THE EFFECT OF UNSTRESSED AFFIXES ON STRESS BEAT LOCATION IN ENGLISH

Robert Allen Fox and Ilse Lehiste

Speech & Hearing Science, Department of Communication, The Ohio State University, 324 Derby Hall; and Department of Linguistics, The Ohio State University, 204 Cunz Hall; Columbus, OH 43210 (This research was supported by Grant #1 R01 NS21121-01 from NINCDS to RAF, principal investigator).

Introduction

Although listeners commonly hear speech as rhythmical (Donovan & Darwin, 1979; Lehiste, 1972) it is not the case that the perception of rhythmicity arises from acoustic onset isochrony. For example, if sequences of monosyllables whose initial consonants differ in manner of articulation, are presented to listeners so that the acoustic onset-to-onset intervals are isochronous, the rhythm of the sequence will sound irregular. These sequences will sound regular to listeners only if systematic deviations from acoustic isochrony are introduced (Morton, Marcus, & Frankish, 1976; Fowler, 1979, 1983). Talkers behave in a similar manner in that when required to produce rhythmic sequences of monosyllables which contain different initial consonants the same kinds of deviations from isochrony are found (Allen, 1972a,b; Rapp, 1971; Fowler, 1979; Fowler & Tassinary, 1981). The term "stress beat" or "perceptual center" has been used in the literature to reference that point (or psychological event) in a stimulus upon which listeners/talkers base their rhythmic judgments.

In the past 15 years, a number of experimental studies have been directed at identifying the parameters which determine the location of this stress beat in both perception and production. Experimental results have supported the assertion that the stress-beat location is not universally linked to any particular articulatory or acoustic event, but rather can be shifted by the acoustic/articulatory characteristics of the entire syllable. For example, we have been engaged in research over the past two years examining the influence of several different phonetic parameters on the location of the stress beat in stressed CV or CVC monosyllables in both production and perception tasks. We have found that final consonant variations can shift the location of the stress beat for both talkers and listeners--an effect opposite in direction, but smaller in degree than, the shifts obtained by Fowler (1979; Fowler & Tassinary, 1981) when manipulating the initial consonant (Fox & Lehiste, 1985a). Similar results have been obtained when the medial vowel was modified (for both listeners and talkers); namely, that the stress-beat location shifts to a point later in the token as vowel duration increases.

The present study is a continuation of this line of inquiry and examines the effect of unstressed prefixes and suffixes upon the stress-beat location of stressed syllables in American English. Although the results to be presented today stem from a production task only (which we considered to be, necessarily, the first step in our research program), we anticipate that the listening tests will again show a similar effect. These data, then, should provide information about the relative timing of syllables in both production and perception and thus will provide relevant information about speech timing for speech recognition purposes.

Method

Talkers: There were three highly practiced American English talkers--two female, one male--naive to the purposes of the experiment.

Stimuli: The basic stimuli consisted of sets of seven-token sequences similar to those used by Fowler (1979; and Fox & Lehiste, 1985a,b). Each sequence was composed of 7 identical tokens, such as:

peer peer peer peer peer peer peer
appear appear appear appear appear appear appear

The tokens were either stressed monosyllables (the basic form) or 2-, 3-, or 4-syllable tokens. The latter were formed from the monosyllable by adding an unstressed prefix (e.g., a-, con-/com-, de-, be-) or an unstressed suffix (e.g., -er, -ing, -able), or both, to the basic form. Where possible, a prefix which, when appended to the basic form, produced a real word was chosen. The syllabic structure of the basic form included CV, CVC, CCVC, and CVCC. The initial and final consonants of the basic form included oral and nasal stops, fricatives, and liquids. In all cases the stressed syllable of the multisyllabic tokens corresponded to the basic form. Altogether 601 different tokens were constructed. The sequences were put into random order and presented to subjects on a CRT screen under the control of a PDP 11/23 computer, one at a time, in blocks of 52 (including distractor sequences).

Procedure: Talkers were instructed to read the 7-token sequence which appeared on the screen and to produce them in a rhythmic fashion "in time" with the timing pulse. If a talker was dissatisfied with his/her production on any trial, the talker was instructed to repeat the sequence. After successful completion of a trial, the talker hit the return key on the terminal which replaced the old sequence with the next sequence. Intersequence intervals were thus self-timed but averaged 2 sec in duration. The timing pulse was a 1000-Hz pulse, 100 msec in duration. The stimulus onset asychrony (SOA) between the timing pulses was 1000 msec. Talkers heard the timing pulses continuously throughout the experiment. Talkers were given a short break after every third block of stimuli.

Measurements: For each different basic form an acoustically defined point in the stressed syllable was selected which would, presumably, not change in its basic nature when prefixes and/or suffixes were added. These points included stop consonant release in those stressed syllables beginning with a stop (e.g., do, bide, pose, tone, broad), onset of medial vowel in those syllabes beginning with a fricative (e.g., cede, seal/-ceal), etc. The onset of this measurement point, relative to the onset of the timing pulse, was determined for each token. Of interest in this study is to determine whether or not the position (in time) of these measurement points shifted as a function of adding unstressed prefixes or suffixes. Although the stress beat does not seem to correspond to any particular acoustic event (cf. Fowler 1979; Marcus, 1981), we assume that if the position of these measurement points shifts in affixed tokens, relative to the unaffixed, basic form, it will indicate a concomitant shift in the stressed syllable's stress beat location.

Results and Discussion

Since the location of the acoustically defined measurement points differs across different stressed syllable types (e.g., those having syllable-initial stops vs. fricatives vs. liquids), it makes little sense to compare them directly across all affix conditions. However, if we take the location of the measurement point relative to the timing pulse in the basic form as a baseline location, we can calculate the shift of the measurement point, relative to the basic

form, for all affixed versions of each basic form. To do this, the onset of the measurement point (relative to the timing pulse) of the basic, monosyllabic form of each token was subtracted from the onset of the measurement point in each of that form's variations. For example, the onset of the stop release (relative to the onset of the timing pulse) of the [p] in peer was subtracted from the stop release onsets (relative to the timing pulse) of the [p] in peer, peerer, peering, appear, appearer, and appearing. The resulting number indicates the shift of the measurement point relative to the basic, monosyllabic form. Shown in Table 1 are the mean shifts obtained for those tokens which were prefixed with a-, de-, and con-. Positive numbers indicate a shift of the measurement point to a position later than that of the basic form, negative numbers a shift to an earlier position.

Table 1. Mean shifts in "measurement points" in affixed conditions. Data are normalized relative to onset of measurement point of unprefixed, unaffixed basic form (with defined shift of 0.0). All data are in msec.

		Suffix			
		None	-er	-ing	-able
a-	no prefix	0.0	-3.7	-2.9	-5.9
(N=137)	prefixed	26.6	22.1	11.5	----
de-	no prefix	0.0	-12.7	-4.3	-7.9
(N=35)	prefixed	34.7	31.6	34.0	----
con-	no prefix	0.0	-14.4	-18.5	-5.2
(N=22)	prefixed	86.3	56.7	59.5	----
MEAN	no prefix	0.0	-6.5	-4.8	-6.2
	prefix	36.0	28.6	22.5	----

Although only of borderline significance in each case ($p<.08$, 1-tailed t-test and Wilcoxon), there is a small mean shift (to an earlier point) of the measurement points in the suffixed forms relative to the unsuffixed forms. This indicates that the location of the stress beat may occur later in the token when additional phonetic elements are appended and the overall duration of the token is increased. This result is consistent with the data obtained by Fox & Lehiste (1985a,b) who demonstrated that such shifts can be obtained by manipulating the medial vowel and final consonants of stressed monosyllables.

There is a much larger mean shift of the measurement point (relative to the basic form) in the opposite direction (i.e., later) when tokens have an unstressed prefix; all these shifts are significant at the .001 level (2-tailed t-tests). This indicates that the addition of a prefix shifts the location of the stress beat to a point earlier in the token. It is interesting to note that the a-, de- and con- prefixes produce a progressively greater shift of the location of the stress beat, respectively. This is most likely explained by the fact that although all three prefixes can be considered "unstressed," they are really not all unstressed to the same degree. The a- prefix usually has the least amount of stress and the con- prefix the most.

In order that these data can be examined globally in terms of the relative contribution of prefix, suffix, and prefix+suffix combinations, analysis of variance was done on a subset of these data--namely the a- data. The dependent variable used in this analysis was the time of the release of the initial stop consonant of the stressed syllable, relative to the onset of the timing pulse. These raw data were used instead of the normalized data because (1) one cell of the normalized data would have a variance of zero and (2) the stop release measure represents the same articulatory and acoustic event in all tokens. Shown in Table 2 are the relevant data averaged over basic forms and subjects.

Table 2. Mean onset of initial stop release of stressed syllable relative to onset of timing pulse in a- prefixed tokens. All measurements are in msec.

	Suffix			
	None	-er	-ing	MEAN
no prefix	-31.9	-39.6	-36.0	-35.8
prefixed	2.0	-1.0	-14.6	-4.2
MEAN	-15.0	-20.0	-25.3	-20.0

A two-way, repeated measures, analysis of variance (using basic form as the random variable) with the factors PREFIX and SUFFIX was done. The results showed a significant main effect due to PREFIX ($F(1,28)=34.2$, $p<.001$), a marginally significant main effect due to SUFFIX ($F(2,28)=2.28$, $p=.05$), but no significant PREFIX by SUFFIX interaction.

These results suggest that the location of the stress beat in stressed syllables in English can be affected by the addition of either an unstressed suffix or an unstressed prefix or both. The effects of such affixes on the stress beat are additive and independent of each other. In addition, the prefix seems to shift the stress beat differentially, as a function of its degree of stress. We are currently analyzing these data in terms of how well the durations of the prefix, affix, stressed syllable, etc. can predict the shifts in the measurement points (and, indirectly, the stress-beat location) and are conducting the appropriate, corresponding listening tests.

References

Allen, G.A. (1972a,b). The location of rhythmic stress beats in English: An experimental study. Parts I and II. Lang. Speech, 15, 72-100, 179-195.

Fowler, C.A. (1979). "Perceptual centers" in speech production and perception. Perc. & Psychophys., 25, 375-388.

Fowler, C.A. (1983). Converging sources of evidence on spoken and perceived rhythms of speech: Cyclic production of vowels in monosyllabic stress feet. J. Exp. Psych.: Gen., 112, 386-412.

Fowler, C.A. & Tassinary, L. (1981). Natural measurement criteria for speech: The anisochrony illusion. In J. Long & A. Baddeley (eds.), Attention and Performance, IX. Erlbaum.

Fox, R.A. & Lehiste, I. (1985a). The effect of final consonant structure on syllable onset location. J. Acous. Soc. Am., Suppl 1, 77, S54 (Abstract).

Fox, R.A. & Lehiste, I. (1985b). The effect of vowel quality variations on stress-beat location. J. Acous. Soc. Am., Suppl. 1, 78, S20 (Abstract).

Donovan, A. & Darwin, C. (1979). The perceived rhythm of speech. Proceedings of the Ninth International Congress of Phonetic Sciences, 2, 268-274.

Lehiste, I. (1972). Rhythmic units and syntactic units in production and perception. J. Acous. Soc. Am., 54, 1228-1234.

Marcus, S. (1981). Acoustic determinants of perceptual center (P-center) location. Perc. Psychophys., 30, 247-256.

Morton, J., Marcus, S., & Frankish, C. (1976). Perceptual centers (P-centers). Psych. Rev., 83, 405-408.