AT&T Labs - Research

subject: **Stress Assignment in Letter to Sound Rules for Speech Synthesis**
Work Project No. 311402-2099
File Case 38794-54

date: **November 27, 2000**

from: **Kenneth W. Church**

11225-850509-05

Technical Memorandum

1. Background

A speech synthesizer is a machine that inputs a text stream and outputs an accoustic signal. One small piece of this problem will be discussed here: words → phonemes. The resulting phonemes are then mapped into a sequence of lpc dyads which are combined with duration and pitch information to produce speech.

text → intonation phrases → words → phonemes → lpc dyads + prosody → accoustics

There are two general approaches to word → phonemes:

- Dictionary Lookup
- Letter to Sound (i.e., sound the word out from basic principles)

Both approaches have their advantages and disadvantages; the dictionary approach fails for unknown words (e.g., proper nouns) and the letter to sound approach fails when the word doesn't follow the rules, which happens all too often in English. Most speech synthesizers adopt a hybrid strategy, using the dictionary when appropriate and letter to sound for the rest.

Some people have suggested to me that modern speech synthesizers should do away with letter to sound rules now that memory prices are dropping so low that it ought to be practical to put every word of English into a tiny box. Actually memory prices are still a major factor in the cost of a machine. But more seriously, it is not possible to completely do away with letter to sound rules because it is not possible to enumerate all of the words of English. A typical college dictionary of 50,000 headwords will account for about 93% of a typical newspaper text. The bulk of the unknown words are proper nouns.

The difficulty with proper nouns is demonstrated by the table below which compares the Brown Corpus with the surnames in the Kansas City Telephone Book. The table answers the question: how much of each corpus would be covered by a dictionary of n words? Thus the first line shows that a dictionary of 2000 words would cover 68% of the Brown Corpus, and a dictionary of 2000 names would cover only 46% of the Kansas City Telephone Book. It should be clear from the table that a dictionary of surnames must be much larger than a typical college dictionary (˜20,000 entries). Moreover, it would be a lot of work to construct such a dictionary since there are no existing computer readable dictionaries for surnames.

Size of Word Dictionary	Brown Corpus	Size of Name Dictionary	Kansas City
2000	68%	2000	46%
4000	78%	4000	57%
6000	83%	6000	63%
8000	86%	8000	68%
10000	89%	10000	72%
12000	91%	12000	75%
14000	92%	14000	77%
16000	94%	16000	79%
18000	95%	18000	81%
20000	95%	20000	83%
22000	96%	22000	84%
24000	97%	24000	86%
26000	97%	26000	87%
28000	98%	28000	88%
30000	98%	30000	89%
32000	98%	32000	90%
34000	99%	34000	91%
36000	99%	36000	91%
38000	99%	38000	92%
40000	99%	40000	93%

Actually, this table overestimates the effectiveness of the dictionary, for practical applications. A fair test would not use the same corpus for both selecting the words to go into the dictionary and for testing the coverage. The scores reported here were computed *post hoc*, a classic statistical error. I tried a more fair test, where a dictionary of 43777 words (the entire Brown Corpus) was tested against a corpus of 10687 words selected from the AP news wire. The results showed 96% coverage, which is slightly lower (as expected) than the 99% figure reported in the table for a 40000 dictionary.

For names, the facts are much more striking as demonstrated in the following table which tests name lists of various sizes against the Bell Laboratories phone book. (As above, the name lists were gathered from the Kansas City Telephone Book.)*

---

* Admittedly, this test is somewhat unfair to the dictionary approach since the ethnic mixture in Kansas City is very different from that found here at Bell Laboratories.

Size of Word List (Kansas)	Coverage of Test Corpus (Bell Labs)
2000	0.496
4000	0.543
6000	0.562
8000	0.571
10000	0.577
20000	0.589
40000	0.595
50000	0.596
60000	0.596
90000	0.597

Note that the asymptote of 60% coverage is quickly reached after only about 5000-1000 words, suggesting (a) that the dictionary approach may only be suitable for the 5000 to 1000 most frequent names because larger dictionaries yield only negligible improvements in performance, and (b) that the dictionary approach has an inherent limitation on coverage of about 60%. To increase the coverage beyond this, it is probably necessary to apply alternative methods such as letter to sound rules.

Over the past year I have been developing a set of letter to sound rules as part of a larger speech synthesis project currently underway at Murray Hill. Only one small piece of my letter to sound rules, orthography → stress, will be discussed here. The output stress assignment is then used to condition a number of rules such as palatalization in the mapping from letters to phonemes.

2. Weight as an Intermediate Level of Representation

Intuitively, stress dependencies come in two flavors: (a) those that apply locally within a syllable, and (b) those that apply globally between syllables. Syllable weight is an attempt to represent the local stress constraints. Syllables are marked either heavy or light, depending only on the local 'shape' (e.g., vowel length and number of post-vocalic consonants). Heavy syllables are more likely to be stressed than light syllables, though the actual outcome depends upon contextual constraints, such as the English main stress rule, which will be discussed shortly.

The notion of weight is derived from Chomsky and Halle's notion of strong and weak clusters [Chomsky and Halle] (SPE). In phonological theory, weight is used as an intermediate level of representation between the input underlying phonological representation and the output stress assignment. In a similar fashion, I will use weight as an intermediate level of representation between the input orthography and the output stress. The orthography → stress problem will be split into two subproblems:

- Orthography → Weight
- Weight → Stress

3. What is Syllable Weight?

Weight is a binary feature (Heavy or Light) assigned to each syllable. The final syllables of the verbs *obey, maintain, erase, torment, collapse*, and *exhaust* are heavy because they end in a long vowel or two consonants. In constrast, the final syllables of *develop, astonish, edit, consider*, and *promise* are light because they end in a short vowel and at most one consonant. More precisely, to compute the weight of a syllable from the underlying phonological representation, strip off the final consonant and then parse the word into syllables (assigning consonants to the right when there is ambiguity).

Example	Weight	Reason
ow-*bey*	heavy final syllable	long vowel
tor-*men*	heavy final syllable	closed syllable
diy-ve-*lo*	light final syllable	open syllable & short vowel

Then, if the syllable is closed (i.e., ends in a consonant as in tor-*men*) or if the vowel is marked underlyingly long (as in ow-*bey*), the syllable is marked heavy. Otherwise, the syllable ends in an open short vowel and it is marked light. Determining syllable weight from the orthography is considerably more difficult than from the underlying phonological form. I will return to this question shortly.

4. Weight → Stress

Global stress assignment rules apply off the weight representation. For example, the main stress rule of English says that verbs have final stress if the final syllable is heavy syllable (e.g., *obey*), and penultimate stress if the final syllable light syllable (e.g., *develop*). The main stress rule works similarly for nouns, except that the final syllable is ignored (extrametrical [Hayes]). Thus, nouns have penultimate stress if the penultimate syllable is heavy (e.g., a*róm*a) and antipenultimate stress if the penultimate syllable is light (e.g., cin*e*ma).

Examples	Penultimate Weight	Reason
a*róm*a	heavy	long vowel
ver*rán*da	heavy	closed syllable
cĭn*e*ma	light	open syllable & short vowel

Adjectives stress just like verbs except suffixes are ignored (extrametrical). Thus monomorphemic adjectives such as *discréet*, *robúst* and *cómmon* stress just like verbs (the final syllable is stressed if it is heavy and otherwise the penultimate syllable is stress) whereas adjectives with single syllable suffixes such as -al, -ous, -ant, -ent and -ive follow the same pattern as regular nouns [Hayes, p. 242].

Stress Pattern of Suffixed Adjectives		
Light Penultimate	**Heavy Penultimate**	**Heavy Penultimate**
munícipal	adjectíval	fratérnal
magnáminous	desírous	treméndous
signíficant	clairvóyant	relúctant
ínnocent	complácent	depéndent
prímitive	condúcive	expénsive

5. Sproat's Weight Table

A large number of phonological studies (e.g., [Chomsky and Halle], [Liberman and Prince], [Hayes]) outline a deterministic procedure for assigning stress from the weight representation and the number of extrametrical syllables (1 for nouns, 0 for verbs). A version of this procedure was implemented by Richard Sproat last summer.

For efficiency purposes, Sproat's program was compiled into a table, which associated each possible input with the appropriate stress pattern.

Sproat's Weight Table		
Weight	Part of Speech	
	Verb	Noun
H	1	1
L	1	1
HH	31	10
HL	10	10
LH	01	10
LL	10	10
HHH	103	310
HHL	310	310
HLH	103	100
HLL	310	100
LHH	103	010
LHL	010	010
LLH	103	100
LLL	010	100
etc.		

Note that the table is extremely small. Assuming that words have up to N syllables and up to E extrametrical syllables, there are $E\sum_{i=1}^{N} 2^i$ possible inputs. For E = 2 and N = 8, the table has only 1020 entries, which is not unreasonable.

6. Analogy with Waltz' Constraint Propagation Paradigm

Recall that Waltz was the first to showed how contraints could be used effectively in his program that analyzed line drawings in order to separate the figure from the ground and to distinguish concave edges from convex ones. He first assigned each line a convex label (+), a concave label (−) or a boundary label (<, >), using only local information. If the local information was ambiguous, he would assign a line two or more labels. Waltz then took advantage of the constraints imposed where multiple lines come together at a common vertex. One would think that there ought to be 4^2 ways to label a vertex of two lines and 4^2 ways to label a vertex of three lines and so on. By this argument, there ought to be 208 ways to label a vertex. But Waltz noted that there were only 18 vetex labelings that were consistent with certain reasonable assumptions about the physical world. Because the inventory of possible labelings was so small, he could disambiguate lines with multiple assignments by checking the junctures at each end of the line to see which of the assignments were consistent with one of the 18 possible junctures. This simple test turned out to be extremely powerful.

Sproat's weight table is very analogous with Waltz' list of vertex constraints; both define an inventory of global contextual constraints on a set of local labels (H and L syllables in this application, and +, −, >, < in Waltz application). Waltz' constraint propagation paradigm depends on a highly constrained inventory of junctures. Recall that only 18 of 208 possible junctures turned out to be grammatical. Similarly, in this application there are very strong grammatical constraints. According to Sproat's table, there are only 51 distinct output stress assignments, a very small number considering that there are 1020 distinct inputs.

Possible Stress Assignments					
1	103	3103	020100	0202013	20020100
3	310	02010	020103	2002010	20020103
01	313	02013	200100	2002013	20202010
31	0100	20010	200103	2020100	20202013
10	0103	20013	202010	2020103	32020100
13	2001	20100	202013	3202010	32020103
010	2010	20103	320100	3202013	
013	2013	32010	320103	02020100	
100	3100	32013	0202010	02020103	

The strength of these constraints will help make up for the fact that the mapping from orthography to weight is usually underdetermined. In terms of information theory, about half of the bits in the weight representation are redundant since log 51 is about half of log 1020. This means that I only have to determine the weight for about half of the syllables in a word in order to assign stress.

The redundancy of the weight representation can also been seen directly from Sproat's weight table as shown below. For a one syllable noun, the weight is irrelevant. For a two syllable noun, the weight of the penultimate is irrelevant. For a three syllable noun, the weight of the antipenultimate syllable is irrelevant if the penultimate is light. For a four syllable noun, the weight of the antipenultimate is irrelevant if the penultimate is light and the weight of the initial two syllables are irrelevant if the penultimate is heavy. These redundancies follow, of course, from general phonological principles of stress assignment.

Weight by Stress (for short Nouns)				
Stress	Weight			
1	L	H		
10	LL	HL		
13	LH	HH		
010	LHL			
310	HHL			
013	LHH			
313	HHH			
100	HLL	LLL		
103	LLH	HLH		
0100	LHLL	LLLL		
3100	HHLL	HLLL		
0103	LLLH	LHLH		
3103	HLLH	HHLH		
2010	LLHL	HHHL	LHHL	HLHL
2013	LHHH	HLHH	LLHH	HHHH

7. Orthography → Weight

For practical purposes, Sproat's table offers a complete solution to the weight → stress subtask. All that remains to be solved is: orthography → weight. Unfortunately, this problem is much more difficult and much less well understood. I'll start by discussing some easy cases, and then introduce the pseudo-weight heuristic which helps in some of the more difficult cases. Fortunately, I don't need a complete solution to orthography → weight since weight → stress is so well constrained.

In easy cases, it is possible to determine the weight directly for the orthography. For example, the weight of *torment* must be "HH" because both syllables are closed (even after stripping off the final consonant). Thus, the stress of *torment* is either "31" or "13" stress depending on whether is has 0 or 1 extrametrical final syllables:*

(stress-from-weights "HH" 0) → ("31") ; verb
(stress-from-weights "HH" 1) → ("13") ; noun

However, most cases are not this easy. Consider a word like *record* where the first syllable might be light if the first vowel is reduced or it might be heavy if the vowel is underlyingly long or if the first syllable includes the /k/. It seems like it is impossible to say anything in a case like this. The weight, it appears is either "LH" or "HH". Even with this ambiguity, there are only three distinct stress assignments: 01, 31, and

* Actually, in practice, the weight determination is complicated by the possibility that *-ment* and *-ent* might be suffixes. Note, for example, that the adjective *dórmant* does not stress like the verb *torménт* because the adjectival suffix *-ant* is extrametrical.

13.

(stress-from-weights "LH" 0) → ("01")
(stress-from-weights "HH" 0) → ("31")
(stress-from-weights "LH" 1) → ("13")
(stress-from-weights "HH" 1) → ("13")

8. Pseduo-Weight

In fact, it is possible now to use the stress to further constrain the weight. Note that if the first syllable of *record* is light it must also be unstressed and if it is heavy it also must be stressed. Thus, the third line above is inconsistent.

I implement this additional constraint by assigning *record* a pseudo-weight of "=H", where the "=" sign indicates that the weight assigment is constrained to be the same as the stress assigment (either heavy & stressed or not heavy & not stressed). I can now determine the possible stress assignments of the pseudo-weight "=H" by filling in the "=" constraint with all possible bindings (H or L) and testing the results to make sure the constraint is met.

(stress-from-weights "LH" 0) → ("01")
(stress-from-weights "HH" 0) → ("31")
(stress-from-weights "LH" 1) → ("13") ; No Good
(stress-from-weights "HH" 1) → ("13")

Of the four logical inputs, the = constraint excludes the third case which would assign the first syllable a stress but not a heavy weight. Thus, there are only three possible input/output relations meeting all of the constraints:*

Weight	Extrametrical Syllables	Stress
LH	0 (verb)	01
HH	0 (verb)	31
HH	1 (noun)	13

All three of these possibilities are grammatical.

The following pseudo-weights are defined:

---

* The noun should probably have the stress 10 rather than the stress 13. I assume that an extrametrical syllable has 3 stress if it is heavy, and 0 stress if it is light. The stress of the extrametrical syllable is very difficult to predict, as discussed in [Ross].

Label	Title	Constraints
H	Heavy	weight = H; stress is unknown
L	Light	weight = L; stress is unknown
=	Unknown	(weight = H) ≡ (stress = 0)
S	Superheavy	weight = H; stress ≠ 0
R	Superlight	weight = L; stress = 0
N	Sonorant	(weight = H) ≡ (stress = 0)
?	Truly Unknown	weight is unknown; stress is unknown

I have already given examples of the labels H, L and =. S and R are used in certain morphological analyses (see below), N is used for examples where Hayes would invoke his rule of Sonorant Destressing (see below), and ? is not used except for demonstrating the program.

The procedure that assigns pseudo-weight to orthography is roughly as outlined below, ignoring morphology, etymological and more special cases than I wish to admit.

1. Tokenize the orthography so that digraphs such as *th, gh, wh, ae, ai, ei*, etc., are single units.
2. Parse the string of tokens into syllables (assigning consonants to the right when the location of the syllable boundary is ambiguous).
3. Strip off the final consonant.
4. For each syllable
    a. Silent *e*, Vocalic *y* and Syllabic Sonorants (e.g., *-le, -er, -re*) are assigned no weight.
    b. Digraphs that are usually realized as long vowels (e.g., *oi*) are marked H.
    c. Syllables ending with sonorant consonants are marked N; other closed syllables are marked H.
    d. Open syllables are marked =.

In practice, I have observed that there are remarkably few stress assignments meeting all of the constraints. After analyzing over 20,000 words, there were no more than 4 possible stress assigments for any particular combination of pseudo-weight and number of extrametrical number of syllables. Most observed combinations had a unique stress assignment, and the average (by observed combination with no frequency normalization) has 1.5 solutions. In short, the constraints are extremely powerful; words like *record* with multiple stress patterns are the exception rather than the rule.

9. Ordering Multiple Solutions

Generally, when there are multiple stress assignments, one of the possible stress assigments is much more plausible than the others. For instance, nouns with the pseudo-weight of "H=L" (e.g., *difference*) have a strong tendency toward antipenultimate stress, even though they could have either 100 or 310 stress depending on the weight of the penultimate. The program takes advantage of this fact by returning a sorted list of solutions, all of which meet the constraints, but the solutions toward the front of the list are deemed

more plausible than the solutions toward the rear of the list.

(stress-from-weights "H=L" 1) → ("100" "310")

Sorting the solution space in this way could be thought of as a kind of default reasoning mechanism. That is, the ordering criterion, in effect, assigns the penultimate syllable a default weight of L, *unless* there is positive evidence to the contrary. Of course, this sorting technique is not as general as an arbitrary default reasoner, but it seems to be general enough for the application. This limited defaulting mechanism is extremely efficient when there are only a few solutions meeting the constraints.

This default mechanism is also used to stress the following nouns

Hottentot	Jackendoff	balderdash
ampersand	Hackensack	Arkansas
Algernon	mackintosh	davenport
merchandise	cavalcade	palindrome
nightingale	Appelbaum	Aberdeen
misanthrope		

where the penultimate syllable ends with a sonorant consonant (n, r, l). According to what has been said so far, these sonorant syllables are closed and so the penultimate syllable should be heavy and should therefore be stressed. Of course, these nouns all have antipenultimate stress, so the rules need to be modified. Hayes suggested a Sonorant Destressing rule which produced the desired results by erasing the foot structure (destressing) over the penultimate syllable so that later rules will reanalyze the syllable as unstressed. I propose instead to assign these sonorant syllables the pseudo-weight of N which is essentially identical to =.* In this way, all of these words will have the pseudo-weight of HNH which is most likely stressed as 103 (the correct answer) even though 313 also meets the constraints, but fair worse on the ordering criteron.

(stress-from-weights "HNH" 1) → ("103" "313")

Contrast the examples above with *Adirondack* where the stress does not back up past the sonorant syllable. The ordering criterion is adjusted to produce the desired results in this case, by assuming that two binary feet (i.e., 2010 stress) are more plausible than one tertiary foot (i.e., 0100 stress).

(weights-from-orthography "Adirondack") → "L=NH"
(stress-from-weights "L=NH") → ("2013" "0103")

It ought to be possible to adjust the ordering criterion in this way to produce (essentially) the same results as Hayes' rules.

---

* N and = used to be identical. I am still not sure the differences are justified. At any rate, the differences are very subtle and certainly not worth going into here.

10. Morphology

Thus far, the discusion has assumed monomorphemic input. Morphological affixes add yet another rich set of constraints. Recall the examples mentioned in the abstract, *degráde/dègradátion* and *télegraph/telégraphy*, which were used to illustrate that stress alternations are conditioned by morphology. This section will discuss how this is handled in the program. The task is divided into two questions: (1) how to parse the word into morphemes, and (2) how to integrate the morphological parse into the rest of stress assignment procedure discussed above.

The morphological parser uses a grammar roughly of the form:

word → level3 (regular-inflection)*
level3 → (level3-prefix)* level2 (level3-suffix)*
level2 → (level2-prefix)* level1 (level2-suffix)*
level1 → (level1-prefix)* (syl)* (level1-suffix)*

where latinate affixes such as *in+, ir+, ac+, +ity, +ion, +ive, -al* are found at level 1, Greek and Germanic affixes such as *hetero#, un#, under#, #ness, #ly* are found at level 2, and compounding is found at level 3. The term *level* refers to Mohanan's theory of Level Ordered Morphology and Phonology [Mohanan] which builds upon a number of well-known differences between + boundary affixes (level 1) and # boundary affixes (level 2).

- Distributional Evidence: It is common to find a level 1 affix inside the scope of a level 2 affix (e.g., *un#in+terned* and *form+al#ly*), but not the other way around (e.g., **in+un#terned* and **form#ly+al*).
- Wordness: Level 2 affixes attach to words, whereas level 1 affixes may attach to fragments. Thus, for example, *in+* and *+al* can attach to fragments as in *intern* and *criminal* in ways that level 2 cannot **un#tern* and **crimin#ness*.
- Stress Alternations: Stress alternations are found at level 1 *párent* → *parént+al* but not at level 2 as demonstrated by *párent#hood*. Level 2 suffixes are called *stress neutral* because they do not move stress.
- Level 1 Phonological Rules: Quite a number of phonological rules apply at level 1 but not at level 2. For instance, the so-called tri-syllabic laxing rule will lax a vowel before a level 1 suffix (e.g., *divine* → *divin+ity*) but not before a level 2 suffix (e.g., *divine#ly* and *divine#ness*). Similarly, the rule that maps /t/ into /s/ in *president* → *presidency* also fails to apply before a level 2 affix: *president#hood* (not **presidence#hood*).

Given evidence such as this, there can be little doubt on the necessity of the level ordering distinction. Level 2 affixes are fairly easy to implement; the parser simply strips off the stress neutral affixes, assigns stress to the parts and then pastes the results back together. For instance, *parenthood* is parsed into *parent* and *#hood*. The pieces are assigned 10 and 3 stress respectively, producing 103 stress when the pieces are recombined. In general, the parsing of level 2 affixes is not very difficult, though there are some cases where it is very difficult to distinguish between a level 1 and level 2 affix. For example, *-able* is level 2 in *changeable* (because of silent e which is not found before level 1 suffixes), but level 1 in *cómparable* (because of the stress shift from *compáre* which is not found before level 2 suffixes). For dealing with a

limited number of affixes like *-able* and *-ment*, there are a number of special purpose diagnostic procedures which decide the appropriate level.

Level 1 suffixes have to be stressed differently. In the lexicon, each level 1 suffix is marked with a weight. Thus, for example, the suffix *+ity* is marked RR. These weights are assigned to the last two syllables, regularless of what would normally be computed. Thus, the word *civil+ity* is assigned the pseudo-weight ==RR which is then assigned the correct stress by the usual methods:

(stress-from-weights "==RR" 1) → ("0100" "3100")

The fact that *+ity* is marked for weight in this way makes it relatively easy for the program to determine the location of the primary stress. Shown below are some sample results of the program's ability to assign primary stress.*

---

* Strings such as *ization, aceous, logist* are really sequences of several affixes. In order to avoid some difficult parsing issues, I decided not to allow more than one level 1 suffix per word. This limitiation requires that I enter sequences of level 1 suffixes into the lexicon.

% Correct Primary Stress	Number of Words Tested	Level 1 Suffix
0.98	726	+ity
0.98	1652	+ion
0.97	345	+ium
0.97	136	+ular
0.97	339	+ical
0.97	236	+eous
0.97	33	+ization
0.98	160	+aceous
0.97	215	+ious
0.96	151	+osis
0.96	2617	+ic
0.96	364	+ial
0.96	169	+meter
0.95	617	+ian
0.95	122	+ify
0.94	17	+bly
0.94	17	+logist
0.94	313	+ish
0.93	56	+istic
0.92	2626	+on
0.92	24	+ionary
0.90	19	+icize
0.88	52	+ency
0.82	1818	+al
0.77	128	+atory
0.77	529	+able

These selected results are biased slightly in favor of the program. Over all, the program correctly assigns primary stress to 82% of the words in the dictionary, and 85% for words ending with a level 1 affix.

Prefixes are more difficult than suffixes. Examples such as *supér+fluous* (level1 1), *súper#conductor* (level 2), and *súper##market* (level 3) illustrate just how difficult it is to assign the prefix to the correct level. Even with the correct parse, it not a simple matter to assign stress. In general, level 2 prefixes are stressed like compounds, assigning primary stress to the left morpheme (e.g., *úndercarriage*) for nouns and to the right for verbs (e.g., *undergó*) and adjectives (e.g., *ùltraconsérvative*), though there seem to be two classes of exceptions. First, in technical terms, under certain conditions [Hayes, pp. 307-309], primary stress can back up onto the prefix: (e.g., *telégraphy*). Secondly, certain level 1 suffixes such as *+ity* seem to induce a remarkable stress shift (e.g., *súper#conductor* and *sùper#conductívity*), in violation of level ordering as far as I can see.

For level 1 suffixes, the program assumes the prefixes are marked light and that they are extrametrical in verbs, but not in nouns. Prefix extrametricality accounts for the well-known alternation *pérmit* (noun) versus *permít* (verb). Both have L= weight (recall the prefix is L), but the noun has initial stress since the final syllable is extrametrical whereas the verb has final stress since the initial syllable is extrametrical. Extrametricality is required here, because otherwise both the noun and verb would receive initial stress.

11. Etymology

The stress rules outlined above work very well for the bulk of the language, but they do have difficulties with certain loan words. For instance, consider the Italian word *tortóni*. By the reasoning outlined above, *tortóni* ought to stress like *cálculi* since both words have the same part of speech and the same syllable weights, but obviously, it doesn't. In fact, almost all Italian loan words have penultimate stress, as illustrated by the Italian surnames: *Aldrighétti, Angelétti, Bellótti, Iannúcci, Italiáno, Lombardíno, Marciáno, Marcóni, Morillo, Olivétti*. It is clear from examples such as these that the stress of Italian loans is not dependent upon the weight of the penultimate syllable, unlike the stress of native English words. Japanese loan words are perhaps even more striking in this respect. They too have a very strong tendency toward penultimate stress when (mis)pronounced by English speakers: *Asahára, Enomóto, Fujimáki, Fujimóto, Fujimúra, Funasáka, Toyóta, Uméda*. One might expect that a loan word would be stressed using either the rules of the the language that it was borrowed from or the rules of the language that it was borrowed into. But neither the rules of Japanese nor the rules of English can account for the penultimate stress in Japanese loans.

I believe that speakers of English adopt what I like to call a *pseudo-foreign accent*. That is, when speakers want to communciate that a word is non-native, they modify certain parameters of the English stress rules in simple ways that produce bizarre ''foreign sounding'' outputs. Thus, if an English speaker wants to indicate that a word is Japanese, he might adopt a pseudo-Japanese accent that marks all syllables heavy regardless of their shape. Thus, *Fujimúra*, on this account, would be assigned penultimate stress because it is a noun and the penultimate syllable is heavy. Of course there are numerous alternative pseudo-Japanese accents that also produce the observed penultimate stress. The current version of the program assumes that Japanese loans have light syllables and no extrametricality. At the present time, I have no arguments for deciding between these two alternative pseudo-Japanese accents.

The pseudo-accent approach presupposes that there is a method for distinguishing native from non-native words, and for identifying the etymological distinctions required for selecting the appropriate pseudo-accent. Ideally, this decision would make use of a number of phonotactic and morphological cues, such as the fact that Japanese has extremely restricted inventory of syllables and that Germanic makes heavy use of morphemes such as *-berg, wein-* and *-stein*. Unfortunately, because I haven't had the time to develop the right model, the relevant etymological distinctions are currently decided by a statistical tri-gram model. Using a number of training sets (gathered from the telephone book, computer readable dictionaries, bibliographies, and so forth), one for each etymological distinction, I estimated a probability $P(xyz|e)$ that each three letter sequence xyz is associated with etymology e. Then, when the program sees a new word w, a straightforward Baysian argument is applied in order to estimate for each etymology a probability $P(e|w)$ based on the three letter sequences in w.

I have only just begun to collect training sets, but already the results appear promising. Probability estimates are shown in the figure below for some common names whose etymology most readers probably know. The current set of etymologies are: Old French (OF), Old English (OE), International Scientific Vocabulary (ISV), Middle French (MF), Middle English (ME), miscellaneous English (core), Latin (L), Gaelic (NBrit), French (Fr), Swedish (Swed), Russian (Rus), Japanese (Jap), Germanic (Ger), and Southern Romance (SRom). Only the top two candidates are shown and only if the probability estimate is 0.05 or better.

As is to be expected, the model is relatively good at fitting the training data. For example, the following names selected from the training data where run through the model and assigned the label Jap with probability 1.00: *Fujimaki, Fujimoto, Fujimura, Fujino, Fujioka, Fujisaki, Fujita, Fujiwara, Fukada, Fukai, Fukanaga, Fukano, Fukase, Fukuchi, Fukuda, Fukuhara, Fukui, Fukuoka, Fukushima, Fukutake, Funakubo, Funasaka*. Of 1238 names on the Japanese training list, only 48 are incorrectly identified by the model: *Abe, Amemiya, Ando, Aya, Baba, Banno, Chino, Denda, Doke, Gamo, Hase, Huke, Ide, Ise, Kume, Kuze, Mano, Maruko, Marumo, Masuko, Mine, Musha, Mutai, Nose, Onoe, Ooe, Osa, Ose, Rai, Sano, Sone, Tabe, Tako, Tarucha, Uo, Utena, Wada* and *Yawata*. As these exceptions demonstrate, the model has relatively more difficulty with short names, for the obvious reason that short names have fewer tri-grams to base the decision on. Perhaps short names should be dealt with in some other way (e.g., an exception dictionary).

I expect the model to improve as the training sets are enlarged. It might be possible to train the model on a very large number of names, so that there is a relatively small probability that the program will be asked to estimate the etymology of a name that was not in one of the training sets. If, for example, the training sets included the 10000 most frequent names, then most of the names the program would be asked about would probably be in one the training sets (assuming that the results reported above for the telephone directories also apply here).

Before concluding, I would like to point out that etymology is not just used for stress assignment. Note, for instance, that orthographic ch and gh are hard in Italian loans *Macchi* and *spaghetti*, in constrast to the general pattern where ch is /ch/ and /gh/ is silent. In general, velar softening seems to be conditionalized by etymology. Thus, for example, /g/ is usually soft before /I/ (as in *ginger*) but not in *girl* and *Gibson* and many other Germanic words. Similarly, other phonological rules (especially vowel shift) seem to be conditionalized by etymology. I hope to include these topics in a longer version of this paper to be written this summer.

12. Concluding Remarks

Stress assignment was formulated in terms of Waltz' constraint propagation paradigm, where syllable weight played the role of Waltz' labels and Sproat's weight table played the role of Waltz' vertex constraints. It was argued that this formalism provided a clean computational framework for dealing with the following four linguistic issues:

- **Syllable Weight**: *obéy / devélop*

Etymology from Orthography				
Acosta	0.96	SRom		
Alvarado	0.92	SRom	0.08	L
Alvarez	1.00	SRom		
Andersen	0.95	Swed		
Beauchamp	0.47	MF	0.45	OF
Bernstein	1.00	Ger		
Calhoun	1.00	NBrit		
Callahan	1.00	NBrit		
Camacho	0.89	SRom		
Camero	0.77	SRom	0.18	L
Campbell	1.00	NBrit		
Castello	1.00	SRom		
Castillo	1.00	SRom		
Castro	0.73	SRom	0.17	L
Cavanaugh	1.00	NBrit		
Chamberlain	0.86	OF	0.13	MF
Chambers	0.37	Core	0.31	MF
Champion	0.73	OF	0.20	MF
Chandler	0.41	OF	0.25	ME
Chavez	1.00	SRom		
Christensen	0.74	Swed	0.15	Ger
Christian	0.63	Core	0.25	Swed
Christiansen	0.81	Swed	0.10	Core
Churchill	0.62	OE	0.17	Core
Faust	0.40	Ger	0.38	OF
Feliciano	1.00	SRom		
Fernandez	1.00	SRom		
Ferrara	0.79	SRom	0.17	L
Ferrell	0.73	SRom	0.08	ME
Flaherty	1.00	NBrit		
Flanagan	0.97	NBrit		
Fuchs	1.00	Ger		
Gallagher	0.67	NBrit	0.33	SRom
Gallo	1.00	SRom		
Galloway	0.65	OE	0.19	ME
Garcia	0.95	SRom		

- **Part of Speech**: *tórment* (n) / *torménт* (v)
- **Morphology**: *degráde / dègradátion*

- **Etymology**: *cálculi / tortóni*

Currently, the program correctly assigns primary stress to 82% of the words in the dictionary.

References

Chomsky, N., and Halle, M., *The Sound Pattern of English*, Harper and Row, 1968.

Hayes, B. P., *A Metrical Theory of Stress Rules*, unpublished Ph.D. thesis, MIT, Cambridge, MA., 1980.

Liberman, M., and Prince, A., *On Stress and Linguistic Rhythm*, Linguistic Inquiry 8, pp. 249-336, 1977.

Mohanan, K., *Lexical Phonology*, MIT Doctoral Dissertation, available for the Indiana University Linguistics Club, 1982.

Ross, J., *A Reanalysis of English Word Stress (Part I)*, in Brame, M., (ed), *Contributions to Generative Phonology*, University of Texas Press, Austin, Texas, 1972.

Waltz, D., *Understanding Line Drawings of Scences with Shadows*, in P. Winston (ed.) *The Psychology of Computer Vision*, McGraw-Hill, NY, 1975.

Abstract

This paper will discuss how to determine word stress from spelling. Stress assignment is a well-established weak point for many speech synthesizers because stress dependencies cannot be determined locally. It is impossible to determine the stress of a word by looking through a five or six character window, as many speech synthesizers do. Well-known examples such as *degráde / dègradátion* and *télegraph / telégraphy* demonstrate that stress dependencies can span over two and three syllables. This paper will present a principled framework for dealing with these long distance dependencies. Stress assignment will be formulated in terms of Waltz' style constraint propagation with four sources of constraints: (1) syllable weight, (2) part of speech, (3) morphology and (4) etymology. Syllable weight is perhaps the most interesting, and will be the main focus of this paper. Most of what follows has been implemented.