vergnaud 1987 and Sietsema 1987 for further details.

Bibliography

- Bao, Z.: 1990, On the Nature of Tone. MIT doctoral dissertation.
- Beckman, M.: 1986, Stress and Non-stress Accent. Netherlands Phonetic Archives, Foris Publications.
- Beckman, M. and J. B. Pierrehumbert: 1986, "Intonational Structure in Japanese and English". Phonology Yearbook 3.
- Birdwhistell, R. L.: 1970, Kinesics and Context. University of Pennsylvania Press, Philadelphia.
- Bolinger, Dwight: 1985, "Two views of accent". Journal of Linguistics 21, 79-123.
- Bolinger, Dwight: 1986, Intonation and its parts. Edward Arnold, London.
- Bolinger, Dwight: 1989, Intonation and its uses. Edward Arnold, London.
- Chang, Jun and B. Shefts: 1964, A Manual of Spoken Tibetan University of Washington Press, Seattle.
- Civera, Marco: 1971, The distinctive phonology of Lhasa Central Tibetan. doctoral dissertation, Georgetown University.
- Condon, W.S. and W.D. Ogston: 1967, "A Segmentation of Behavior". Journal of Psychiatric Research 5, 221-235.
- Cruttenden, Alan: 1981, "Falls and rises: meanings and universals". Journal of Linguistics 17, 1-178.
- Dawson, W.: 1980, Tibetan Phonology. University of Washington doctoral dissertation.
- Erteschik-Shir, Nomi, and Shalom Lappin: 1983, "Under stress: a functional explanation of English sentence stress". Journal of Linguistics 19, 419-453.
- Fink, Raymond B. and R. J. Demarest: 1978, Laryngeal Biomechanics. Harvard University Press, Cambridge.
- Fromkin, V. (ed.): 1978, Tone a linguistic survey. Academic Press, New York.

- Goldstein, M.: 1984, English-Tibetan Dictionary of Modern Tibetan. University of California Press.
- Goldstein, M. and N. Nornang: 1984, Modern Spoken Tibetan Lhasa Dialect. Ratna Pustak Bhandar, Kathmandu.
- Guessenhoven, C.: 1983, "Focus, mode, and the nucleus". Journal of Linguistics 19, 377-417.
- Hadar, U.: 1989, "Gestural Modulation of Speech Production: the Role of Head Movement". Language and Communication 9.4.
- Halle, M.: 1990, "Respecting Metrical Structure". Natural Language and Linguistic Theory.
- Halle, M. and R. Vergnaud: 1987, An Essay on Stress. MIT Press, Cambridge.
- Hari, A. M.: 1980, An Investigation of the Tones of Lhasa Tibetan. Summer Institute of Linguistics, Huntington Beach.
- Hirano, M. and K. Stevens: 1981, Vocal Fold Physiology. University of Tokyo Press, Tokyo.
- Hirschberg, J. and J. Pierrehumbert: 1986, "Intonational Structuring of Discourse". Proceedings of the 24th Meeting of the Association for Computational Linguistics, New York. 136-144.
- Ho, A.T.: 1976, "The Acoustic Variation of Mandarin Tones", Phonetica 33.
- Ho, A.T.: 1976, "Mandarin Tones in Relation to Sentence Intonation and Grammatical Structure", Journal of Chinese Linguistics 4.
- Hoa, Monique: 1983, L'Accentuation en Pekinois. University of Paris doctoral dissertation, published by Centre de Recherches Linguistiques, Paris.
- Hockett, C.F.: 1947, "Peiping Phonology". in Readings in Linguistics, Chicago, 1966.
- Hombert, J.: 1978, "Consonant Types, Vowel Quality, and Tone. In Fromkin 1978.
- Kjellin, O.: 1976, "A Phonetic Description of Tibetan with a review of the literature". Research Institute of Logopedics and Phoniatrics 10, University of Tokyo.

- Kratochvil, P.: 1969, "Syllabic Volume as Acoustic Correlate of Perceptual Prominence in Peking Dialect". Unicorn 5.
- Kratochvil, P.: 1974, "Stress Shift Mechanism and its Role in Peking Dialect". Modern Asian Studies, 1974.
- Ladd, R.: 1983, "Phonological features of intonational peaks". Language 59 721-759.
- Lieberman, M.: 1975, The Intonation System of English. MIT doctoral dissertation. Reproduced by the University of Indiana Linguistics Club, Bloomington.
- Lieberman, P.: 1966, Intonation, Perception, and Language MIT Research Monograph 38.
- Oakeshott-Taylor, John: 1984, "Factuality and Intonation". Journal of Linguistics 20, 1-21.
- Ohala, J.: 1978: "Production of Tone". In Fromkin 1978.
- O'Shaughnessy, D.: 1976, Modelling Fundamental Frequency, and its Relationship to Syntax, Semantics, and Phonetics. MIT doctoral dissertation.
- Ossorio, J.: 1982, Tsang Tibetan Phonology. University of Colorado doctoral dissertation.
- Pierrehumbert, J. and D. Talkin: 1989, "Lenition of /h/ and Glottal Stop". Ms. Linguistics Research Department, AT&T Bell Laboratories.
- Pierrehumbert, J.: 1980, The phonology and phonetics of English intonation. doctoral dissertation, MIT, Cambridge, MA.
- Rietveld, A.C.M., and C. Gussenhoven: 1985, "On the relation between pitch excursion size and prominence". Journal of Phonetics 13, 299-308.
- Rochemont, M.: 1986, Focus in Generative Grammar. John Benjamins Publishing Co., Amsterdam.
- Ross, Elliott D. et al: 1986, "The effect of affect on various acoustic measures of prosody in tone and non-tone languages: A comparison based on computer analysis of voice". Journal of Phonetics 14, 283-302.
- Sampson, G.R.: 1974, Is there a universal phonetic alphabet? Language 50, 236-259.
- Schmidt, R.A.: 1975, "A schema theory of discrete motor skill learning". Psychological Review 82, 225-260.

- Schuh, R.G.: 1978, "Tone Rules". In Fromkin 1978.
- Selkirk, E.: 1984, Phonology and Syntax. MIT Press, Cambridge.
- Shih, Chi-lin: 1986, The Prosodic Domain of Tone Sandhi in Chinese. UCSD doctoral dissertation.
- Shih, Chi-lin: 1987, "The Phonetics of the Chinese Tonal System". AT&T Bell Labs Technical Memo.
- Sietsema, B.: 1989, Metrical Dependencies in Tone Assignment. MIT doctoral dissertation.
- Sprigg, R.K.: 1954, "Verbal phrases in Lhasa Tibetan" I-III. Bulletin of the School of Oriental and African Studies XVI.
- Thonden, L.: 1984: Modern Tibetan Language. Library of Tibetan Works and Archives, Dharamsala.
- Umeda, N.: 1982, "F0 declination is situation dependent". Journal of Phonetics 10 279-291.
- Willems, Nico: 1982, English intonation from a Dutch point of view. Netherlands Phonetic Archives, Foris Publications.
- Xu, S.: 1956, "Shuangyin zhuici de zhongyin guilu". Zhongguo Yuwen 2.
- Yamada, Eiji: 1989/1990, "Stress Assignment in Tokyo Japanese". Ms. Fukuoka University/MIT.
- Yip, M.: 1980, The Tonal Phonology of Chinese. MIT doctoral dissertation / Indiana University Linguistics Club, Bloomington, Indiana.
- Yip, M.: 1981, "The interaction of tone and stress in Molinos Mixtec". Ms. Brandeis University.
- Yip, M.: 1982, "Word and Phrase Stress in Mandarin". paper given at the XVth International Conference on Sino-Tibetan Languages and Linguistics, Peking.
- Yip, M.: 1989, comments and discussion on "Ga-Gyoo: Featural and Prosodic Characteristics". (Ito and Mester), MIT Conference on Features and Under-specification.
- Yip, M.: 1989, "Tonal Register in East Asian Languages". to appear in: K. Snider and H. van der Hulst, eds., The Representation of Tonal Register, Foris, Dordrecht.