ASSIGNMENT OF SYLLABLE STRESS IN A DEMISYLLABLE-BASED TEXT-TO-SPEECH SYNTHESIS SYSTEM

Suzanne C. Urbanczyk and Stephen J. Eady

Department of Linguistics and
Centre for Speech Technology Research
University of Victoria

The purpose of this paper is to describe an algorithm that has been developed to determine the stress pattern of polysyllabic words in English. Stress is one determinant of word pronunciation in English, and the algorithm is an important component of a demisyllable-based text-to-speech synthesis system. The algorithm described here is a computationally efficient and accurate system for assigning primary stress to English words.

1. INTRODUCTION

The conversion of written English text into intelligible speech by means of a computer involves a number of different steps. The first of these is the conversion of text to some abstract linguistic form that represents the pronunciation of each word in the text. Within this stage there are two components that rely on each other in obtaining a high degree of accuracy for later stages in the synthesis procedure. The first of these is the conversion of English letters into phonemes. The second is the determination of the stress pattern for polysyllabic words. This paper describes an algorithm that has been developed to predict English stress. The algorithm is currently being used in a demisyllable-based text-to-speech synthesis system (Eady *et. al.*, 1988).

2. SRESS IN ENGLISH

The correct assignment of stress to words in English is an important aspect of any text-to-speech synthesis system. Stress in English can be defined as the relative prominence of the syllables in a word, and it has an important influence on pronunciation. For example, it is a difference in stress that causes the difference in pronunciation between the noun "*OBject*" (with primary stress located on the first syllable) and the verb "*obJECT*" (with primary stress located on the second syllable). This difference in stress is manifested by variations in the pitch, duration and amplitude of each syllable in a word. There are varying degrees of stress in English words and the main concern of this paper is to determine the placement of the primary or main stress in words (i.e., the most prominent syllable).

3. AN ALGORITHM FOR STRESS ASSIGNMENT

While English is notorious for its complicated stress patterns, there are certain features of the language that can be exploited to help predict stress placement in a word. These features can be used in an algorithm for predicting the location of primary stress. In particular, it has been noted that the morphological and phonological structure of an

English word can be used to predict the location of the primary-stressed syllable. For example, it is known that certain English suffixes influence stress. Thus, the suffix -eer (as in the word *enginEER*) attracts primary stress to itself. On the other hand, words containing the suffix -ical (as in *CHEMical*) cause the primary stressed syllable to immediately precede the suffix.

Previous work on this problem has shown that a simple rule for the assignment of stress to English words can operate as accurately as other more complex algorithms (Bernstein and Nessly, 1981). However, there are some problems with this approach, because it does not account for the variations in stress patterns that are predictable from the morphological composition of words in English. In particular, it has been noted that the most effective way of predicting the primary stress of a word is accomplished by incorporating a basic stress rule with some level of morphological decomposition (Klatt, 1987).

Following the analysis by Fudge (1985), on the effect of affixes (i.e., prefixes and suffixes) on stress patterns in English, an algorithm has been developed that will locate the primary stressed syllable. The algorithm has two components: one makes use of a basic stress rule, the other involves the analysis of words into their constituent morphemes. These components will be discussed in detail in the following sections. This in turn will be followed by a description of how the two components interact in the design of the algorithm.

3.1 Basic Stress Rule

Much work has been done to formally capture the systematic variations in the location of main stress for English words (Chomsky and Halle, 1968; Liberman and Prince, 1977; Selkirk, 1984). This previous work has pointed to the concept of "syllable weight" as being important for stress determination in a word. Syllable weight refers to the phonological structure of each syllable. A "heavy" syllable is one that ends in a consonant cluster (i.e., more than one consonant). A "light" syllable is one that ends with a single consonant. This concept is important for stress placement, because it is usually a heavy syllable that carries primary stress. Light syllables are usually unstressed.

Another important observation about basic stress patterns in English is that primary stress rarely occurs on the last syllable of a word. It usually occurs on the penultimate (second to last) or on the antepenultimate (third to last) syllable. The choice of which of these syllables receives primary stress depends on the concept of syllable weight. In particular, stress placement depends on the weight of the penultimate syllable. The following examples illustrate this point.

1. verANDa	2. PAMela
AlASKa	CINema
SamANTHa	aPOCalypse

Note that in 1., the main stress falls on the penultimate syllable, while in 2., the main stress falls on the antepenultimate syllable. This difference in stress assignment is due to a difference in syllable weight for the penultimate syllable of each word. The penultimate syllables in the words of list 1. (i.e., and, ask, anth) are heavy and are thus stressed. For the words in 2., on the other hand, the penultimate syllables, (i.e., el, em, al) are light and are not stressed.

The basic stress rule developed in this algorithm makes use of these generalizations in determining the location of the main stress in each word. The rule is stated as follows:

1. locate the penultimate syllable of the word.
2. if it is heavy, then stress it.
3. otherwise stress the antepenultimate syllable.

This stress rule is capable of assigning stress accurately in conditions where no stress-affecting affixes are located. It is easily stated and easily encoded into an algorithm.

3.2 Effect of Affixes on Stress Placement

Fudge (1985) has produced a comprehensive analysis of the effect of affixes on the stress patterns of words in English. How an affix determines the location of primary stress is known as its *accentual property*. In his analysis of English stress patterns, Fudge has proposed a number of categories capturing these accentual properties. There are four different categories, including both suffixes and prefixes. The categories are described below, and examples of each type are shown in Table 1.

3.2.1 Autostressed Suffixes

As in the example given above, the suffix -eer attracts stress to itself. Suffixes of this type are referred to as *autostressed* suffixes. Autostressed suffixes may be one or more syllables in length. Other examples of autostressed suffixes are listed in Table 1.

3.2.2 Pre-stressed Suffixes

The largest class of stress-affecting suffixes are those that cause stress to be located on a syllable preceding the suffix. These *pre-stressed* suffixes can be subdivided into groups, based on which preceding syllable receives the stress. There are three types of *pre-stressed* suffixes: *pre-stressed 1*, *pre-stressed 2*, and *pre-stressed 1/2*. Examples of each are given in Table 1.

Pre-stressed 1 suffixes are those that stress the syllable immediately preceding the suffix. An example of this is the suffix -ic, which predicts the stress in the words *matheMATic*, *sporADic* and *humanISTic*.

Pre-stressed 2 suffixes cause stress to be placed two syllables prior to the suffix. An example of this is the suffix -tude, which correctly assigns stress to words such as *ATtitude* and *SOLitude*.

The third type of *pre-stressed* suffix is the *pre-stressed 1/2* suffix which locates the stress on either the syllable immediately preceding the suffix or two syllables prior to the suffix. The decision as to which of the two syllables before the suffix will receive stress is based on the notion of syllable weight (as described above). In particular, if the syllable immediately preceding the pre-stressed 1/2 suffix is heavy, then it will carry primary stress. If it is light, then stress will be assigned two syllables preceding the suffix. An example of this type of suffix is -al as in the words *uniVERSal* and *orIGinal*. In the first example, the syllable preceding the suffix (i.e., VERS) is heavy and receives stress. In the second example, the syllable in question (i.e., in) is light, and so stress is assigned to the preceding syllable (i.e., RIG). Other examples of pre-stressed 1/2 suffixes are shown in Table 1.

Table 1

Accentual Properties of English Affixes.

Affix Type	**Affix**	**Example**
Suffixes		
Autostressed	-ette	cigarETTE
	-ation	imaginATion
	-self	himSELF
Pre-stressed 1	-ic	sporADic
	-ssion	perMISSion
	-metry	geOMetry
Pre-stressed 2	-ene	acETylene
	-gon	PARagon
	-tude	SOLitude
Pre-stressed 1/2	-al	uniVERSal
		oRIGinal
	-is	syNOPsis
		GENesis
Stress-Neutral	-less	BOTtomless
	-ish	YELlowish
	-dom	MARtyrdom
Prefixes		
Stress-Repellent	ex-	exERT
	ac-	acCOUNT
	af-	afFECT

3.2.3 Stress-neutral Suffixes

The last group of suffixes are referred to as *stress-neutral*. While *stress-neutral* suffixes do not directly influence the placement of primary stress, they do play a role in the stress location. *Stress-neutral* suffixes do not have an accentual property associated with them. This can be seen in how they can attach to words without affecting the stress pattern. For example, when the suffix -ment is attached to the word *GOVern* to create the form *GOVernment*, there is no change in the stress pattern. However, if the suffix -ment is not located prior to the determination of stress, when the basic stress rule is applied it will incorrectly assign main stress on the penultimate syllable as follows: *govERNment*. This is because the penultimate syllable (ERN) is considered heavy and will be assigned stress according to the basic stress rule. Therefore it is important to strip off all *stress-neutral* suffixes prior to the application of the basic stress rule. In this way, *stress-neutral* suffixes are important in the development of a more accurate system of stress assignment.

3.2.4 Prefixes

Prefixes are important in the determination of stress in that there are some with the accentual property of repelling stress. These, appropriately, are referred to as stress-repellent prefixes. An example of this is the prefix ex- in the words *exPEL*, *exCITE* and *exERT*. In these words, we would normally expect to have stress on the first syllable (in accord with the basic stress rule). However, the presence of the stress-repellent prefix causes stress to be placed on the second syllable in each case. Note that these prefixes have a bearing on stress placement, only if there are no stress-affecting suffixes attached to the word in question.

3.3 Design of the Algorithm

The algorithm developed for assignment of primary stress is based on the basic stress rule and the stress-affecting affixes that were just described. As mentioned previously, the stress-assignment algorithm is just one component in the conversion of text to speech. Thus, its development has been influenced by other components. In particular, it is closely linked to the letter-to-sound rules which determine the phonetic pronunciation of each syllable. In fact, the stress-assignment algorithm works simultaneously with the letter-to-sound rules to determine the pronunciation of a word. The input to these two components is English orthographic text. The output is a phonetic transcription of each word along with assignment of primary stress on a particular syllable.

The stress-assignment component works as follows. When a word is encountered, the program attempts to detect the presence of suffixes. If there is a suffix, one of two things occurs. If the suffix is *stress-neutral*, then it is necessary to check for further suffixes. This continues until either a stress-affecting suffix is located, or until no more suffixes are found. The algorithm then does one of two things to determine the stress. If the suffix is not *stress-neutral*, then stress is located based on the *accentual properties* of the suffix. In the event that no stress-affecting suffixes are located, then the program looks for stress-affecting prefixes. If no suffixes or prefixes are found, or if only *stress-neutral* suffixes are found, then the location of the primary stressed syllable is determined by the basic stress rule (described above).

3.4 Sample Derivations

To clarify how the algorithm works some sample derivations will be given that illustrate some of the various components of the algorithm. The first example is the word ALASKA. When the algorithm checks for the presence of suffixes or prefixes it finds none. Therefore, stress must be assigned by the basic stress rule. This operates by examining the penultimate syllable to see if it ends in a consonant cluster. Because the syllable is ASK, and it ends in a consonant cluster, it correctly receives the main stress.

The second example is the word HIMSELF. When the algorithm checks for the presence of suffixes or prefixes it finds that the word ends in the suffix -SELF. This is noted to be an autostressed suffix. Therefore the main stress is located according to this accentual property. The main stress is correctly located on the final syllable of the word.

4. LEVEL OF ACCURACY

The algorithm described here has been evaluated using a method that has been used to assess other such algorithms. A corpus of 475 polysyllabic words was compiled from the Brown Corpus of most frequent words of English (Kucera and Francis, 1967). The list contains the most frequent polysyllabic words of English. The algorithm was tested on this list in order to determine its accuracy and to compare it to the accuracy of other systems. After each word was tested, it was evaluated as to whether the stress was located on the correct syllable. When a doubtful stress pattern was encountered, the Canadian Gage Dictionary was used as a reference.

For this corpus of 475 most frequent words it was found that the program currently operates at a level of 85% accuracy. This is as accurate or more accurate than other systems discussed in the literature. Using the same corpus of words, Hunnicutt (1976) reported an accuracy of about 68%. Bernstein and Nessly (1981) by using a simple basic stress rule alone reported an accuracy of 75%. Church (1985), by taking advantage of the accentual properties of suffixes plus a basic stress rule has achieved an accuracy level of 82%.

A further test of accuracy for the algorithm described here was run on a text that will be used as part of an information delivery system under development at the Centre for Speech Technology Research. The text consisted of 1318 words, 145 of which were unique occurrences of polysyllabic words. The algorithm maintained its accuracy level of 85%.

This high level of accuracy was achieved without the use of an "exceptions dictionary". Klatt (1987) has noted that when an algorithm that predicts stress at a level of 85% accuracy is combined with an exceptions dictionary of 2000 words, this results in an overall accuracy of 97%. This suggests that the use of such an exceptions dictionary with the stress-assignment algorithm described here will produce a level of accuracy that will be quite adequate for text-to-speech synthesis.

5. SUMMARY

In summary, the algorithm presented here represents a computationally efficient and accurate system for the task of assigning primary stress to English words. Combined with a set of letter-to-sound rules, it allows one to type any English word and have as output a phonemic representation of the word with the location of the primary stressed syllable.

The algorithm has been designed so that if there are any further changes that need to be made to increase the accuracy, this can be accomplished with great facility.

6. ACKNOWLEDGEMENTS

The work reported in this document has been funded by a grant from the Science Council of British Columbia.

REFERENCES

Bernstein, J. and Nessly, L. (1981). "Performance comparison of component algorithms for the phonemicization of orthography." *Proceedings of the 19th Meeting of the Association of Computational Linguistics*: 19-22.

Chomsky, N. and Halle, M. (1968). The Sound Pattern of English. Harper and Row. New York.

Church, K.W. (1985). "Stress Assignment in Letter-to-Sound Rules for Speech Synthesis." *Proceedings of the 23rd Meeting of the Association for Computational Linguistics*: 246-253.

Eady, S.J., T.M.S. Hemphill, J.R. Woolsey, and J.A.W. Clayards. (1989). "Development of a demisyllable-based speech synthesis system." *Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, this volume.

Fudge, Eric. (1985). English Word Stress. George Allen and Unwin Publishers.

Hunnicutt, Sharon. (1976). "Phonological Rules for a Text-to-Speech System." *American Journal of Computational Linguistics*. Microfiche 57. 1-72.

Klatt, Dennis (1987). "Review of text-to-speech conversion for English." *Journal of the Acoustic Society of America*, 82(3). 737-793.

Kucera, H. and Francis, W.N. (1967). Computational Analysis of Presentday American English. Brown U.P. Providence, R.I.

Liberman, M. and A. Prince. (1977). "On stress and linguistic rhythm". *Linguistic Inquiry* 8: 249-336.

Selkirk, E.O. (1984). Phonology and Syntax: the Relation between Sound and Structure. MIT Press, Cambridge, Massachusetts.