STRESS

By: **A.E. Berinstein**

WPP 47

Nov. 1979

A Cross-Linguistic Study on the Perception and Production of Stress

Ava E. Berinstein

UCLA Working Papers in Phonetics 47

November 1979

University of California, Los Angeles

for

my mother and my father

Table of Contents

List of Tables

List of Figures

ACKNOWLEDGMENTS

There are so many people who, in so many different ways, have helped me and influenced my way of thinking, that, only with a certain reserve am I enumerating those that come to mind; since, to name everyone, would surely be an impossibility.

My thanks to the director of the Proyecto Lingüístico Francisco Maroquín in Huehuetenango, Guatemala and to all the teachers there. Special thanks to Sarah Violeta Pérez for aiding in the Spanish explanations and audiotaped instructions of the perception experiment. Thanks to Señora Echeverría, director of the Escuela Nacíonal de Verano Numero 1, and Señor Romeo Arroyave, director of the Instituto Basico "3 de Junio", in Patzicia, Guatemala and to the mayor of Patzicia, Juan Miculax Xalix, alcalde Municipal Interno. Stephen Stewart, director of the language program at the Centro San Benito, in Cobán, Guatemala, the teachers and students there, Flora Ac Caal who assisted in explaining the perception experiment to the K'ekchi participants and to Maria Teresa de Masias, director of Canton las Casas, in Cobán, Guatemala. Special thanks to Francis Eachus and Ruth Carlson who have helped me in so many ways, both in and out of Guatemala, that its trivial to mention their aid in recruiting subjects, finding testing areas and revision of word lists. To them, I can only express love and gratitude for everything.

An initial statistical overview of the perception experiment was made possible by "Alice" whose access and understanding of, I owe to Susan Carey-Block, Cornelia Parkes and Ed Walker.

Grants from the National Science Foundation and the National Institute of Health to the UCLA phonetics laboratory have provided financial support for me as a graduate student.

Thanks to the UCLA phonetics group, Bill Bright, Jeff Elman, Jonas Nartey, Paola Adriani Neiman, Mary Smith, Renée Wellin, and, for several good suggestions (over the years), thanks to Cathe Browman, Louis Goldstein, Steven Greenberg, Jean-Marie Hombert, Peter Ladefoged, Ian Maddieson, Jim Wright and especially, Vicki Fromkin and her wonderful enthusiasm.

I have profited enormously from discussions and suggestions from Vincent van Heuven who, as well, designed the regression analysis and helped in all statistical analyses presented in this version of the paper. To him, a very special thank you.

Then, there's a thank you of another sort, to Adagio; the espresso, champagne, Ginestet, fresh flowers and candlelight, to all those that work there, and especially to Ron, who created it.

ABSTRACT OF THE THESIS

A Cross-Linguistic Study on the

Perception and Production of Stress

by

Ava E. Berinstein

Master of Arts in Linguistics

University of California, Los Angeles, 1979

Professor Peter Ladefoged, Chair

This thesis discusses two experimental studies concerned with cross-linguistic properties of stress as revealed in speech perception and production. One aim was to investigate the extent to which languages differ in the relative importance of duration in the perception and production of stress. The hypothesis tested was that such a difference exists, that it is non-random, and that it differs according to the phonological structure of languages.

One experiment examined the contribution of increased vowel duration to the perception of stress. It was found that subjects' responses showed a bias in favor of a particular stress position, and that this bias differed systematically across the languages studied: English showed initial syllable bias, Spanish no position bias, and K'ekchi (a Mayan language spoken in Guatemala) a final syllable bias. These differences are directly traceable to the stress patterns and phonology of these languages.

As a follow up to the original perception experiment, the acoustic correlates of stress in K'ekchi were investigated. Since variations

in vowel or syllable duration did not play a significant role in stress perception in K'ekchi, although it did for English, it was hypothesized that in the production of K'ekchi stressed syllables, duration would not prove to be a major acoustic factor. The quantitative analysis of Fo,intensity, and duration showed that stressed syllables were significantly higher in pitch and amplitude but not longer than unstressed syllables. This may be explained by the fact that there is phonemic vowel length in K'ekchi, preventing the use of duration from being used non-contrastively as a correlate of stress. The hypothesis that duration will not be an important cue for stress in a language with such a length contrast is thus supported.

This hypothesis was further tested by comparing the perceptual responses of the K'ekchi speakers with those of Cakchiquel, another Mayan language, using the same length varying stimuli. Both K'ekchi and Cakchiquel have fixed final stress but they differ in that while K'ekchi has phonemic vowel length, Cakchiquel does not. This phonological difference, as predicted, resulted in different responses: Cakchiquel speakers used variations in duration as a perceptual cue for stress whereas the K'ekchi speakers did not.

1.0 Introduction

The phenomenon of phonetic stress has long interested phoneticans and linguists. Stress production has been associated with greater muscular effort. Fromkin and Ladefoged (1966), in a study following Stetson (1951)using the technique of electromyography, found small increases in subglottal pressure and increased activity of internal intercostal muscles preceding phonetically stressed syllables. Catford (1977) suggests that "initiator power" should be equated with phonetic stress.

A number of acoustic correlates have been associated with stressed syllables; a rise of fundamental frequency (Medress et al, 1971), longer duration, higher intensity, and fuller vowel quality (Lieberman, 1967). However, these correlates do not seem to be equally important as perceptual cues. This has been demonstrated in experiments using synthetic speech in which each acoustic correlate has been varied individually to determine which, if any, is more predictive.

Fry (1958) found that in English, fundamental frequency is more important than duration, and that duration is more important than intensity (Fry 1955). This finding for English need not, of course, hold for all languages. Thus, phoneticians have generally stated perceptual correlates of stress as general tendencies rather than exceptionless universals, as illustrated by Lehiste:

> "The perception of stressedness appears to be based on a number of factors, the most influential of which is fundamental frequency, other phonetic correlates of stress, besides fundamental frequency and intensity, include vowel quality and duration." (Lehiste, 1970:153)

However, in most discussions of stress, a relative ranking of the cues is implied with fundamental frequency considered the most important. Bolinger (1958) argues for a hierarchy of perceptual correlates with change of pitch as the most important cue, followed by greater duration, and finally by greater intensity.

Some linguists suggest that this hierarchy is universal. Hyman (1977) implies this when he states:

> "The experiments of Fry (1955, 1958), Bolinger (1958), Jasem (1959), and others have demonstrated that pitch may universally best fit our perceptual capabilities followed by duration, and finally by intensity" (Hyman, 1977, p.40).

A major question that arises is the extent to which the relative importance of a given acoustic correlate of stress is cross-linguistically uniform and the extent to which it varies from language to language.

In the attempt to answer this question, two hypotheses will be discussed, H-1 may be stated as follows:

H-1: There exists a universal hierarchy of acoustic cues for perception of stress, such that in all languages the most important cue will be change of Fo, followed in importance by increased duration, and followed, lastly, by the increased intensity of the stressed syllable.

This seems to be the hypothesis suggested by Bolinger for English and generalized to all languages by Hyman. If this hypothesis is correct, the linguistic structure of the language would have no bearing on the hierarchy of importance of the three correlates. This hypothesis is easily falsifiable by showing that in a given language, any one of the cues has a ranking different from that predicted.

H-2: An alternative hypothesis, H-2, may be stated as:
Change in Fo, increased duration, and increased intensity, in that order, constitute the unmarked universal hierarchy for perception of stress in languages with no phonetic contrasts in tone or vowel length; in languages with such contrasts the perceptual cue correlated with that contrast (i.e. Fo with tone and duration with length) will be superseded by the other cues in the hierarchy.

This hypothesis would thus predict that in a language with a phonemic length contrast, duration will be the least important cue, Fo the most important, and intensity, second in importance. Similarly in a tone language, variations of Fo of the type used in making the contrast will be the least important cue for determining stressed syllables, duration will be the most important, and intensity second in importance.[1]

The aim of this thesis is to test these hypotheses. Should it be found that even in one language with a contrast in tone or length, the hierarchy predicted by H-1 fails, H-1 will be invalidated. If the prediction is according to H-2, there will be some support for the notion that languages differ in the importance of each correlate of stress, and that they differ in a non-random way, according to their phonological structure.

Chapter 2, as part of the general question of the thesis will discuss the result of an experiment investigating the extent to which speakers of languages with inherently different stress assignment rules use duration to discriminate a stressed syllable in a syllable sequence. Listeners from three language groups served as subjects. The Mayan language, K'ekchi (final stress), English (variable stress) and Spanish (variable, but predominantly penultimate stress). It was found that subjects' responses showed a bias in favor of a particular position and that the bias differed systematically across languages.

Chapter 3 discusses a study of the acoustic correlates of stress in K'ekchi. Pitch movement, pitch height, peak vowel intensity and vowel duration were measured. Using regression formulas with coefficients derived from stepwise multiple regression, it was possible to predict stress in 94% of the cases across three conditions defined by phonological environment. In these predictions, the change in fundamental frequency made the greatest contribution to the variance accounted for, intensity made the next greatest contribution and duration had virtually no effect.

Chapter 4 explicitly tests H-2 by comparing the results of the experiment in which speakers of two Mayan languages of the Quichean family were subjects. Both languages have final stress but K'ekchi, has a phonemic length contrast, and Cakchiquel does not. This experiment, it will be shown, supports H-2, in that Cakchiquel speakers used duration as a perceptual cue for stress while K'ekchi speakers did not.

1

While there were no tone languages in this sample, acoustic evidence from Siamese (Gandour, 1974) does support the prediction of H-2. In Siamese stressed syllables are longer in duration than unstressed syllables and there is much less pitch displacement in unstressed syllables.

However, in a tone language, if the contrast for example, is between level tones - the precition of H-2 is that level tones will provide the weakest cue to a stressed syllable. That is NOT to say that a gliding tone or a change of Fo is the weakest cue also, i.e. it may well be the case that in languages with phonemic pitch and stress - different uses of pitch are relevant to each. It is probably for this reason that Pike and Scott (1975) find the characterization of Fore pitch patterns (a language of New Guinea) somewhat ambiguous between tone and stress.

CHAPTER TWO

A cross-linguistic study on the contribution of duration to the perception of stress[1]

2.0 Introduction

The two hypotheses mentioned briefly above counterpose a universal hierarchy of perceptual cues to a language specific hierarchy. H-2, if validated, would show the extent to which stress is perceived according to language dependent cues. If one finds, for example, that similar stimuli are judged differently by subjects from different language groups, and that these differences correlate with specific aspects of the language phonological system, then such findings would have some bearing on this question.

It has, in fact, been suggested that this is true, i.e. that speakers of a language with a favored position for lexical stress will be inclined to perceive stress in that position even when the acoustic characteristics of the stimuli contain no special cues for that position.

> "Knocks produced at even intervals,with every third louder, are perceived as groups of three separated by a pause. The pause is usually claimed by a Czech to fall before the louder knock, by a Frenchman to fall after the louder; while a Pole hears the pause one knock after the louder. The different perceptions correspond exactly to the position of the word stress in the languages involved. In Czech the stress is on the initial syllable, in French, on the final and in Polish, on the penult. When the knocks are produced with equal loudness but with a louder interval after every third, the Czech attributes greater loudness to the first knock, the Pole, to the second and Frenchman, to the third." (Jakobson et al. 1951; 10,11).

In other words, according to Jakobson and his colleagues, all other things being equal, the syllable position that speakers expect to hear as stress, here referred to as the bias position, will depend on the phonological and phonetic stress regularities in a language.

The experiment discussed in this chapter attempted to test this further by using speakers from three different languages, each with a different stress pattern, and by manipulating duration to see what the effect of this might be.[2]

2.1 Experiment 1.

2.2 Method

Subjects

The subjects composed speakers from three language groups: English (free stress), Spanish (free stress, but predominantly penultimate) and the Mayan language K'ekchi (fixed final stress). In K'ekchi, there is also phonemic vowel length.

A. English

36 native speakers of English were tested in a classroom; 31 were students in an introductory phonetics class, 2 were professors at UCLA and 3 were graduate students participating in a seminar on Intonation. Of this group, 26 were females and 10 were males. The mean age was 22.

B. Spanish

22 monolingual Spanish speaking persons from Huehuetenango, Guatemala, were tested in a classroom. Of this group, 18 were females and 8 were males. All participants were teachers at the Proyecto Lingüístico Francisco Marroquín, a school organized for teaching Spanish to foreign students. The mean age was 23.

C. K'ekchi

K'ekchi is a Mayan language of Guatemala. It is a language of the Quichean subgroup. All 31 speakers were from Cobán, Department of Alta Vera Paz, Guatemala. Of this group there were 25 bilingual (K'ekchi-Spanish) sixth grade students; 13 females and 12 males. The ages ranged from 11 to 16 with an average age of 14. 3 bilingual (K'ekchi-Spanish) teachers from the Centro San Benito in Cobán, Guatemala whose average age was 26, and 3 monolingual K'ekchi students from the Centro San Benito whose average age was 25, also participated. All groups were tested in classrooms. All subjects were from K'ekchi families.

Stimulus Material

The stimuli (sequences of four segmentally identical CV syllables) were synthesized on an OVE III speech synthesizer whose parameter values were controlled by a PDP 12 computer. The basic syllable was [bɩ] with formant center frequencies at 400, 1900 and 2550 Hz for the steady state portion of the vowel. There was one basic vowel duration of 100 msec which was used for three of the four syllables while the other was one of six durations: 70, 100, 120, 140, 160 or 200.

In order not to confound the design of the experiment by other prosodic factors, especially the fundamental frequency variations as a function of syllable duration, it was decided to synthesize the stimuli on a monotone at 100 Hz. A basic syllable duration of 140 msec. (a 40 msec CV transition followed by a 100 msec. steady state vowel portion) was chosen to minimize the influence of the monotonicity of the fundamental frequency. A speech rate of 7.1 syllables per second is fast, but still within the capabilities of the human articulatory mechanism.

For example, Hudgins and Stetson (1937) instructed speakers of **English** to produce the sequence pupupupú as fast as they could. They found a range of durations for the unstressed syllables between 5.7 and 7.7 per second with a mean of 6.7. Further, given that short /u/ in English, is 20 msec. longer than short /I/ (cf. Peterson and Lehiste, 1960), an unstressed syllable duration of 140 msec (7.1 syllables per second) is well within this range.

The position of the deviating syllable was orthogonally combined with each of the five deviating values, yielding 4 x 5 + 1 stimulus types , which were recorded with four second interstimulus intervals on audiotape in two different random orders.

To determine what the stress response bias is of a given language, it is assumed that if speakers have some abstract representation of stress patterns (i.e. of an expected stress position), that it will be possible to override this perceptual bias by manipulating duration, thereby attaining the effect of stress in that position at significantly shorter durations than any other. For this reason, four CV syllables, as opposed to two or three are used. Four syllables allow all possible stress biases; i.e. initial- distinct from antepenultimate, penultimate and final. It therefore will be unambiguous as to whether the results are due to a boundary effect or internal word structure. Furthermore, as has been noted in the literature by several investigators- there is a general preponderance for listeners to perceive the first of a group of sounds as louder or accented (Woodrow, 1951 , Fraise, 1963) relative to the others. Recent perceptual studies also indicate that the first of two positions, **of equal duration, is perceived as longer than the second when only Fo is varied. (Lehiste 1976, Pisoni, 1976). Chuang et. al (1977) for example, proposed as an explanation for these and similar phenomena that there exists a psychophysical time order bias and that it is for this reason that the first position is perceived as lower in pitch, louder in intensity and longer in duration. Crucially, this work is based on perceptual tasks which employed the manipulation of only two positions. The questions that remain to be asked are: 1) is it "accidental" that the time-order-error (TOE) and the expectation of an initial accent in English coincide? and 2) if there exists a TOE, do all languages (i.e. regardless of linguistic structure) have a bias for initial position? (which is implied by Chuang's proposal).**

The point here is that if two syllables are used, the result is completely ambiguous as to whether the preference for initial position is due to a) an expected initial stress, b) an expected phrase/word final lengthening (Klatt 1975 , Oller 1973 , Lindblom 1968), or c) a TOE. It is for this reason that the task **described** here uses four syllables.

Procedure

The stimuli were presented to speakers of three different languages: English, Spanish and K'ekchi. The English group received the following written and auditaped instructions:

"In this experiment, you will hear the syllable BI in sequences of four, as: BI BI BI BI. On your answer sheet are the numbers 1 2 3 4 corresponding to each of the syllables: first, second, third and fourth of the sequence. Please circle the number on your answer sheet corresponding to the position of the syllable that you hear as stressed. For example, if the third syllable were stressed, then you would circle the number 3, as: 1 2 (3) 4. Mark one and only one stress. There may be some sequences which you find more difficult than others. If this should happen, please guess. We have found in the past that people's guesses are quite consistent; therefore, if you aren't completely sure, make a guess, but mark only one stress per sequence.

We will now do some practice sequences. Your answers should begin in column one and proceed to the bottom of the page. If you have any questions, please ask them now."

These instructions were translated into Spanish and K'ekchi. They were played on a Sony TC 142 Recorder to both the Spanish and K'ekchi groups and on a Rheem Califone 70 TC Recorder to the English group. The taperecorded instructions (in each native language) were followed by a verbal explanation and a question period before the subsequent practice session. A native speaker assisted in the description of the task and the question-answer period for the Spanish and K'ekchi groups. Instructions and explanations therefore, were always given in the speaker's native language, as were the answer sheets.

As the concept "stress" was not as familiar to either the Spanish or the K'ekchi groups as it was to the English (phonetics) students, this was something which had to be taught. Specifically, the use of length as a cue for "prominence" or "accent" (i.e. the syllable which is somehow distinguishable from all the rest) was unfamiliar to the K'ekchi. The practice stimuli were played two times for the Spanish group and from two to three times for the K'ekchi groups, until everyone felt confident about the task. The practice session consisted of the complete set of 21 stimuli. The first 15 of these constituted the "official practice session". This was followed by a final period for questions before the experiment was given. The remaining six sequences of the practice stimuli were discarded from the results but were presented under the auspices "experiment". The experimental tape included a short tone which was heard at the end of every 15 sequences.

2.3 Results

The contribution of differences in duration at each of the 4 syllable positions for each of the three languages investigated will be examined. The conditions containing four equal 100 msec vowels (hereafter, the control condition) or a 70 msec deviant vowel duration (i.e. a 30 msec decrease with respect to the reference duration) will be analyzed separately.

K'ekchi Results

Figure 2.1 contains the results of the experiment in terms of percentage of perceived stressed syllables as a function of the position of the temporally deviant syllable.

It will be observed in Figure 2.1 that no matter how much the duration of the temporally deviant syllable (henceforth, the target syllable) in the sequence differs from the 100 msec. reference value, the percentage of perceived stresses corresponding to that position in the sequence is roughly the same. However, the percentage of perceived stresses increases monotonically with the sequential positions of the target syllable; 16, 22, 30 and 43% for the target syllable in first, second, third and final position respectively. These results express a perceptual bias in favor of stress in word final position.

In order to establish the statistical significance of these results, the K'ekchi data were submitted to a classical two-way analysis of variance with duration of the target syllable and target position as factors, assuming fixed effects. The criterion variable was the percentage of judgments that a lengthened syllable was stressed.

The results of the analysis of variance show that target syllable position constitutes a highly significant effect, $F(3, 974) = 15.0$, $p. < .001$.

A posteriori tests for contrasts indicate that only the difference in percent correctly perceived stress between first and second positions is insignificant by the Student-Newman Keuls procedure with a .05 level of significance. All other differences among positions however, reach significance.

Duration of the target syllable does not exert an effect, $F(3, 974)=1.1$, n.s. There is no interaction between the two factors, $F(9,974)= 0.5$, n.s.

Spanish Results

Figure 2.2 which is analogous to figure 2.1 presents the results for the Spanish data.

The same statistical analyses as described above, were performed on the Spanish data. Figure 2.2 shows that there is no general effect due to position of the target syllable, $F(3, 687) < 1$, n.s. The percentage of perceived stresses for the target syllable in first, second, third and final positions are 34, 34, 31 and 31% respectively. These results suggest that there is no perceptual bias for a particular position in the word in Spanish.

There is a general increase of percentage of perceived stresses as the target syllable assumes larger durations: 27, 26, 39 and 39% for 120, 140, 160, and 200 msec. respectively, which constitutes a significant effect, $F(3, 687) = 3.8$, $p = .010$.

There is no interaction between the two factors, $F(9, 687) = 1.0$, $p = .427$ n.s.

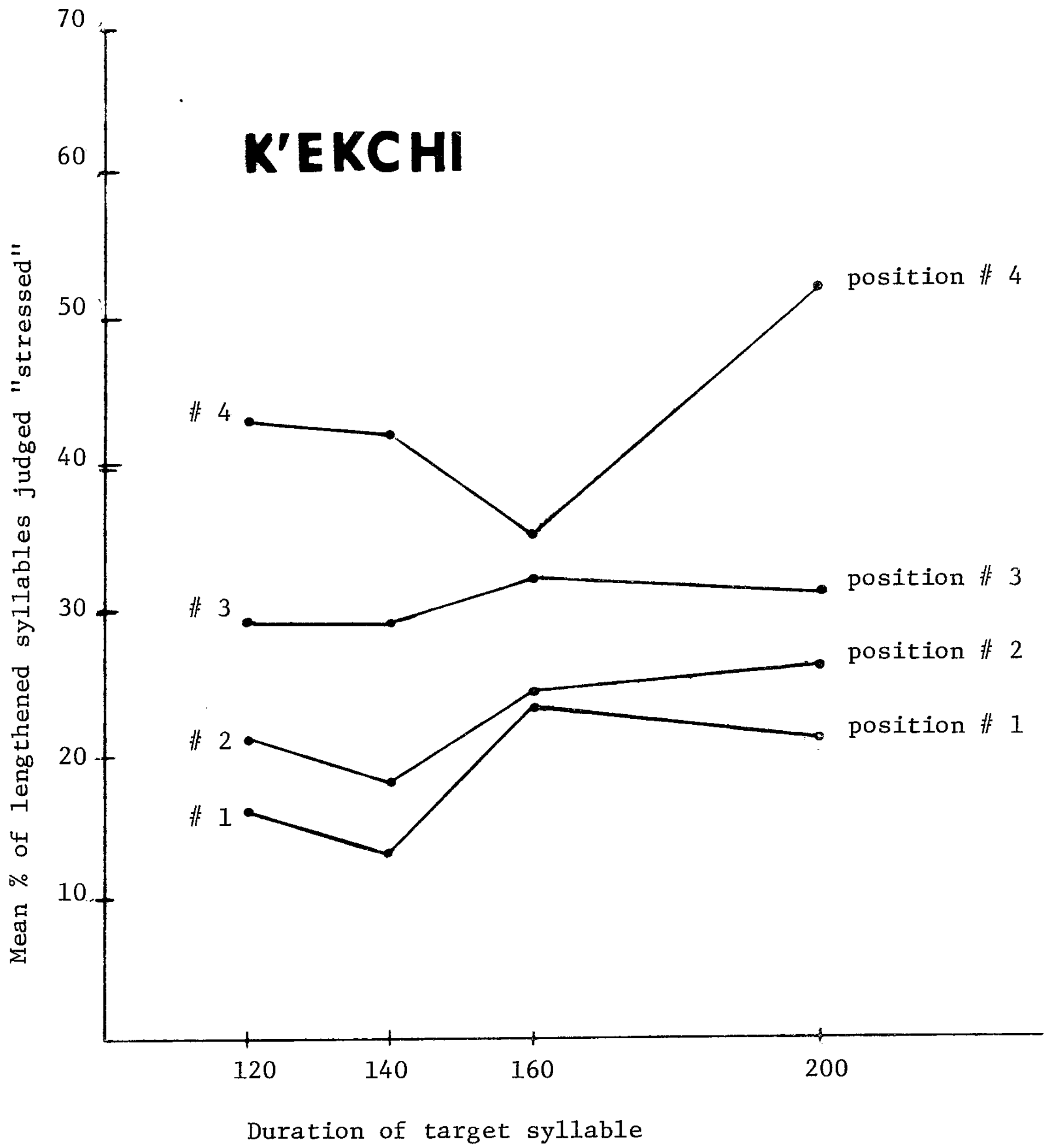

Figure 2.1. Mean percentage of lengthened syllables judged stressed as a function of duration conditions; 120, 140, 160, and 200.

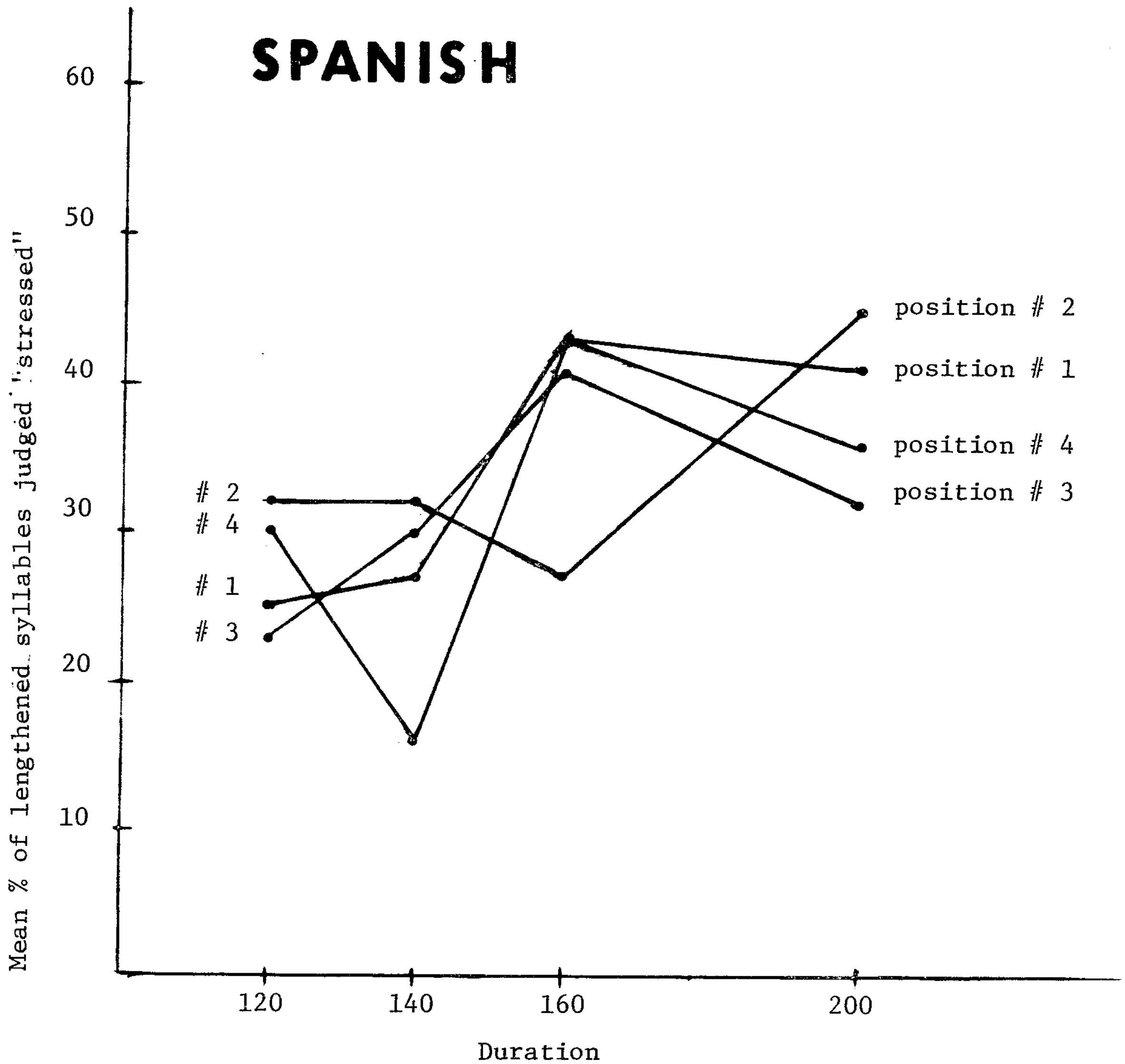

Figure 2.2. Mean percentage of lengthened syllables judged stressed as a function of duration conditions; 120, 140, 160 and 200.

English Results

Figure 2.3 graphs the effects of position and duration of the long syllable on percentage perceived as stressed.

Analysis of variance shows that target syllable position is a highly significant effect: (F(3,1136) = 65.0, p <.001. The percentage of perceived stresses decreases monotonically: 86, 67, 59, and 46% for the target syllable in first, second, third and final positions respectively. These results clearly express a pronounced perceptual bias in favor of stress in word initial position.

Duration of the target syllable also exerts a highly significant effect: F(3, 1136) = 225.5 p <.001. Here, the percentage of perceived stresses increases monotonically as the target syllable assumes longer durations: 31, 44, 89 and 94% for 120, 140, 160 and 200 msec. respectively.

These effects together, explain 42% of the variance: 10 and 32% for position and duration respectively.

It is obvious from figure 2.3 that there is little improvement in the scores when going from 160 to 200 msec (possibly due to a ceiling effect). Also, there is only a moderate difference in percent correctly perceived stress between the 120 and 140 msec conditions. However, the increase in performance level when going from 140 to 160 is much greater. It appears that it is the sharp increase in performance for positions two, three and four, rather than for position number one, that causes this effect. As a result we find a highly significant interaction between position and duration of the longer syllable, F(9, 1136)= 9.4, p <.001.

2.4 Differences between the languages

Summarizing the results so far, we have found that the position of the deviant syllable and its duration are not used to the same extent in the three languages.

- In K'ekchi position is highly influential, but duration is hardly used.
- In Spanish we found the reverse to be true: duration clearly influenced stress perception irrespective of position.
- In English both position and duration are effective, with duration the stronger of the two effects.
- Perceptual biases for a particular stressed syllable position in a word also differed across languages: K'ekchi preferring the final, Spanish-none, and English the initial.

The extent to which positional cues as reflected in the perceptual response bias vary in the languages investigated, will be established in a two way analysis of variance with language and syllable position as factors. Figure 2.4 presents the effects of target syllable position across durations, separated out for the three languages.

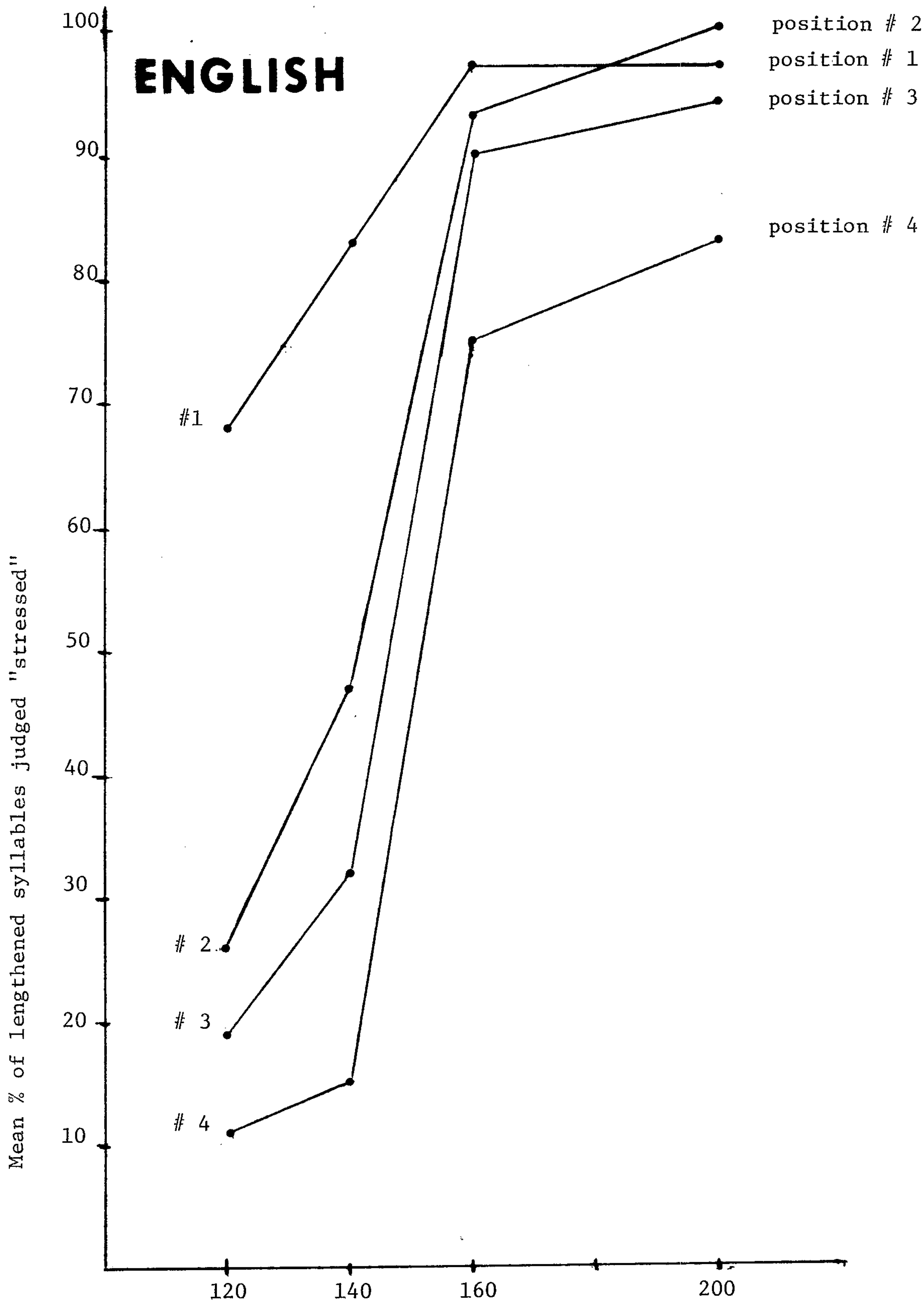

Figure 2.3. Mean percentage of lengthened syllables judged stressed as a function of duration conditions; 120, 140, 160 and 200.

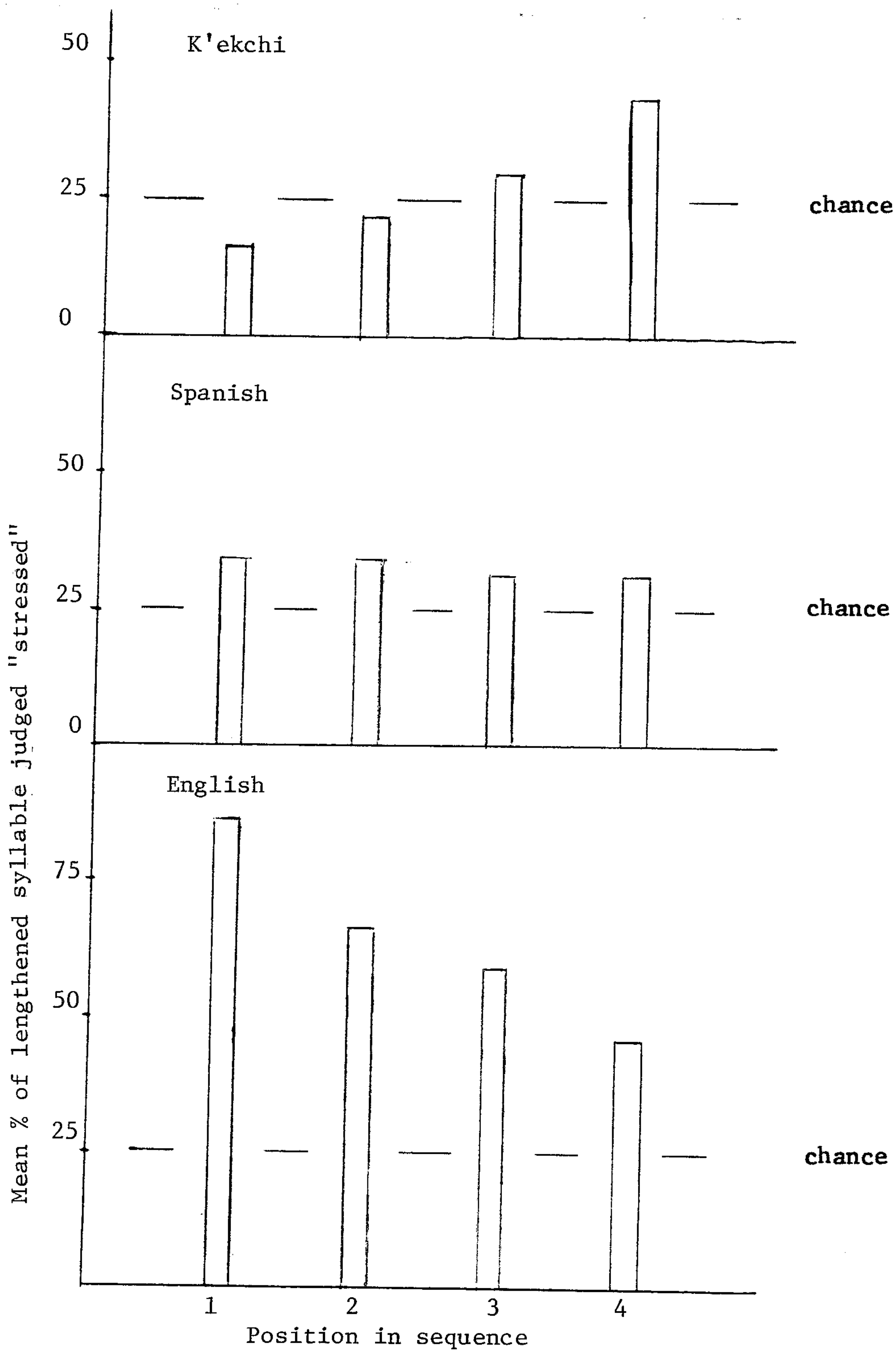

Figure 2.4. Mean percentage of lengthened syllable judged stressed as a function of position for K'ekchi, Spanish and English.

Spanish has roughly the same percentage of correctly identified target syllables at each of the four syllable positions. English performance is very good for targets in initial positions, but drops abruptly for non-initial target positions. K'ekchi, on the other hand, reveals bias in favor of the final syllable and a steady increase of percent correct identifications with target syllable positions toward the end of the word.

English performance is always well above chance. Spanish performance is slighly above chance. K'ekchi performance is below chance in the first two positions, and (considerably) above chance only in position four.

Analysis of variance indicates that there is a highly significant language-position interaction, $F(9,4249)$ 22.0 $p. <.001$, which substantiates the claim that the bias for perceived stress location differs strongly across the three languages.[3]

The extent to which duration cues facilitate the perception of a stressed syllable will be established in a two way analysis of variance with language and duration of the deviant syllable as factors. Figure 2.5 summarizes the effects of durational increments of the target syllable in the three languages.

Clearly, the largest effect is seen in English where increments of 20 and 40 msec lead to a small improvement in performance, but to near perfect identification of the target with 60 to 100 msec increments. A small effect is obtained in Spanish, where performance at chance level is found for 120 and 140 msec. targets, and performance significantly above chance for the two longer conditions. K'ekchi, finally, shows no effect of duration at all; even when the target syllable is twice as long as the unstressed syllables, correct identification is only 7 points above chance which is still short of significance.

Analysis of variance indicates that the difference between the three languages is a highly significant effect, $F(3,4249) = 195.6$, $p <.001$. More significantly, however, the language-duration interaction is highly significant too, $F(9,4249) = 32.2$, $p <.001$, which corrorborates our hypothesis that durational cues are not used to the same extent in each of the languages concerned.

Tables 1, 2 and 3 allow us to see the bias from a different perspective, i.e. regardless of ("correct") identification of the temporally deviant syllable with "stress". These tables represent the mean percentage of all judgments per position as a function of the duration condition separated out for the three languages; K'ekchi, Spanish and English respectively.

In Table 1 we see that in K'ekchi there are more responses for position 4 than any other (41%) irrespective of duration. In Table 2 we see that Spanish stress judgments are fairly equally distributed for each position regardless of duration. In Table 3 we see that English stress judgments are predominantly initial in the shorter duration conditions, 63% and 59% for the 120 and 140 conditions respectively. The fewest overall responses are in final position (15%), and the most responses are in initial position (45%).

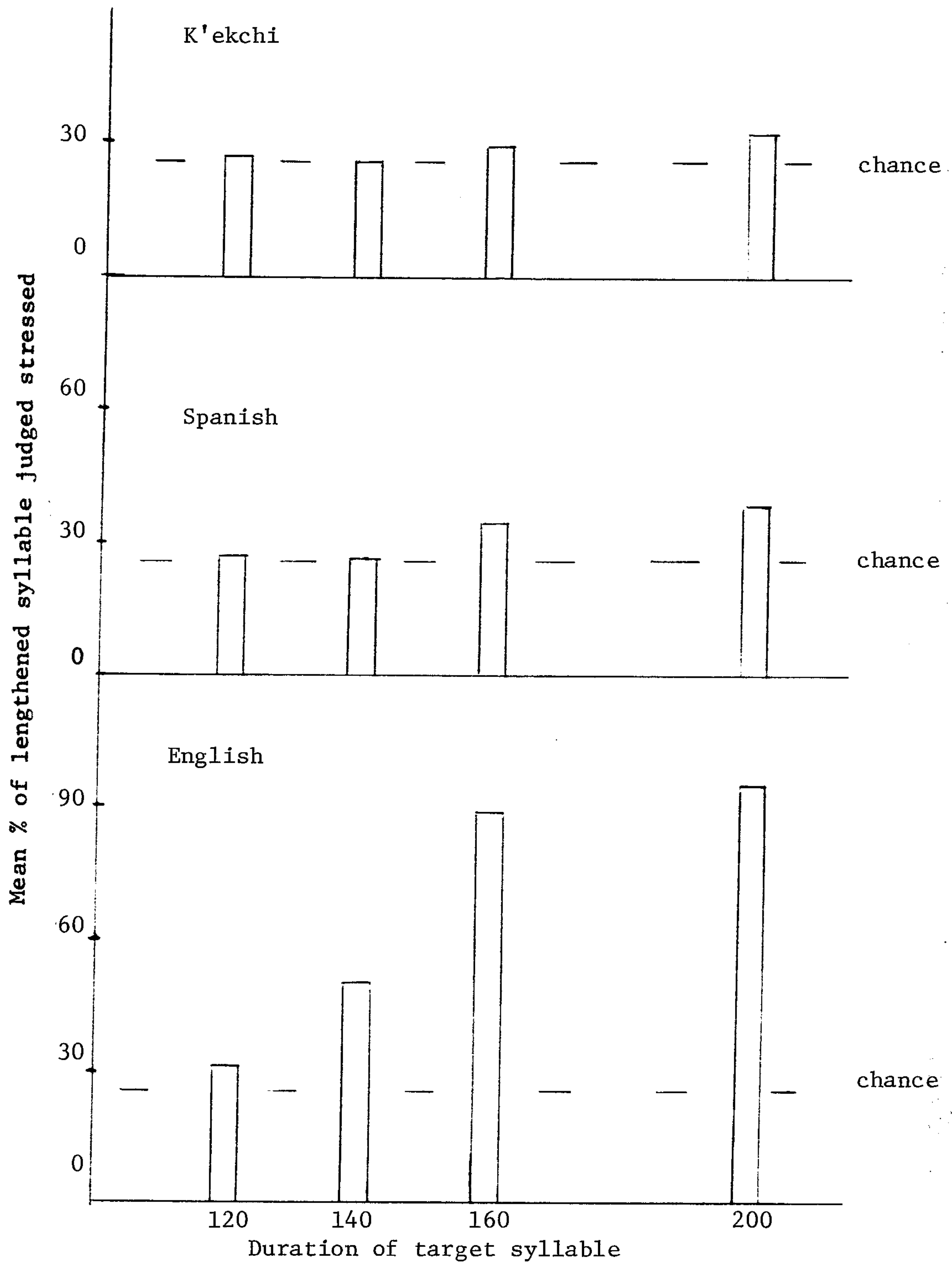

Figure 2.5. Mean percentage of lengthened syllables judged stress as a function of duration condition in K'ekchi, Spanish and English.

Mean % of stress responses/position

Duration Condition

	120	140	160	200	
1	14	12	15	13	= 17%
2	19	19	21	20	= 20%
3	26	26	27	24	= 26%
4	41	43	37	43	= 41%

Table 2.1

Percentage of stress responses/position for duration conditions, 120, 140, 160, and 200 in K'ekchi.

Duration Condition

	120	140	160	200	
1	22	26	27	24	= 25%
2	27	31	25	30	= 29%
3	25	22	23	24	= 24%
4	26	18	25	22	= 23%

Table 2.2

Percentage of stress responses/position for duration conditions, 120, 140, 160, and 200 in Spanish.

Mean % of stress responses/position

Duration Condition

	120	140	160	200	
1	63	59	30	26	= 45%
2	15	20	26	27	= 22%
3	12	11	25	26	= 19%
4	10	10	19	21	= 15%

Table 2.3

Percentage of stress responses/position for duration conditions, 120, 140, 160, and 200 in English.

2.5 Analysis of the control data

The control stimuli- i.e. that condition in which all 4 syllables were of equal duration, allows further verification of the bias differences between the languages concerned. We predict that essentially the same bias differences will be found here, as were found in the previous section. Figure 2.6 represents the percentage of stress judgments in the control condition as a function of syllable position for each of the languages.

The results for the control condition for English and K'ekchi suggest a pattern that favors one position over the others. This bias was seen also in the data as a whole. Figure 2.6, similar to the pattern discussed earlier in Figure 2.4, illustrates an initial bias for English, a final bias for K'ekchi and virtually no bias for Spanish.

If there were no significant position bias, subjects' mean stress responses in the control condition should not differ significantly from 2.5 the average stress position of a four syllable stimulus. This hypothesis was tested. It was found that the average responded stress position in the control sequences for English was 1.5, for K'ekchi 3.2, and for Spanish 2.6. One-way analysis of variance of perceived stress position for the control stimuli yields a significant effect for language $F(3, 265) = 36.1$, $p < .001$, with significant differences at the .05 level between each of the three languages , as determined by the Newman Keuls procedure. Further, the direction of the position bias is confirmed in Figures 2.6 for the control data and 2.4 for the data as a whole; i.e. in English, subjects' means were significantly less than 2.5- indicating a bias towards the beginning of the word. In K'ekchi, subjects' means were significantly greater than 2.5- indicating a bias towards the end of the word. Finally, in Spanish, subjects' means did not significantly differ from 2.5- indicating no bias for initial or final positions. Given that the percentage of responses for each of the syllable positions in Spanish were found to be insignificantly different from each other- it is clear that 2.6 is indicative of no bias, rather than a medial bias.

In sum, the control data were found to be a good predictor of the perceptual biases for the languages investigated. This is best revealed by comparing the results for the data as a whole (figure 2.4) to the results of the control data (figure 2.6).

2.6 Analysis of the 70 msec. condition

A final indication as to the relative importance of durational cues versus cognitive bias in stress perception may be provided by an analysis of the stimuli in which the deviant syllable is shorter than the other syllables.

In figure 2.7 the percentage of perceived stress locations is plotted as a function of the position of the 70 msec syllable (which corresponds to the darkened bar) separated out for the three languages.

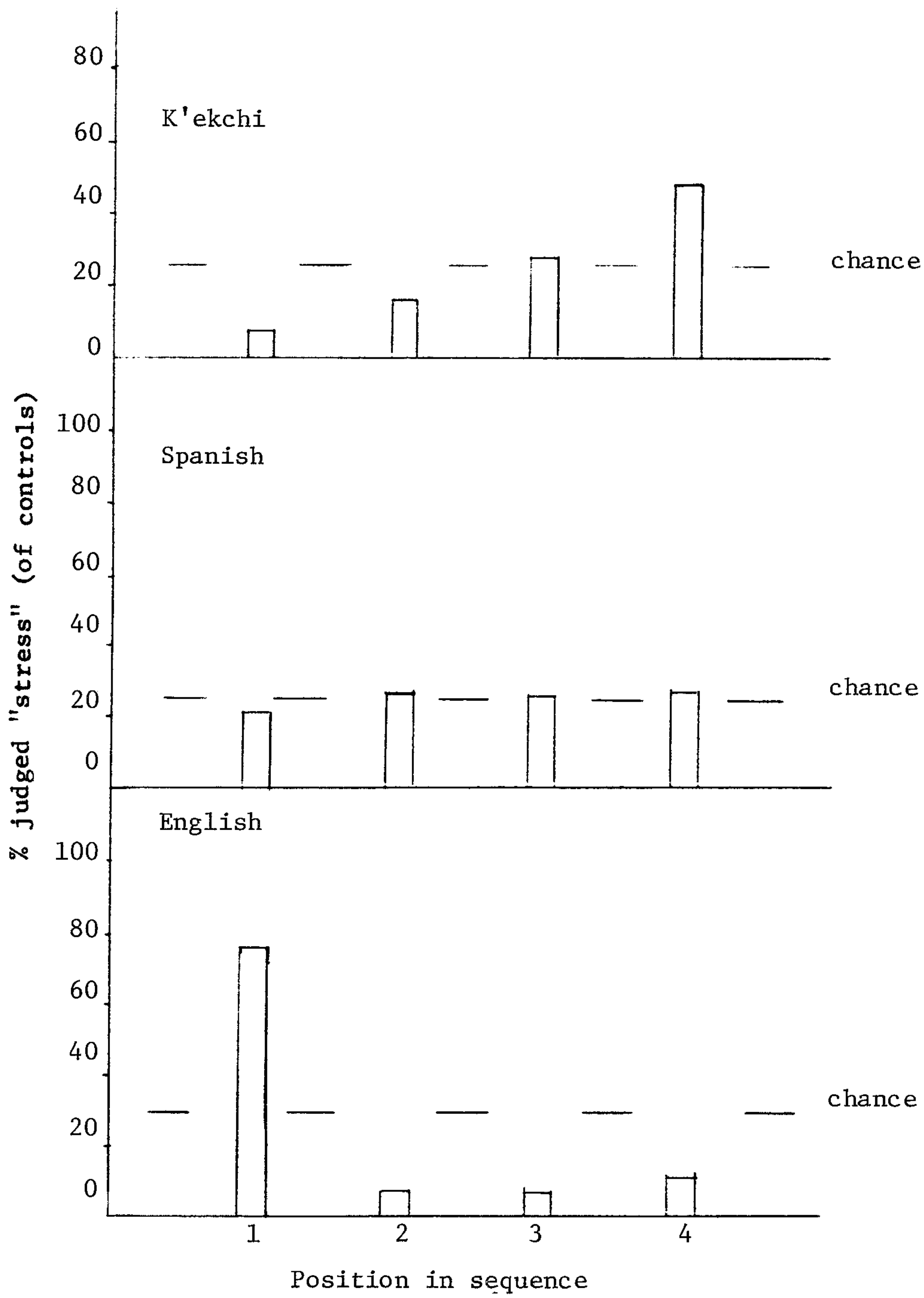

Figure 2.6. Percentage of stress judgments in the control condition per position in K'ekchi, Spanish and English.

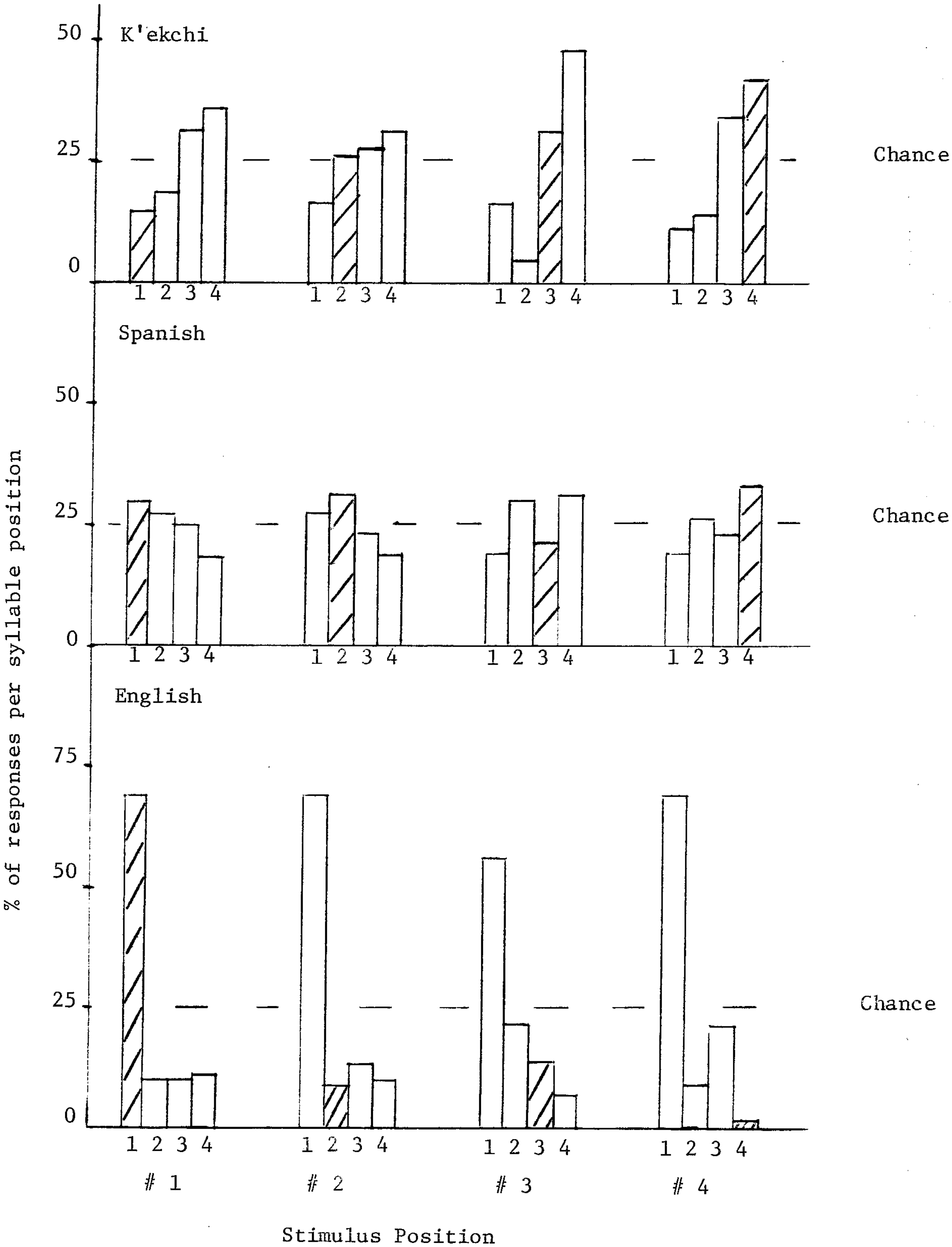

Figure 2.7. Percentage of responses/syllable position as a function of stimulus position and language for the 70 condition.

The results for K'ekchi are generally predictable from the data reported above. The 70 msec target was identified as the stressed syllable in roughly the same proportions as in the control conditions irrespective of the position of the 70 msec syllable. Note for example that when the deviant syllable was in fourth position there were more stress responses for that position than any other. This indicates that the final position bias, rather than durational differences (either shortening or lengthening) is most crucial to the perception of a stressed syllable in K'ekchi.

The data for Spanish are more or less random, indicating that the 30 msec decrement went largely unnoticed. In fact, contrary to expectations, there is a slight increase of stress responses associated with the shortened syllable for all positions except penultimate. These data suggest that there is no position bias and that there is little or no influence of a decrease in duration on the perception of a stressed syllable in Spanish.

In English, as predicted, there is a very clear preponderance of perceived stress in initial position. However, a second effect emerges in that the syllable immediately preceding the shortened syllable receives the largest number of stress judgments for <u>non</u>initial syllables. Let us hypothesize that shortening a particular syllable creates the impression of greater duration on the syllable immediately preceding it. Note for example that when the final syllable is shortened there is a sharp increase of penultimate stress responses. Further, this is significantly different from the percentage of stress responses for that position in the control data (21% as compared to 6%). $X^2 = 7.34$, $df = 1$, $p < .01$. Similarly, when the penultimate syllable is shortened there is a sharp increase of antepenultimate stress responses; again, significantly different from the percentage of stress judgments for that position when compared to the control data (22% as compared to 7%), $X^2 = 6.8$, $df = 1$, $p < .05$. That the initial syllable receives the most stress responses when the antepenult is shortened is consistent with this leftward trend- but clearly predictable given that the initial position is independently favored (perhaps reflecting an additive effect), thus when the deviant syllable is in initial position the pattern is essentially the same as the control data.

These results indicate that English speakers use durational information in the perception of stress, and that the perception of a stressed syllable is based largely on its length relative to the other syllables in the sequence. Thus English speakers seem to have certain expectations about timing information. If the acoustic stimuli does not coincide with our expectations (in this case, expecting the stressed syllable to be longer than the rest) certain adjustments must be made by the listener, as evidenced by the strategy employed here.

2.7 Conclusions and discussion

2.7.1 A first conclusion that may be reached is that languages differ in the extent to which duration is used as a cue to stress.

Evidence for the conclusion comes from:

a) the conditons in which the deviant syllable had an increment of 20,40, 60 or 100 msec from the basic reference duration, and
b) the condition in which the deviant syllable had a decrement of 30 msec from the basic duration.

Note that the results of the two-way analysis of variance with language and duration of the (incremented) deviant syllable as factors indicate that the difference between the three languages is a highly significant effect, $F(3, 4269) = 195.6$, $p < .001$. Moreover, the language-duration interaction was highly significant $F(9, 4249) = 32.2$, $p < .001$, which supports our conclusion that languages differ in the extent to which duration is used as a cue to stress.

In K'ekchi it was found that increases in duration had no influence on the perception of stress. Why should this be the case? It is hypothesized here (and further supported in chapters 3 and 4) that there is a high functional load of duration for the long/short phonemic distinction in the language and it is for this reason that duration is not simultaneously used as a cue for stress.[4]

In Spanish it was found that increases in duration do have a slight influence on the perception of stress. These results, as with K'ekchi are directly relatable to the linguistic structure of the language. In Spanish the variability of vowel data is very restricted. Of particular relevance therefore, is Delattres (1966) observation that: "One receives very different rhythmic impressions from languages whose longest syllables (as a class) are about three and a half times longer (English) two and a half times longer (German, French) or one and a half times longer (Spanish) than the shortest syllables (as a class)." (pp.188-189).

In this experiment, the 20 and 40 msec increments had no effect in Spanish, and a very small effect in English. The 60 and 100 msec increments had a large effect in English, but only a limited effect in Spanish. An explanation for this is that such increments (1.6-2.0 : 1) are a novel experience to a Spanish listener and will not immediately be associated with the linguistic function of stress, given the fact that there is an upper limit of 1.5: 1 to variability in Spanish vowel durations as opposed to English 3.5 : 1 (cf. Delattre, 1966).

In English it was found that both position and duration have a significant influence on the perception of stress. Position one had by far the greatest percentage of stress responses and position four, the least. Further, in English, unlike Spanish, not only did durational increments influence stress perception, but also, durational decrements.

These results follow if we assume that speakers expect the stressed syllable to be longer than any other in the word and that this expectation is dependent upon the relation of the syllables to each other.

We will loosely refer to this as the "Relative Timing Principle". This principle will be of some import in Chapter 3. It should be pointed out that the increase of stress responses for the syllable adjacent and to the left of the shortened syllable (in the 70 condition), is in accordance with the direction of the suggested position bias (e.g. towards the beginning of the word).

As enumerated earlier, there exists considerable variability in the vowel durations in English with the maximum ratio on the order of 3.5 : 1 (Delattre 1966). Thus, in English production there are large durational differences but, contrary to K'ekchi, duration is only marginally relevant for phonemic segmental distinctions; it therefore can be used for other purposes, as stress (or syllable/phrase final lengthening).

2.7.2 A second conclusion is that languages differ in regards to the bias syllable.

Evidence for this conclusion is drawn from the control conditions and the data as a whole. The results of the two-way analysis of variance with language and position of the deviant syllable as factors indicate that the language-position interaction is a highly significant effect, $F(9,4249) = 22.0$, $p < .001$ this supports the conclusion that languages differ in bias.

In the K'ekchi group, the percentage of "stress" responses in the control data clearly indicates a preference for final position, as seen in Figure 2.6 . This same pattern is exemplified in the data as a whole where the percentage of correct responses increases across the word with 16, 22, 30 and 43% for 1st, 2nd, 3rd and 4th positions respectively.

One reason for these results may be because there is a fixed final stress in K'ekchi. This does not mean that in a language with such fixed stress, acoustic cues will not be used. What I am arguing for here, and will develop in Chapter 3, is that duration as a cue for stress is not used in K'ekchi because the average duration difference between stressed and unstressed vowels is statistically insignificant (see Figure 5, Chapter 3.). Thus, the final position bias in K'ekchi is due to the fixed final stress of the language and the irrelevance of duration to the distinction between stressed and non-stressed syllables.

In the Spanish group, the percentage of stress responses in the control data (Figure 2.6) are statistically insignificant from each other as determined by the Newman Keuls procedure at a .05 level of significance. The data as a whole reflect the same pattern; as seen in Figure 2.4, responses are slightly above chance and the percentage of correct responses/position do not statistically differ from each other.

Unlike the dominant penultimate stress pattern of the language, Spanish showed no bias for any position over any other. This is not too surprising when one considers bisyllables as; líbro 'book', libró 'she/he delivered', triplets as: término, 'term', termíno 'I terminate' and terminó 'she/he terminated' and quadruplets as: dígaselo 'tell it to her/him', mecánico 'mechanic,' homenáje,'homage' and encendedór 'cigarette lighter', which argue that stress is in part morphologically determined. Given this we might predict that Spanish speakers would not have a cognitive bias to perceive stress in a particular syllable position. This, in fact, is the case. Delattre (1966) states that in Spanish"there are large categories whose lengths are not greatly affected by stress or position-unstressed open syllables are hardly longer in final position (18.52) than in non-final position (18.16): and non-final open syllables are only slightly longer in stressed syllables (20.23) than in unstressed syllables (18.16)." Thus, for Spanish speakers, (unlike English) there is no final length expectation - so fourth position is still in. One may hypothesize that for a Spanish speaker the expectancy is that the syllables will be of near equal duration; therefore, listeners have no predilection to perceive one position as any longer than any other (or to expect stress in a particular position),and they don't.

English was the only language investigated which exerted a highly significant interaction of position and duration, $F(9,1136) = 9.4$, $p < .001$, therefore the explanation for the initial position bias is somewhat more complicated and raises a number of issues.

When all stimuli were of equal duration (see Figure **2.6)the percentage** of stress judgments is overwhelmingly initial 76%, as compared to 7%, 6%, and 11% for 2nd, 3rd and 4th positions respectively. The mean percent correct is also greater in first position than any other and decreases across the word as exemplified in Figure **2.4.**That position **four is least well** perceived as stressed is also exemplified in Figure **2.3.**

The reason for this may be that when increased length is perceived, it will be perceived as stress. However, when increases in length are not perceived, there will be an initial bias, similar to the control condition.

This pattern of results would be explained if we assume that the final long syllable is not perceived as long at the same duration value as other positions are perceived as long. This is in accordance with other researchers-Lindblom (1968), Oller (1973), Lehiste (1973), Klatt (1975), and Klatt and Cooper (1975), who found that listeners expect the final syllable to be longer than the others. Because of this expectation, the final syllable is not perceived as long, until it is very long (relative to the other vowels in the word). Therefore, when length is not perceived (in the shorter conditions), overall responses parallel the control condition.

In summary, we have seen that if English listeners perceive any position as longer, they report it as stressed; if they don't, they report position number one. This initial bias interacts with listeners expectation of final lengthening. That is, while there is a tendency to perceive length accurately at other positions at equivalent durations, that length is not perceived in final positions make those sequences perceptually equivalent to the controls; thus, stress is perceived as initial (see Table 3, for example).

The final lengthening expectation in English explains why there are fewer stress judgments in final position, it does not explain why there are more in first. We are left therefore with one remaining question. Why is it that listeners perceive an initial stress when they don't perceive a length difference? The result is a rather striking one and has been evidenced in other experiments as well (Lehiste 1975, 1976, Smith 1978) although not discussed in these terms.[5]

Given that we know English listeners may impose a rhythmic structure on an incoming speech signal (Nakatani and Schaffer 1976) and that when such a structure is imposed on a group, the initial item is perceived as accented (Woodrow 1951, Fraise 1963) it follows that when processing incoming speech signals which don't provide the necessary acoustic information for predicting segmentation- we impose a rhythmic structure that will. In short, this proposal is suggesting that rhythm is used as an organizing principle in the perception of English and it is for this reason that the first syllable is perceived as longer, louder, or accented relative to the other syllables in the word. I do not want to imply however, that this phenomena is limited to English or that a language with a fixed stress rule can not simultaneously have a fixed rhythmic pattern- rather that rhythmic principles are instantiated in the phonology of a language in specific ways and that they are not only integrated at the phonological level (Allen, 1975) but at the syntactic level (which won't be discussed here) as well.

Yet another reason for the initial stress bias in English is the Germanic stress rule which would lead listeners to expect stress on the first unreduced vowel of a sequence. In the experimental stimuli the vowels all had the same quality and none of them could be considered as more reduced than any of the others. This would lead listeners to expect an initial stress.

In summary, several factors contribute to the initial stress bias:

1) The Relative Timing Principle (i.e. listeners expect the stressed syllable to be longer than any other in the word).

2) Final lengthening expectation (listeners expect the final syllable to be longer).

3) The tendency to organize rhythm groups with initial stresses.

4) The Germanic stress rule (listeners expect initial stress).

This explanation is contrary to that which has been proposed by Chuang and Wang (1976, 1977), and Wang, Lehiste, Chuang and Darnovsky (1976). For example it is argued by Chuang and Wang (1977) that the first of two stimuli is perceived as lower in pitch, louder in instensity, and longer in duration due to a psychophysical time order error (TOE). If this were the case however, strictly speaking, all languages should have initial biases. We have already seen two counterexamples in the present experiement, e.g. K'ekchi and Spanish. It might plausibly be argued that certain language particular factors override the TOE to some extent. However, upon examining the independent contribution of duration to the TOE in Chuangs' data we find that D_{pse}=7.4 ms., which is lower than the difference limen for duration. Given this, it seems unlikely that the bias reported for English by Chuang and Wang (1977) is a psychophysical TOE, and more likely, an artifact of the language investigated.

Footnotes

1) This chapter is an extended version of a paper presented at the 95th meeting of the Acoustical Society of America, "A cross-linguistic study on the perception of stress", by A. Berinstein, Providence, R. I. May 1978.

2) For excellent surveys on the relationship of duration and stress the reader is referred to Klatt (1976), for perception and production of stress, to Lehiste (1970).

3) Actually a fourth language, Cakchiquel, was involved in the experiment which will be reported on in Chapter 4. The analysis of variance included the presence of a four-level language variable which is reflected in the degree of freedom specification of F-ratios involving language as an effect or interaction.

4) More generally the use of duration is specific to the phonemic length contrast in the language. It is for this reason (I would argue) that changes in Fo do not cue a phonemic length distinction in Swedish (Rosen 1977b).

5) Rosen (1977a) has noted a similar position bias which interacts with the effect of fundamental frequency patterns on perceived duration.

CHAPTER THREE

A hierarchy of acoustic correlates to stress[1]

3.0 Introduction

It is suggested in Chapter 2 that duration may not be a significant correlate of stress in languages in which vowel length is phonemic. As shown above, and in Berinstein (1978a), the perceptual data on K'ekchi support this view, since the systematic variation of syllable length in K'ekchi did not affect perceptual judgements of stress. Since such variations in English resulted in highly significant differences in stress perception judgments, we have some further support for the language dependency of acoustic cues of stress.

One may then hypothesize that speakers use the other acoustic characteristics such as length both for producing of stresses versus unstressed syllables as well as perception. The mistakes of length made by foreign learners (Jones 1969:243) demonstrate that speakers do this, and these errors are consistent with speakers' language particular stress patterns.

It is important, however, to investigate further the relationship between the production and perception of stress. As a first step in this effort it is necessary to determine what the acoustic correlates of stress are in a language where the posited universal hierarchy seems to fail. A second experiment, to be discussed in this chapter, was designed for this purpose; to determine the acoustic correlates of stress in K'ekchi, and to determine the relative importance of these acoustic parameters (rank-ordered) in signalling the distinction between stressed and unstressed syllables in the language.

3.1 Experiment II

3.2 Method

Subjects

Ten male adults, all of whom were literate, bilingual speakers of K'ekchi and Spanish, whose average age was 40 years, and who were all experienced linguistics consultants served as voluntary subjects. They all used K'ekchi on a daily basis in their personal and professional life. Subjects were paid for their time.

Stimulus Material

The ten subjects recorded words in K'ekchi controlled for stress, vowel quality, phonemic vowel length and number of syllables in the word. The words were embedded in a K'ekchi phrase: [li a:tin a'in ______ naštye] "this word________means". The list of words included short and long versions of the vowels /i,e,a,o,u/ in both stressed and unstressed syllables.

One difficulty presented itself since K'ekchi has fixed final stress, and so it is impossible to obtain minimal stress pairs such as the English pairs 'contrast (N)/con'trast (V) or 'pervert (N)/ per'vert (V). Therefore the stimuli were constructed to allow comparisons of stressed and unstressed vowels under two different but unfortunately non-optimal conditions. A between word condition permitted a comparison of the stressed vowel of a monosyllabic word with an unstressed vowel of the same root with a stressed suffix added, as in: [kút] 'throw (it).' [Imperative] versus [kut+úk] 'to throw' [stem + infinitive suffix].

This comparison is non-optimal because, as is shown in the literature, vowel duration decreases with number of syllables (Elert, 1964: 137 for Swedish ; Sharf, 1962: 26 for English, Meyer and Gombocz, 1909: 138 for Hungarian, and Menzareth and Oleza 1928: 68 for Spanish). However, it is amply documented that shortening due to polysyllabicity is much less pronounced for unstressed syllables preceding a stress than for unstressed syllables following a stress within the same word (Nooteboom 1972, 1973, Lindblom & Rapp 1973).

The stressed and unstressed syllables in the between word comparison are perfectly matched in terms of their segmental structure. However, it is conceivable that adding a suffix may have led to resyllabification of the original syllable final consonant, so that in the between word comparison open and closed syllables are potentially confounded. Delattre (1966) has demonstrated that appreciable differences exist in French, German, English and Spanish between open and closed syllables.

A within word comparison is one in which the stressed vowel of a disyllabic word is compared with the unstressed vowel in the same word, as [kutúk] ; where the underlined vowels are compared. The vowel in the suffix was chosen to be as much like the stem vowel as possible. It was especially difficult to obtain pairs of the same long vowel in the within word condition, in fact, only one completely suitable word was found in this category [kʉ:tú:ŋk].

The within word condition is also non-optimal because a comparison is made between syllables in different positions in the word. We know from studies (Oller 1973, Klatt 1975) on other languages showing word final lengthening that there is at least a potential effect of syllable position. It is an inevitable consequence of fixed final stress that the potential role of stress and position cannot be disentangled. It is for this reason that linguists have often referred to French as a language without phonemic stress (Malmberg,1966). It would, however, be difficult to argue that the K'ekchi data only reflect a word-final position effect, as both fundamental frequency and vowel intensity were found to vary as well in word final (i.e. stressed) position.

A set of twenty words was constructed by variation of vowel quality (/a,e,i,o,u/), stress (+stress, -stress) and phonemic length (long/short). Table 1 lists the twenty words, their phonetic transcription and glosses. Stress is indicated by an accent mark on the vowel. The words (in the first column) are written in K'ekchi orthography where an underline is used to signify phonemic length.

The sentences were typed out in standard K'ekchi orthography- which is strongly based on Spanish orthography- in a list on one sheet of paper, with filler sentences at the beginning and end of the list.

Procedure

The subjects received non-standardized oral instructions in Spanish to read aloud the list of stimulus sentences. The responses were recorded in a sound treated room on a Sony TC-142 2-track cassette recorder through a Sony ECM 250 condenser microphone. Subjects were admonished to observe a fixed distance from the microphone. A practice session in which the entire list was read, preceded the experiment. After familiarization with the test items and the use of the microphone, the list of sentences was recorded three times.

	Words	Phonetic transcriptions	Gloss
1a	cut	[kút]	throw it
b	cutuc	[kutúk]	to throw
1'a	cut	[kú:t]	prop, support
b	cutunc	[ku:tú:ŋk]	to prop
2a	tik	[tíq]	add it
b	tikoc	[tiqók]	to add
2'a	tic	[tí:k]	straight
b	ticoc'	[ti:kók']	to straighten
3a	quet	[két]	hit it
b	quetoc	[ketók]	to hit
3'a	tep	[té:p]	place
b	tepal	[tepál]	district
4a	cot	[kót]	rope it
b	cotoc	[kotók]	to rope
4'a	pop	[pó:p]	mat
b	popol	[po:pól]	authorities
5a	cak	[káq]	red
b	cakcak	[kaqkáq]	very red
5'a	cak	[ká:q]	thunder
b	cakeb	[ka:qɛ́b]	explosions

Table 1. K'ekchi test items (column 1) with corresponding phonetic transcription (column 2) and English translation (column 3).

3.3 Analysis

It was decided only to incorporate the utterances obtained in the final reading of the list, so that the data are based on 200 word tokens (10 Ss * 20 word types).

Each of the responses was copied onto a fulltrack Voice Identification VII Spectograph and both wide and narrow band spectographs were made. Filter settings were 45 and 300 Hz for narrow and wide band spectographs respectively, and the full vertical scale comprised 8kHz. Intensity traces of the 0 -3.5 kHz portion of the spectrum were included on the wide band spectrograms.

Vowel durations were measured. Onset and offset of glottal excitation of the vowels (which were always between voiceless consonants) were expressed in terms of integral millimeters and later converted to msec. by computer. Peak intensity was measured with an attempted precision of 1 dB. Fundamental frequency was measured by tracing the 10th harmonic in the narrow band spectograms. Three Fo values were measured.

1) the peak Fo

2) the lowest Fo value preceding the peak value

3) the lowest Fo value following the peak.

Statistical analyses were performed on vowel durations (after conversion to msec), peak intensities, peak Fo, and Fo ratio.

The Fo ratio parameter was introduced to capture the potential role of fundamental frequency dynamics, i.e. rises and falls, complementary to the peak frequency. For reasons of simplicity only, rises and falls but no complex Fo contours were accepted in the data. Whenever a syllable contained a rise-fall, which almost exclusively happened with long vowels - the fall part was ignored. The ratio was then computed by dividing the peak Fo, (Fo max) by the lowest Fo value (Fo min) within the vowel. When Fo max followed Fo min the contour was a rise, and was given a positive sign, when Fo min followed Fo max, the contour was a fall, and received the negative sign. Finally, all positive values were diminished by 1, and all negative values were incremented by 1, so that a level Fo contour was expressed by 0, a rise of one octave by +1, and a one octave fall by -1.

Interestingly, this definition of Fo intervals turned out to supply a better predictor of intended stress than using intervals expressed in terms of semitones, i.e. log transformed to fit a musical scale by multiplying the ratio of Fo max over Fo min by $12\sqrt{2}$. For this reason no semitone conversion will be reported here.

The data were punched on cards and further analyzed by computer, using the Statistical Package for the Social Sciences (SPSS Nie et. al. 1975).

3.4 Results

Four acoustic parameters used by speakers in producing the stress/ nonstress distinction were measured. **Fo ratio**, **peak Fo**, **peak vowel intensity** and **duration**.

Fo Ratio and Peak Fo

Figure 3.1 presents the mean Fo ratios ± 1 standard deviation pooled across the five different vowels, but separated out for long and short vowels in three conditions: (1) (stressed) monosyllables, (2) the unstressed root in a disyllable and (3) in the stressed position as suffix in a disyllable.

Most importantly the unstressed syllables (condition 2) are characterized by typically level or slightly falling Fo contours, with relatively little variance. We see also that the mean Fo ratio is equivalent for both short and long unstressed vowels.

The stressed syllables (conditions 1 and 3) are typically correlated with rising Fo patterns, although there is some overlap for short stressed suffixes and unstressed roots.

A three-way analysis of variance on Fo ratios with vowel length, stress condition and vowel quality as factors reveals that the differences due to stress are highly significant $F(2,270) = 10.8$, $p < .001$. There is also a significant difference between stressed short versus stressed long vowels in both monosyllabic and disyllabic words $F(1,270) = 5.8$, $p = .017$. This difference shows up in the Fo ratio because in a short vowel there is less time to reach the target value so the excursion is smaller. Notice that this difference does not show up between unstressed short and long vowels. This is because unstressed vowels tend to be fairly level.

We can draw the following conclusions:

1) stressed vowels distinct from unstressed vowels can be characterized by a rise in Fo
2) long stressed vowels distinct from short stressed **vowels have considerably** steeper Fo rises and
3) long unstressed vowels are not distinct from short unstressed vowels in terms of their mean Fo ratio.

The point here is that stressed and unstressed vowels are highly identifiable on the basis of Fo dynamics alone, i.e. independent of phonemic length or word length.

Figure 3.2 analogous to the previous graph, plots the peak **Fo as a** function of stress condition and vowel length. Although there is overlap between stressed and unstressed vowels, there is considerable difference in terms of peak Fo between stressed vowels (158 Hz) for conditions 1 and 3 and unstressed vowels (125 Hz) condition 2. Notice that there are hardly any differences between long and short vowels within a given condition. Also, the peak Fo values do not differ for stressed vowels in monosyllabic **and disyllabic words.**

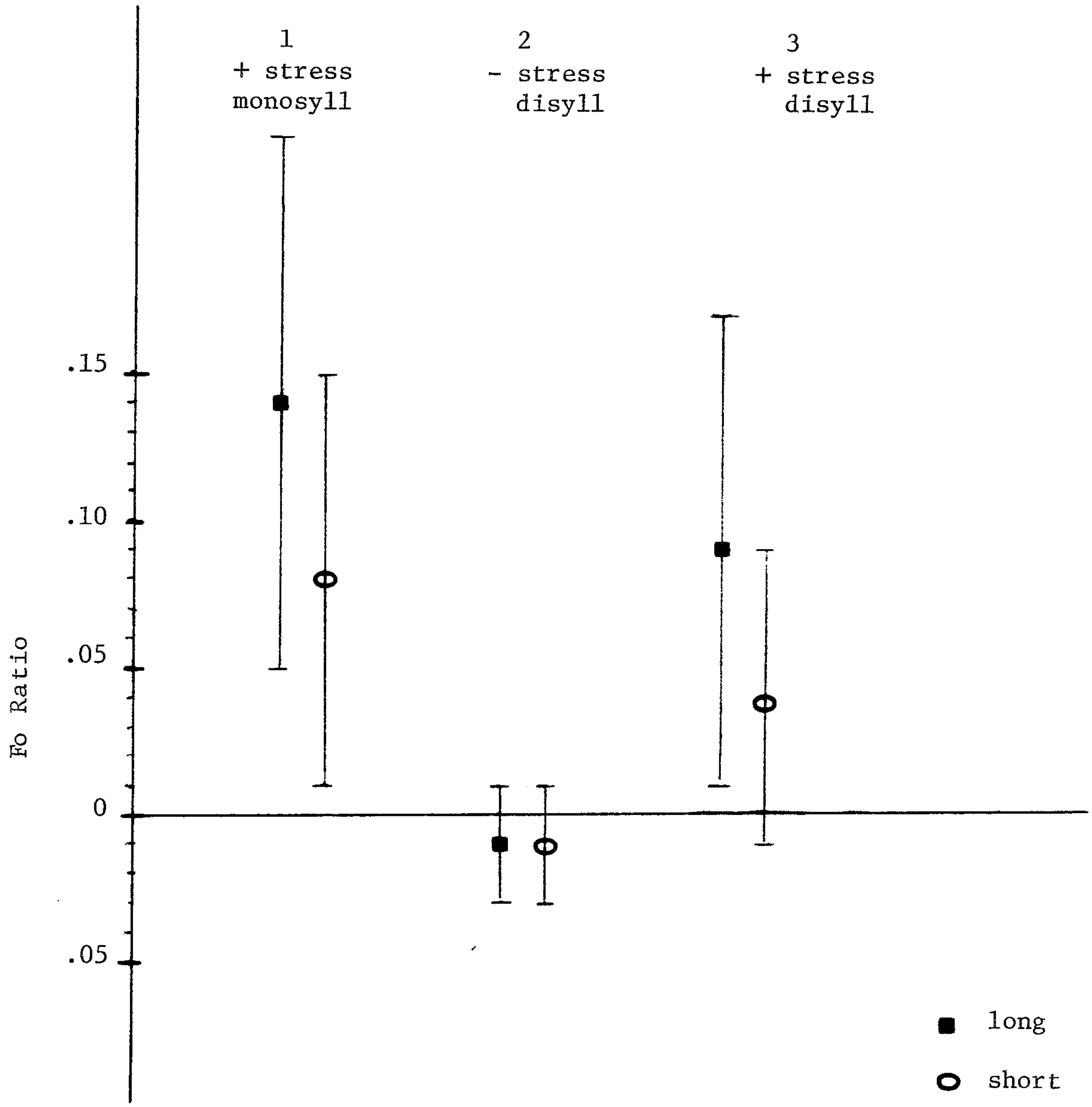

Figure 3.1. F_o ratio as a function of stress condition and vowel length.

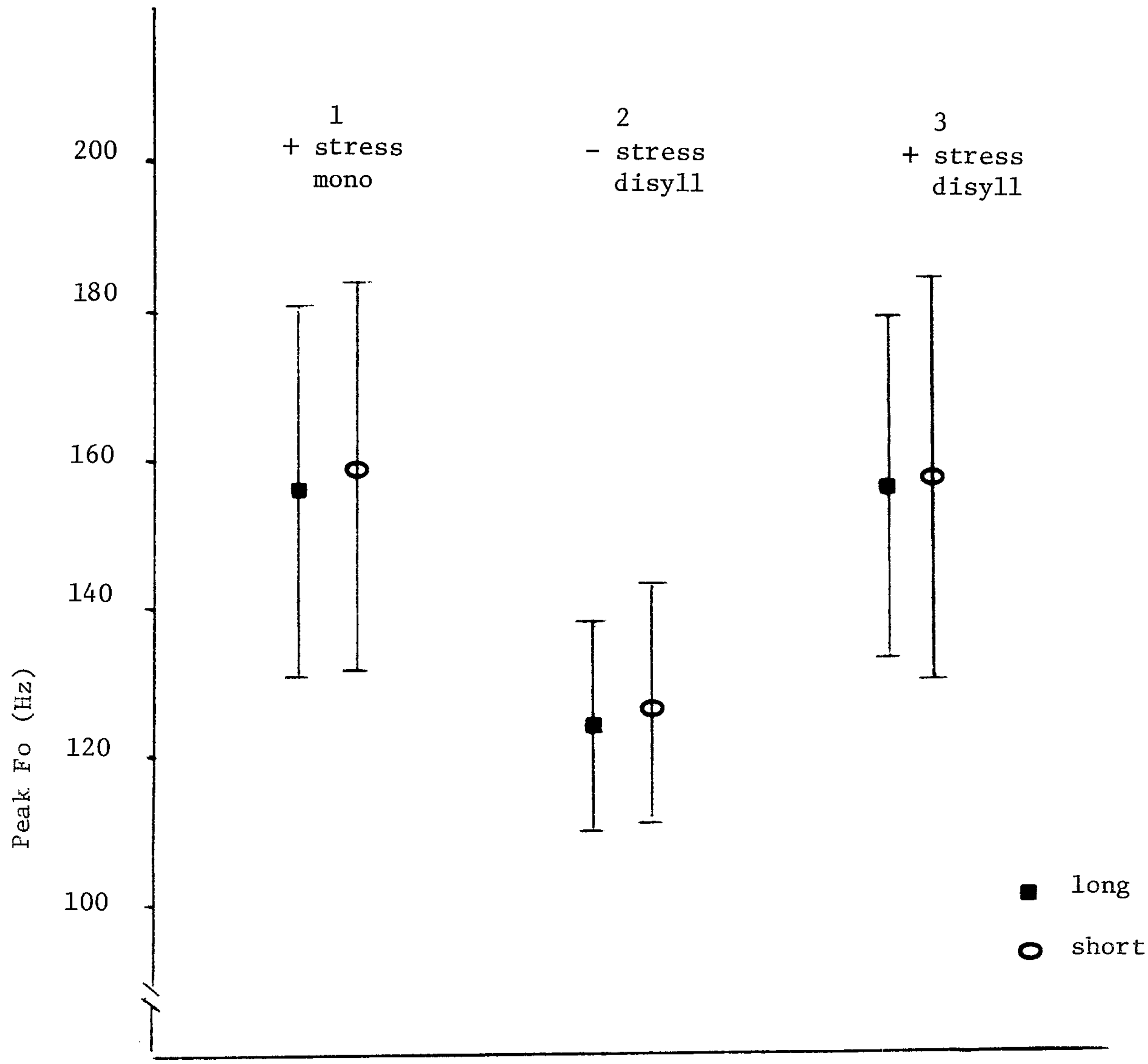

Figure 3.2 Peak Fo as a function of stress condition and vowel length.

We can draw the following conclusions:

1) Peak Fo values for stressed vowels (in monosyllabic and disyllabic words) are significantly higher than those of unstressed vowels $F(2,270) = 61.2$, $p < .001$.
2) Peak Fo values for long versus short vowels do not significantly differ from each other $F(1,270)$, $p < 1$, n.s.

The point here is that pitch height is conditioned by stress, not number of syllables or phonemic length.

Peak Vowel Intensity

Peak intensity differences as correlates of stress and vowel length are presented in Figure 3.3

The average intensity difference between stressed and unstressed vowels is approximately 4 db. A secondary effect is provided by phonemic length in that short vowels on the average, are 1 db stronger than long vowels $F(1, 269) = 10.3$, $p = .001$. Stress and length do not interact, rather, they are purely additive $F(2, 269) < 1$, n.s., since it is predictable that short stressed vowels will have the greatest intensity.

In conclusion we see that peak vowel intensity of stressed vowels is similar in mono and disyllabic words, but these values differ significantly from the peak values for unstressed vowels $F(2,269) = 105.1$, $p < .001$. The point again is that the contribution of stress to intensity is independent of phonemic length or word type.

Duration

Figure 3.4 represents vowel duration in msec as a function of stress condition and phonemic vowel length. Not too surprisingly, there is a large durational difference between phonemically long versus phonemically short vowels. If increased duration is highly correlated with stress, stressed syllables should be longer than unstressed syllables just as Fo ratios, peak Fo and peak intensities are greater for stressed syllables than for unstressed syllables regardless of word type or phonemic length. Typically, short vowels have durations on the order of 75 msec., long vowels are approximately 190 msec. However, the question that is at issue is the extent to which duration plays a role in stress distinctions.

Analysis of variance indicates that the phonemic length opposition has by far the most significant effect on duration $F(1,269)=877.1$, $p < .001$, explaining 56% of the variance. According to the analysis of variance, stress condition also influences duration $F(2,269) = 120.1$ $p < .001$ explaining 15% of the variance. Inspection of the graph reveals that this is not due to systematic duration differences between the two stress conditions on the one hand and the unstressed condition on the other; rather, the stress effect is attributable for the major part to the highly deviant mean of the long stressed syllable in disyllabic words.

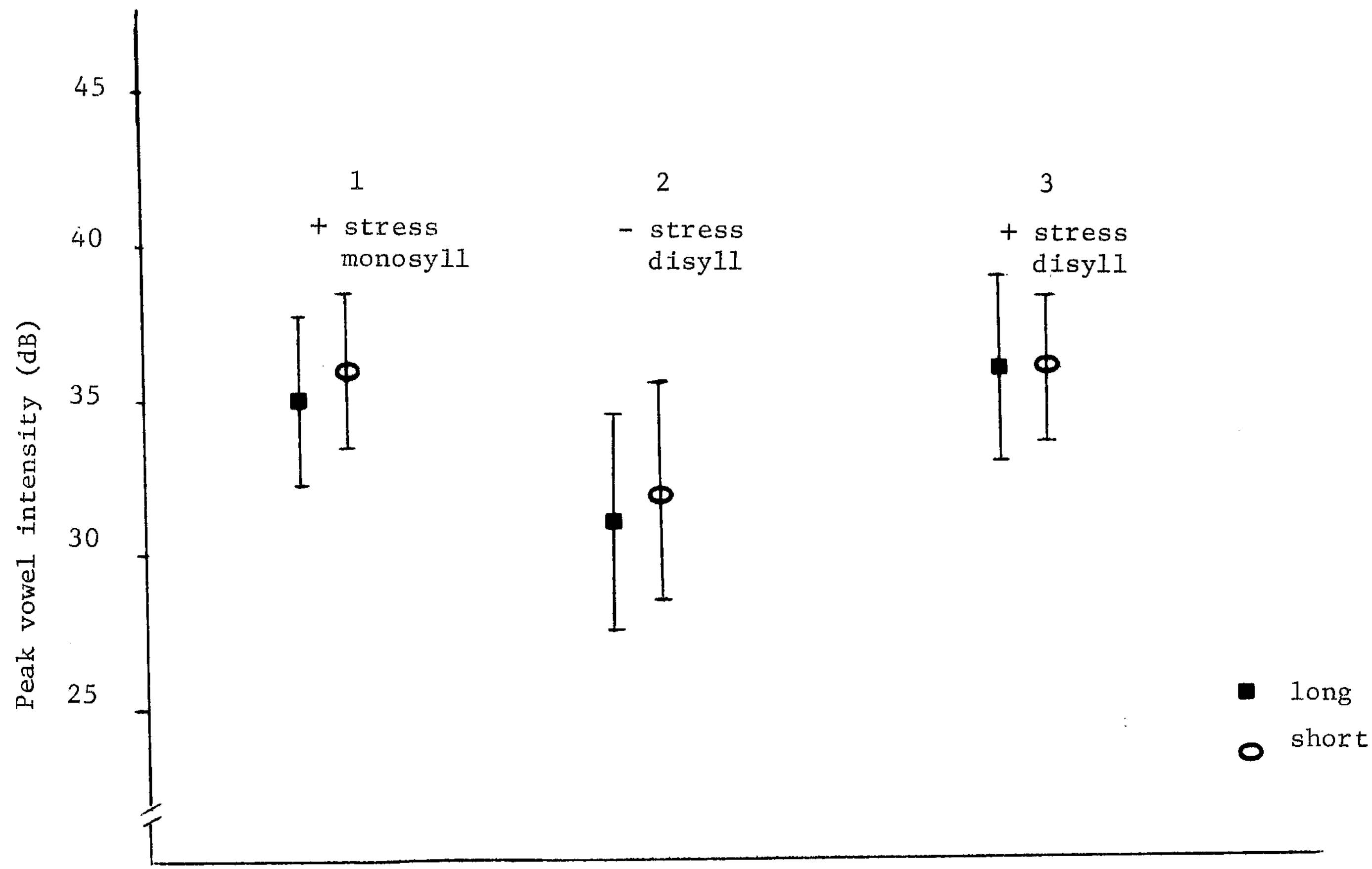

Figure 3.3. Mean peak intensities for stressed and unstressed vowels in monosyllables and disyllables for long and short vowels.

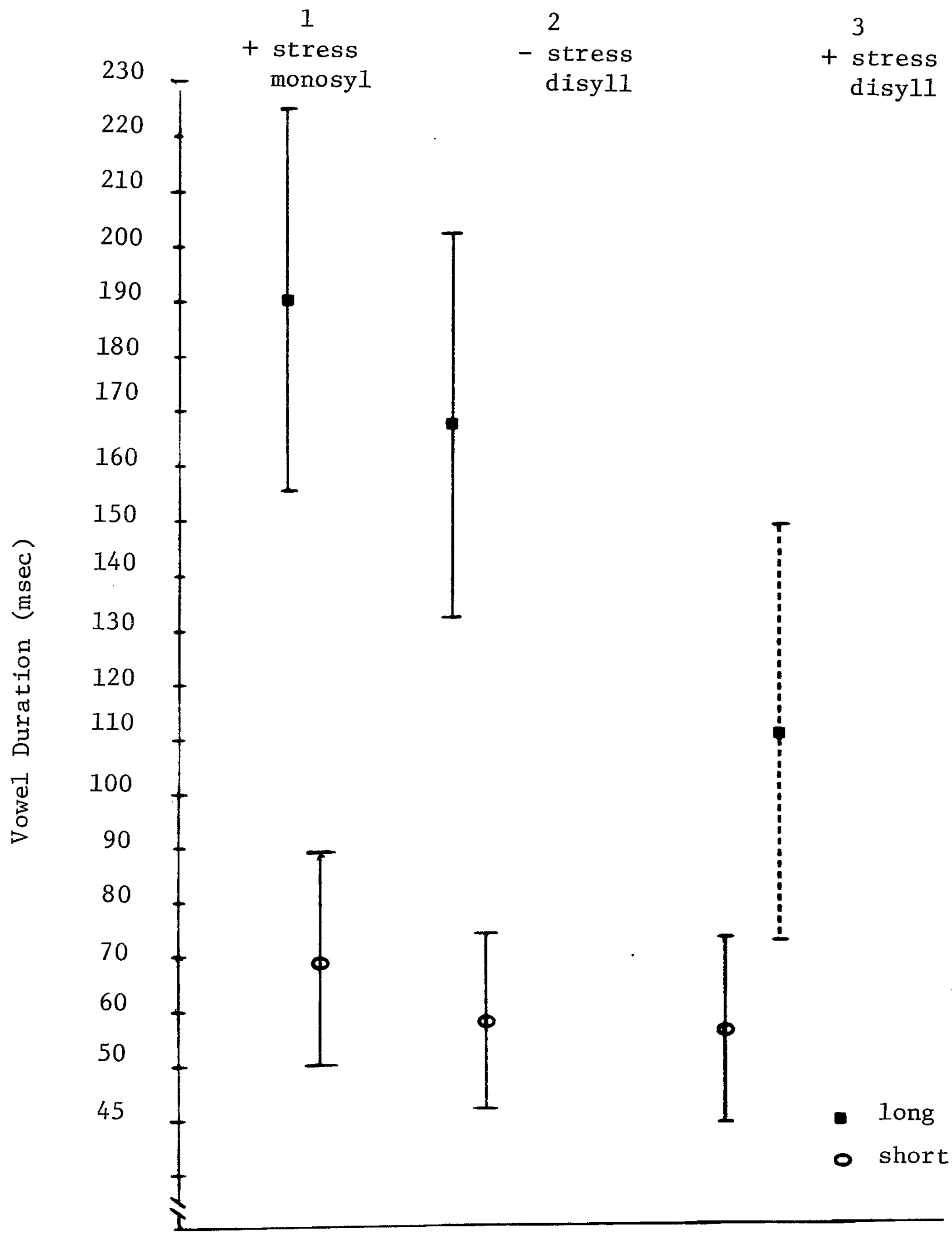

Figure 3.4. Mean duration (msec) for stressed and unstressed vowels in monosyllables and disyllables for long and short vowels.

As mentioned earlier, it was not possible to have a simple length contrast in the within-word stress condition (the comparisons of 2 to 3 on the graph). The stressed vowels in the final syllable in disyllabic words in K'ekchi were practically always short. For this reason a much lower mean duration is indicated in the graph for the long vowels in stress condition 3. The dotted line in Figure 3.4 is representative of the duration of the stressed final vowel of ku:tú:ŋk, only. It is unclear, in fact, whether or not ú: in this word is phonemically long or phonetically lengthened due to the sonorant cluster; although, probably the latter. The point, however, is the same: stressed vowels need not be and in general, are not longer than unstressed vowels. This is obviously the case in words with non-final phonemically long vowels. Thus, the average duration of the unstressed long vowel is 57 msec. longer than the stressed "long" vowel of a disyllable. Moreover, in the short vowel-within word conditions we see that there is no duration difference due to stress. In fact, the average duration of the unstressed short vowels (68 msec) is slightly longer than the stressed short vowels (66 msec). T-tests indicate that this difference is statistically insignificant. $T(98) = .56$, n.s. Were duration correlated with stress, we would predict precisely the opposite (e.g. that stressed short vowels would be significantly longer than unstressed short vowels). From this we can only conclude that differences in duration within words are due to the phonemic length distinction, not to stress.

Further tests for a-posteriori contrasts reveal that (short) stress condition 1 does differ systematically from (short) stress conditions 2 and 3, as evidenced by a one-way analysis of variance $F(2,147) = 8.7$, $p < .001$. Conditions 2 and 3, however, do not differ from each other, as determined by the Newman Keuls procedure at a .05 level of significance. It is clear therefore that the significance of the analysis of variance is due to the fact that stress condition 1 is different from the others. From this we can only conclude that the differences in duration are due in part to word type. Notice for example, that the stressed vowels of the monosyllable (stress condition #1) are on the average 12 msec. longer than those of the disyllables (stress conditions 2 and 3). The point here is that duration is poorly correlated with stress; rather, it is dependent on phonemic length and word type.

3.5 Discussion

So far we have seen that in K'ekchi certain acoustic parameters-Fo ratio, peak Fo, and peak intensity-are correlated well with stress; whereas, duration is not. Also, we saw in the perception experiment (Chapter 2) that duration was not a good cue for stress. We may hypothesize now that the reason duration is not important to the stress/nonstress distinction is because it is used to signal a phonemic length contrast in the language. Recall that the proposed universal unmarked hierarchy of perceptual correlates of stress was primarily indicated by a change in fundamental frequency, secondarily by duration and finally by intensity, unless a language had a phonemic contrast of tone or length in which case the subjective cue correlating to that contrast would be superseded by the other cues.

Given this, we should predict that in K'ekchi stress will primarily be indicated by a change of fundamental frequency, secondarily by intensity and lastly by duration, since K'ekchi has a phonemic length contrast.

It might be of interest to look at a comparison of the relation of syllable duration as a function of stress and word position in three languages. Figure 3.5 presents these data from English and Spanish (Delattre, 1966), and K'ekchi. It is obvious that the three languages make durational distinctions but in K'ekchi such differences are primarily conditioned by phonemic length contrasts rather than stress or word position. While this does not support the predicted hierarchy by itself, it seems clear that duration cannot be the second most important cue for stress in K'ekchi when it is not important in production at all.

3.6 Establishing a Hierarchy of Cues for Stress

As has been abundantly documented in the literature, the stress/nonstress distinction may be cued by a variety of acoustic phenomena. It has also been documented that some cue or cues may be more influential in signalling the distinction than certain others. Yet, the procedures by which the differences in cue value were derived, have so far been crude, and often unrevealing. In the classic studies by Fry (1955, 1958) the conclusion that duration is more important than amplitude, was arrived at on the basis of the observation that in a perceptual experiment in which intensities and duration were varied, durational differences could compensate a wider range of intensity differences than vice versa.

In the present production experiment, we have isolated a number of acoustic phenomena that vary in a more or less systematic fashion in accordance with the stressed/nonstressed distinction as intended by the speakers. It is apparent in the graphs (Figures 3:1,2,3,4) that some acoustic differences are more systematically related to stress than others. This could informally be observed in the degree of overlap in the measurements on intended stresses versus nonstressed vowels. The degree of overlap can be, and has been, further expressed numerically through analysis of variance. The percentage of the variance explained by the factor stress may serve as an index of the potential effectivity of the acoustic phenomenon in cueing stress.

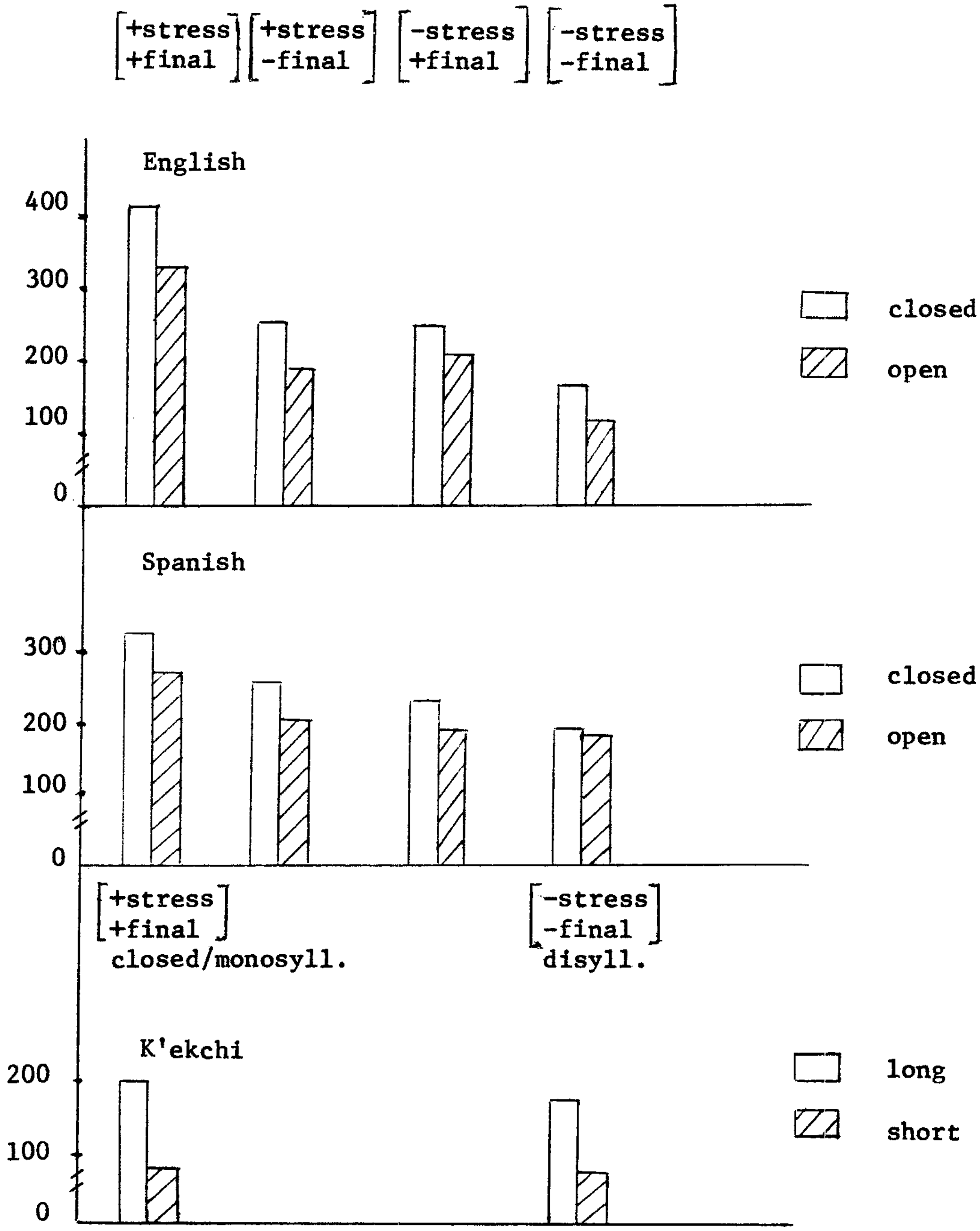

*Figure 3.5. Duration (msec) as a function of stress(+stress/-stress) and position (+final/-final) in English, Spanish and K'ekchi.

*English and Spanish data adapted from Delattre (1966).

However, this technique does not allow for the possibility that some acoustic differences may occur in parallel, or statistically, that due to high intercorrelations the independent contribution of a particular acoustic parameter as a cue to the stressed/nonstressed distinction may be low. A more adequate statistical technique to express the extent to which each one of a number of acoustic parameters contributes independently to a binary distinction is provided by multiple regression analysis which allows for a straight-forward numerical prediction on the basis of the raw acoustic parameter units as to which of the two relevant stress categories is signalled.

Multiple stepwise regression analyses were performed on the data as a whole, as well as separately for the within and between word conditions relating stress to peak intensity, vowel duration, peak Fo and Fo ratio. In the analysis the acoustic parameters, or predictor variables were entered in terms of their original units. i.e. dB, Hz, and msec. Intended stress was coded Ø for nonstress and 1 for stress. The resulting regression formula yields a value typically, but not necessarily, between Ø and 1; these values can be interpreted as the predicted degree of stress on a particular vowel. To the extent that the formula yields predictions that are not exactly 0 for unstressed vowels and 1 for stressed vowels, we are left with unexplained or residual variance. Consequently, the percentage of the variance explained by the formula is an important criterion for the adequacy of the regression solution. In our case, however, a more meaningful test for the adequacy of the regression formula is to establish how often it enables us to correctly classify a vowel as stressed or unstressed on the basis of the acoustic predictor variables. In this test the predicted degree of stress may vary from 0 to 1, but for perfect discriminability there should be no overlap of values for predicted stress on the one hand, and no stress on the other.

When, on the basis of such tests, one is convinced of the adequacy of a regression solution, one can then rank-order the acoustic predictors in terms of how much each contributes to the prediction of stress. This rank-ordering can be done, 1) on the basis of the raw correlations between stress/no stress and each of the predictors, which for reasons given earlier, may be a deceptive criterion, and 2) by comparing the percentage of the total variance independently accounted for by each variable. In this latter case, the predictor with the highest raw correlation to stress will explain more variance than any other predictor, the next variable which need not necessarily have been the second largest raw correlation coefficient, will then explain more additional variance than any of the remaining predictors and so forth. This solution provides also a way of explicitly testing the proposed hierarchy of acoustic cues in K'ekchi. We will predict that Fo ratio will explain the most variance and duration the least.

Tables 3.2, 3.3, and 3.4 contain the raw correlation matrices for the between and within word subsets separated out for phonemic length. It should be pointed out that the between word comparison is completely balanced for long and short vowels. (Tables 3.2 and 3.3 respectively). However, for reasons mentioned in Section 3.2, it would be unfair to include long vowels in the within word comparison, as they are not properly balanced by long vowels having the same quality.

	FR	I	F	D	STRESS
FR		0.373	0.636	0.294	0.745
I			0.445	0.593	0.513
F				0.431	0.628
D					0.317

Table 3.2. Raw correlation matrix of four predictor variables: Frequency ratio (FR), Peak vowel Intensity (I), Peak Fo (F) and Duration (D) for the long vowel between word stress condition.

	FR	I	F	D	STRESS
FR		0.335	0.561	0.314	0.620
I			0.382	0.360	0.570
F				0.495	0.584
D					0.309

Table 3.3. Raw correlation matrix of four predictor variables: Frequency ratio (FR), Peak vowel Intensity (I), Peak Fo (F) and Duration (D) for the short vowel between word stress condition.

	FR	I	F	D	STRESS
FR		0.353	0.615	0.062	0.544
I			0.419	0.206	0.636
F				0.301	0.558
D					-0.058

Table 3.4 Raw correlation matrix of four predictor variables: Frequency ratio (FR), Peak vowel Intensity (I), Peak Fo (F), and Duration (D), for the short vowel within stress condition.

For this reason, only short vowels have been included in the within word comparisons (Table 3.4).

It will be observed that the intensity and frequency parameters correlate much better with stress than duration does, regardless of the data sets included in the analysis. However, some of these parameters particularly frequency and frequency ratio have high intercorrelations, so that multiple regression is needed to partial out the shared explained variance due to interdependence of predictors.

The following formulas capture the optimal solutions for predicting the degree of stress on the basis of the four acoustic predictors for the between-long data set (1), between-short (2), and within-short (3), where FR refers to the frequency ratio, I to peak intensity (dB), F to peak frequency (Hz), and D to duration (msec). When a particular predictor is not mentioned in a formula, its contribution is nonsignificant, i.e. it can be omitted from the formula without effecting the outcome of the predictions.

(1) $P(\text{stress}) = 2.623\ FR + .038\ I + .004\ F - .001\ D - 1.219$

(2) $P(\text{stress}) \doteq 2.431\ FR + .050\ I + .005\ F - 1.939$

(3) $P(\text{stress}) = 3.859\ FR + .066\ I + .004\ F - .008\ D - 1.798$

Figure 3.6 graphically represents the independent contributions of each of the four predictors to the stress/no stress distinction. It contains a plot of the percentage of variance additionally explained by each predictor variable.

It turns out that the solution is fairly stable across the data subsets. Consistent with our prediction, duration has little or no independent contribution, as could be expected on the basis of the low simple correlation with stress. Interestingly, peak frequency is a less effective predictor than the frequency ratio , which indicates that the dynamics of the Fo (rises, falls) are more important correlates of stress in K'ekchi, than the peak frequency. Due to high intercorrelations of the two frequency variables, the independent contribution of peak Fo is extremely low, once the ratio parameter is entered in the regression analysis, which explains why it may at one time contribute even less than the duration variable. Frequency ratio and intensity provide the best predictors of stress, depending on the particular data subset on which the analysis is performed. The influence of the conditions will be discussed in detail in Section 3.7. Again, due to sizeable intercorrelation of Fo movement and intensity-both of which have been claimed as involuntary acoustic correlates of increased subglottal airpressure in the production of a stressed syllable- (cf. Lehiste 1970 and references there), the independent contribution of either parameter declines sharply vis a vis its raw correlation coefficient (simple r), when the other has been entered in the analysis at an earlier stage.

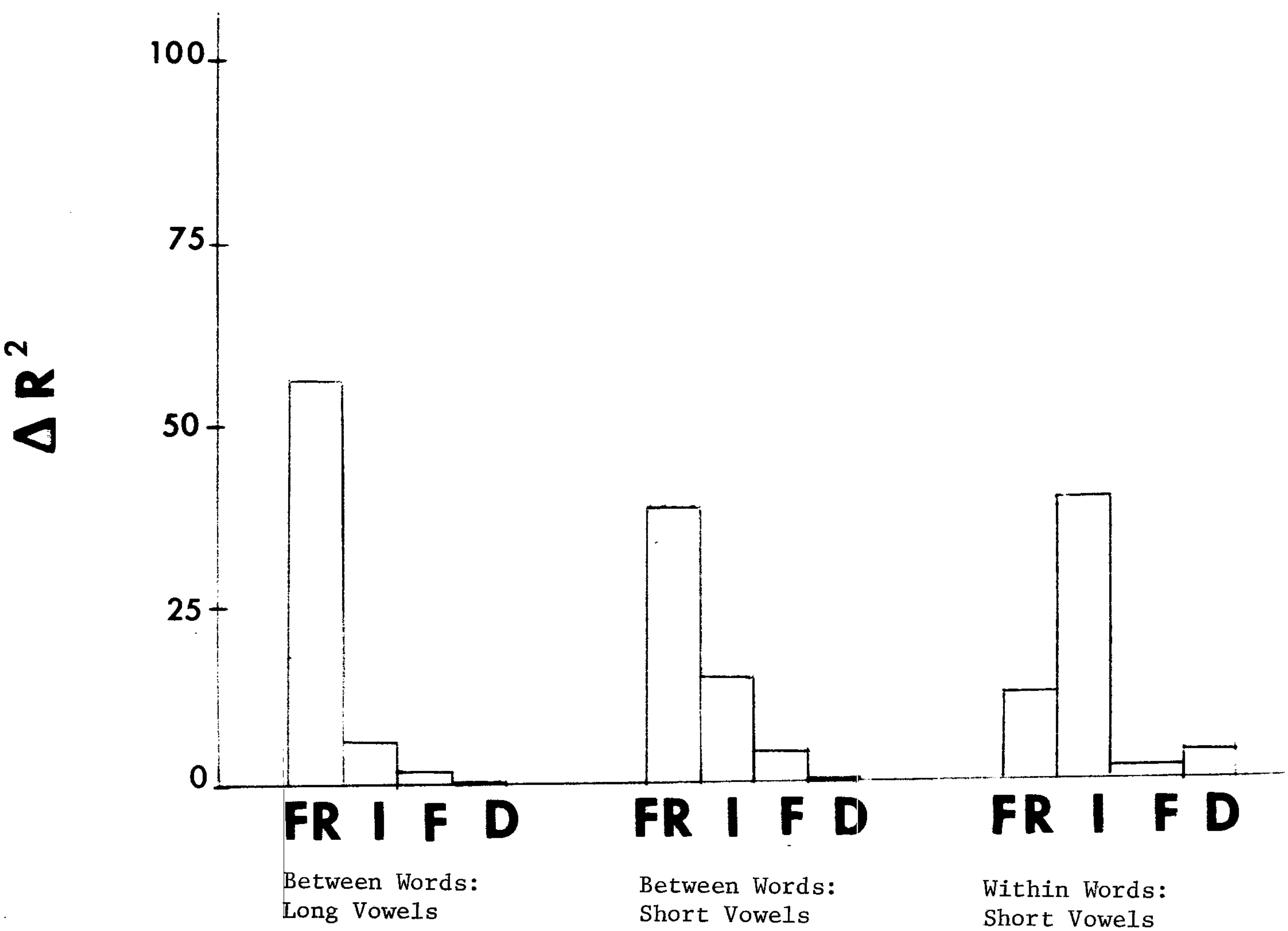

Figure 3.6. Percentage of the variance explained by each predictor variable to stress, in K'ekchi, for three conditions where FR is frequency ratio (Hz), I is intensity (dB),F is peak Fo (Hz) and D is duration (msec.).

The regression formulae (1) -(3) account for 57 to 64% of the variance. The additional explanatory power of durational differences is marginal (.01%) compared with intensity or frequency.

Let us, finally, attempt to establish the adequacy of the solution by examining the formulas' efficiency in separating the intended stresses from the unstressed syllables. Inspection of the actual data when using optimal boundaries forces us to shift the boundary somewhat towards the 0-extreme for the between word conditions, where the criterion used was .460. The number of incorrect classifications is very small, as is illustrated in Table 3.5.

Inspection of error percentages in table 3.5 shows that the regression formulas permit excellent classification of stressed and unstressed syllables, with an error rate ranging from 12 to as low as 2%. This is a level of adequacy comparable to results reported in the literature on automatic stress detection in English (Lea 1977, Sargent 1975). It should be pointed out, however, that the cirterion for stress in these studies was listeners' perception rather than speaker intention. In the case of K'ekchi it would seem unnecessary to subject the spoken material to a stress judgment experiment, as any naturally produced polysyllabic word will have stress on the ultimate syllable.

3.7 Discussion and Conclusions

As predicted by the perceptual results given in Chapter 2, duration is irrelevant in the production of stress in K'ekchi. Frequency ratio and amplitude were found to be the best acoustic cues. I would like to argue that FR is generally the better predictor on the grounds that 1) it is the best predictor in two out of three data subsets and 2) the predictive power of FR is relatively limited in the short vowel conditions only. Referring to Figure 3.1 recall; (a) that short vowels on average have a less pronounced Fo rise than the corresponding long vowels and (b) that the Fo contours are more level on average for the within word comparisons than for the between word cases. ~~In one condition the short-within word~~, there is considerable overlap of Fo ratios for stressed and unstressed vowels. It is only in that case that intensity differences provide the better stress cue.

In conclusion, the importance of acoustic correlates of stress in K'ekchi is indicated primarily by a change of fundamental frequency. secondarily by increased intensity and finally by duration.[3] A stressed syllable will be indicated by a rise in pitch and greater amplitude relative to an unstressed syllable, with a rise in pitch the stronger of the two cues. The role of duration is negligible in signaling this distinction (accounting for .01% of the variance). These results, taken together with the failure of K'ekchi speakers to respond to duration as a cue for the perception of stress indicate the interrelation of perception and production in this language. Further, this language supports the hypotheses that:

Data Subsets		Regression equation Criterion for Stress	Correct Classification	Stress Predicted No stress	No stress Predicted Stress
Long Vowel between	(1)	$P(str) \geq .460$	98%	2%	0%
Short Vowel between	(2)	$P(str) \geq .460$	88%	8%	4%
Short Vowel within	(3)	$P(str) \geq .500$	95%	2%	3%

Table 3.5 Percentage of correct and incorrect classified stresses predicted by the regression formulae in three conditions defined by phonological environment.

1) Duration has little or no cue value to the stress/no stress distinction in a language in which there are long and short vowels distinguished solely by length.

2) In languages with no phonemic contrasts in tone or length, the "unmarked hierarchy" for cues relevant to the perception of stress will be primarily indicated by a change of Fo, next by duration and lastly by intensity. In languages with such contrasts, the subjective cue correlating to that contrast, will be superceded by the other cues of the hierarchy.

Footnotes:

1 This is an extended version of a paper presented at the 96th Meeting of the Acoustical Society of America, "Acoustic correlates of stress in K'ekchi", by A. Berinstein, Honolulu, Hawaii, Nov. 1978.

2 There are obviously no $\begin{bmatrix}\text{+stress}\\ \text{-final}\end{bmatrix}$ or $\begin{bmatrix}\text{-stress}\\ \text{+final}\end{bmatrix}$ categories in Figure 3.5 for K'ekchi, since, all (and only) final syllables are stressed.

3 It should be noted here that contrary to Bolingers' (1962) claim, change of Fo in K'ekchi which is associated with major syntactic boundaries and all final syllables, is a RISE, not a fall (see Berinstein 1979, "Timing Differences and so-called 'Isochrony'", for details), thus, the pitch baseline in K'ekchi is generally an inclination rather than a declination.

CHAPTER FOUR

A distinction between a "position bias" and a fixed location of stress

4.0 Introduction

In the experiments described above evidence was presented in opposition to the universal hierarchy hypothesis. In addition it was found that in K'ekchi, a language with phonemic length, duration was a poor cue for stress, either production wise or perceptually. It might be the case however, that this is due to the fact that there is final fixed stress in K'ekchi. Should we find a language which also has final fixed stress and no phonemic length, and should duration also be a poor cue, one might then argue that the key factor is the fixed stress position rather than the phonemic length. To distinguish these hypotheses the perception experiment was repeated with speakers of Cakchiquel, a related Mayan language, which has fixed final stress and no phonemic length. The results were then compared to the K'ekchi data.

4.1 Method

The same stimulus material and procedures were used as in the experiment described in Chapter 2.

4.1.1 Subjects

46 bilingual Cakchiquel and Spanish speaking students from Patzicia, Department of Chimaltenango, Guatemala, participated. Of this group there were 30 sixth grade students; 11 females and 19 males whose average age was 14, and 16 highschool students; 10 females and 6 males whose average age was 17. All students were tested in classrooms and were from Cakchiquel families.

[note: the K'ekchi subjects are detailed in Chapter 2, Section 2.1.1].

4.2 Results

In order to establish the statistical significance of the stress judgments, the Cakchiquel data were submitted to a classical two-way analysis of variance with duration and position of the target syllable as factors, assuming fixed effects. The criterion variable was the percentage of lengthened syllables judged "stressed".

Results of the analysis of variance show that target syllable position constitutes a highly significant effect, $F(3, 1452) = 5.6$, $p = .001$.

Duration of the target syllable also exerts a significant effect. $F(3, 1452) = 4.3$, $p = .005$.

Finally, the interaction between target syllable position and duration is insignificant, $F(9, 1452) = 1.5$, $p = .132$, n.s.

Figure 4.1 contains the results of the experiment in terms of percentage of perceived stress syllables as a function of the temporally deviant target syllable. As indicated in Fig. 4.1 the greatest percentage of lengthened syllables perceived as stressed is in the longest condition (the 200), and the least, in the shortest (120). The mean percentage of perceived stresses as the target syllable assumes larger durations is 24%, 29%, 25% and 35% for the 120, 140 160 and 200 conditions respectively.

4.3 A Comparison of Cakchiquel and K'ekchi

The fact that duration exerted a significant effect in Cakchiquel but did not in K'ekchi supports the hypothesis that differences in duration are used differently in each of the languages.

The Cakchiquel and K'ekchi data were further submitted to a three-way analysis of variance with language, duration and position as factors. Of importance is the fact that the two-way interaction of language and duration was not statistically significant, $F(3, 2426) = 1.03$, $p = .375$, n.s., however, the interaction of language and position exerted a highly significant effect, $F(3,2426) = 9.60$, $p < .001$.

Prima Facie we would expect a final position bias in both K'ekchi and Cakchiquel if durational differences were independent of the perception of stress. That durational differences influence the perception of stress differently in the two languages is support for the original hypothesis.

In addition, the position biases are found to differ in Cakchiquel and K'ekchi. The Cakchiquel data is indicative of a penultimate position bias and the K'ekchi data, a final position bias.

A similar pattern is reflected in the stress judgments for the data as a whole. Figure 4.3 represents the mean percentage of lengthened syllables judged "stressed" as a function of syllable position for Cakchiquel and K'ekchi. In Cakchiquel, the antepenultimate and penultimate positions are more accurately associated with "stress" when lengthened than the final position. The mean percentage of stress judgments for these positions were 32% and 33% respectively, as compared to 21% and 26% for first and final positions respectively. In K'ekchi the mean percentage of stress judgments increases monotonically across the word, 18,22,30 and 43% for 1st, 2nd, 3rd and final positions respectively.

A-posteriori tests for contrasts indicate that only the difference in percent "correctly" perceived stress between second and third positions is statistically nonsignificant by the Newman-Keuls procedure at a .05 level of significance in Cakchiquel, and only positions 1 and 2 are statistically insignificant in K'ekchi. All other positions reach significance.

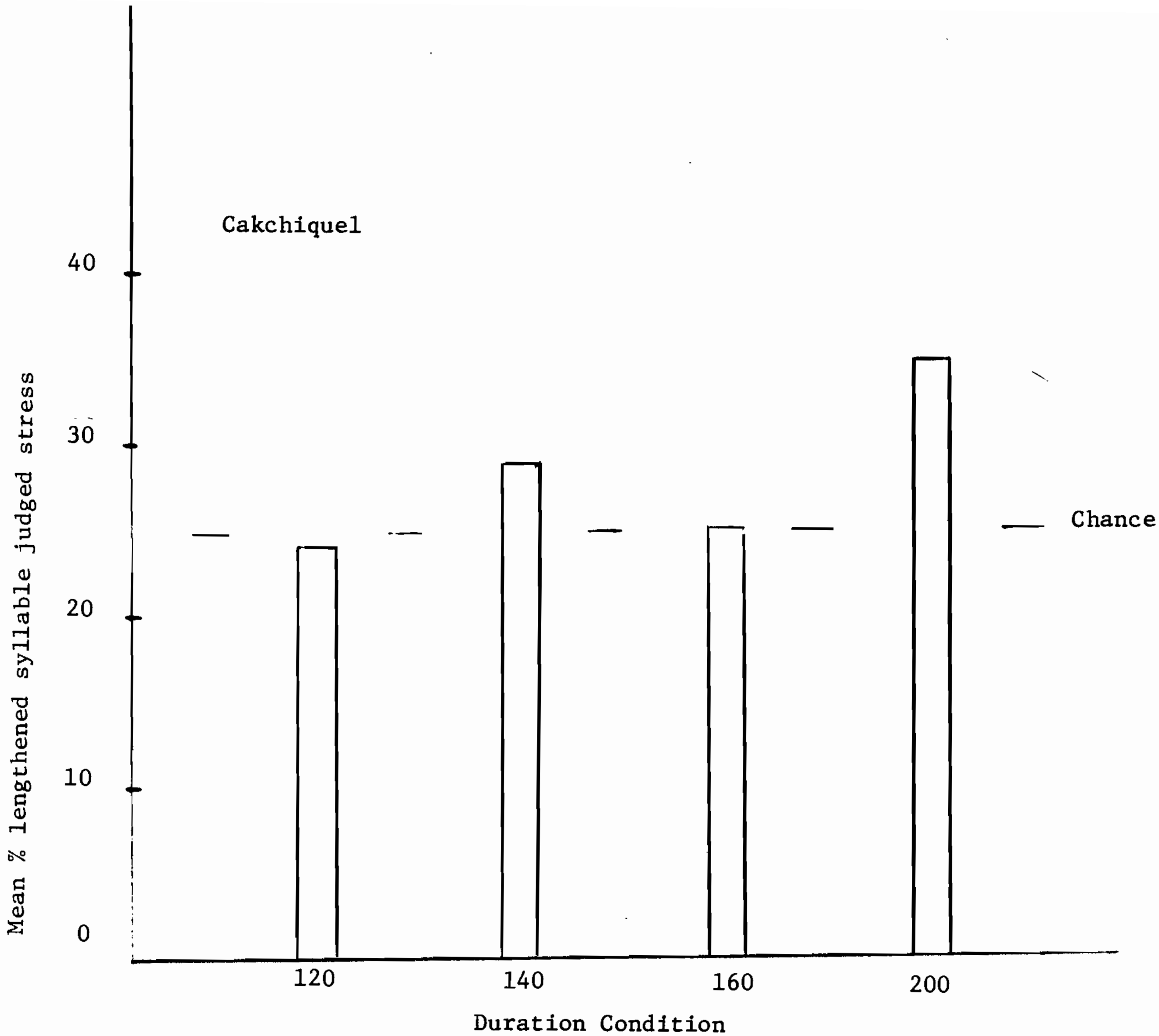

Figure 4.1. Mean percentage of lengthened syllables judged stress as a function of duration conditions.

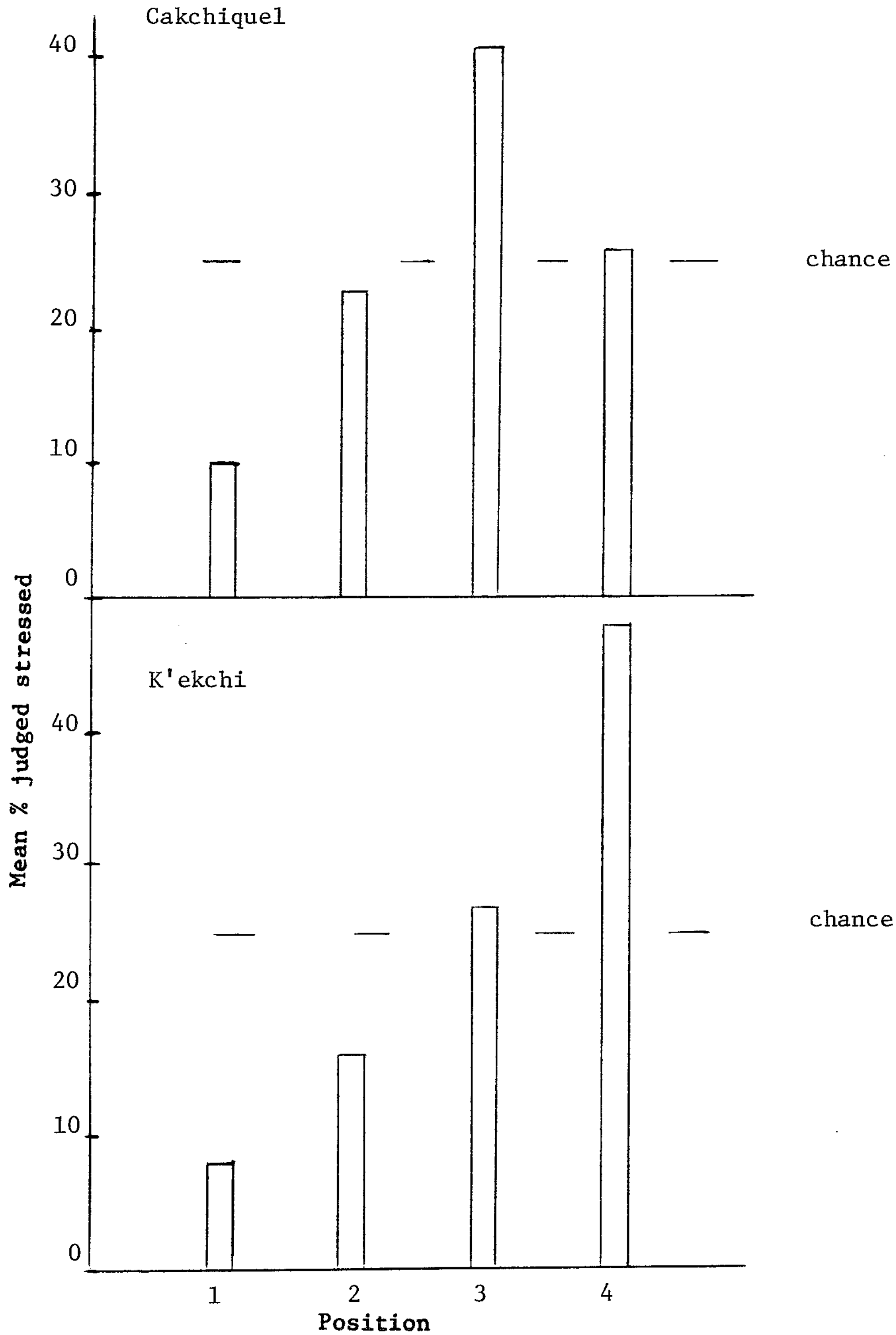

Figure 4.2. Percentage of stress responses per position in the control condition in Cakchiquel and K'ekchi.

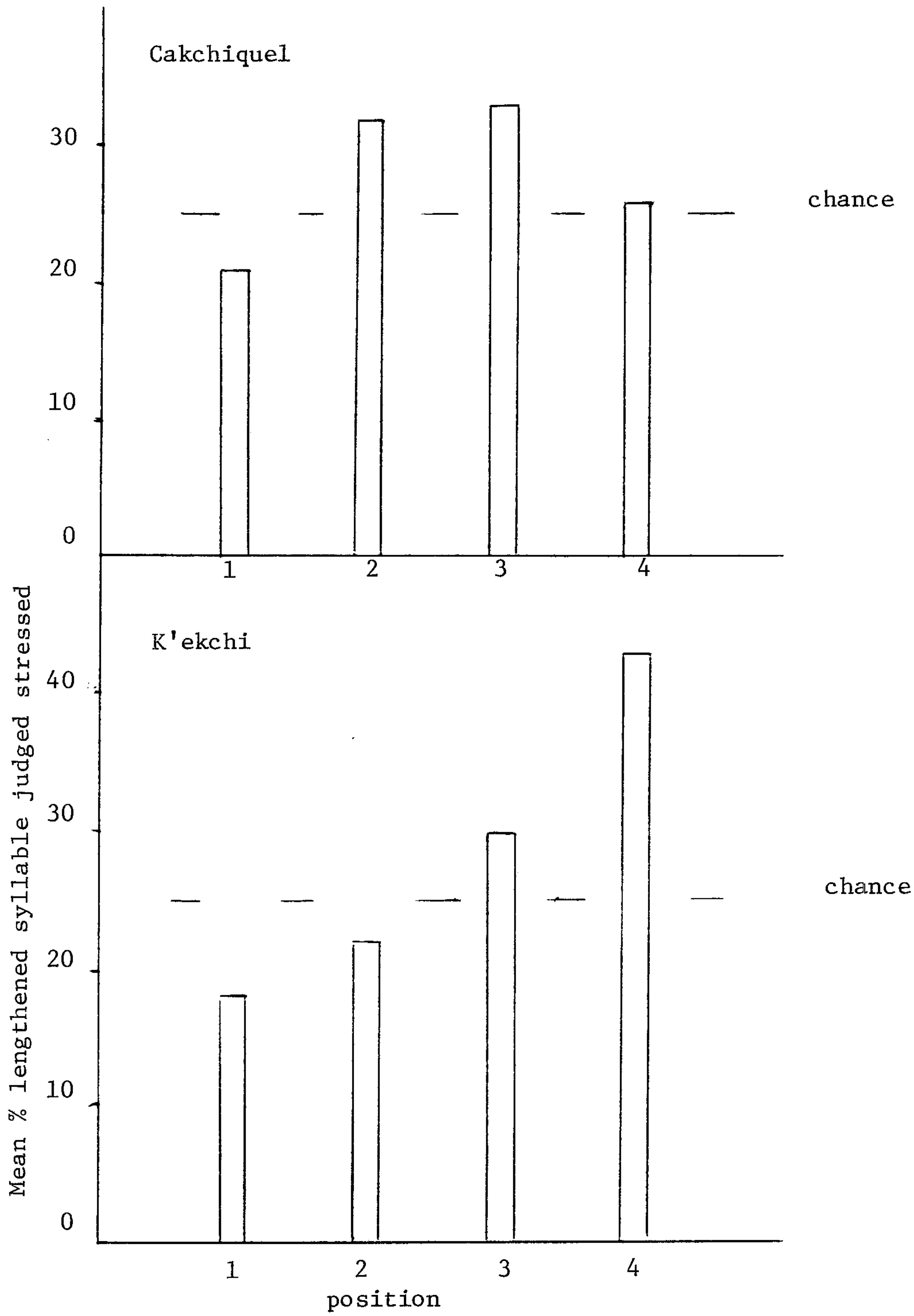

Figure 4.3. Percentage of stress responses in Cakchiquel and K'ekchi as a function of position.

Finally, the average response position in the two languages differed significantly at a .05 level as determined by the Newman-Keuls procedure. The average response position in Cakchiquel was 2.60, indicative of a medial bias. The average response position for K'ekchi was 2.94.

4.4 Discussion and Conclusions

It is clear that the extent to which duration is used in K'ekchi and Cakchiquel differ significantly and that these results allow us to reject the alternative hypothesis that the unimportance of duration to the stress/no stress distinction is due to the fixed position of stress.

We may now ask, First -Why do Cakchiquel speakers use duration and Second -Why is there a medial bias in Cakchiquel, a language with fixed final stress?

The answer to the first question is that in Cakchiquel, stress is correlated with longer duration, and duration may be used to signal the stress/no stress distinction because, unlike K'ekchi, duration does not signal a phonemic length contrast.

The answer to the second question is that Cakchiquel speakers in associating greater length with stress, expect that the final syllable (the normal location of stress) will be longer than any other syllable in the word. The medial bias in Cakchiquel is in part due to this final lengthening expectation and the fact that the final position must be cued.

There is, however, one contradiction in terms. I had assumed that if a language had a bias for a particular position, then stress judgments should be attained in that position at significantly shorter durations than any other. This is clearly not reflected in the Cakchiquel data, since positions two and three have a greater percentage of stress judgments than position four (e.g. Fig. 4.3).

Is there an explanation for this? I will argue that there is an explanation, and that the assumption that I had formerly made is only partially correct. The "incorrectness" stems from the assumption that "fixed location of stress=bias position". This however, is only true in languages in which the stress position need not be cued (as in K'ekchi). The position in which stress regularly occurs is only one attribute of stress properties. The other characteristics include the expectation that the syllable which occurs in that position will also be higher in fundamental and/or longer in duration and/or greater in intensity than any other syllable in the word (or whatever the properties of "stress" are in the particular language). The bias cannot be shifted to another position unless the acoustic expectation is compensated for. Thus the bias in English can be shifted from initial position, if the reference duration for first, second, third and fourth positions is 100, 100, 100 and 200 msec respectively, instead of 100, 100, 100 and 100.

That is, if the final lengthening expectation is built into the stimuli, then (as a pilot study indicates), the bias will be predominantly anti-penultimate. Further, we might predict that were a three syllable word tested under these conditions, the bias position would be penultimate in English.

In Cakchiquel, the fact that there is no condition in which the majority of perceived stresses is final, albeit that is the position in which stress normally occurs, is because the position cue is less effective than the duration cue, as an indicator of stress. This need not be the case for every language, just as the ranking of acoustic cues to stress perception need not be the same in every language. The "acoustic compensation" which would shift stress judgments to final position in Cakchiquel is greater duration and fundamental frequency. That is, speakers not only expect greater length, they also expect a change of Fo. Thus, the final position bias is not reflected in Cakchiquel because of a deficiency in acoustic stress cues in that position.

This suggests that in the analysis of "stress properties" and their relative importance, one should either make adjustments for acoustic expectations, or utilize some correction factor for bias. This understanding can influence the interpretation of the results remarkably. For example in summarizing a group of perceptual stress studies, Lehiste (1970) indicates that fundamental frequency provided relatively stronger cues for the perception of stress than did either duration or intensity. In many cases, however, it was unclear as to whether duration played a stronger role than intensity, particularly in languages where length was contrastive.

There was, however, another instance where it was unclear what the relative importance of duration and intensity were; French. In French, according to production data in Delattre (1966), the average duration of a stressed open syllable is 246 msec. and of an unstressed open syllable, 137 msec. As "stress" is final, this duration difference is also one of final versus nonfinal positions. In a study by Rigault (1962) the word papa and the phrase "Qu'est-ce que vous faites?" was synthesized and systematic variations of fundamental frequency, duration and intensity were produced. We know that in order for the word papa to be perceived as the French word "father", the last syllable should be at least 100 msec. longer than the first. We can predict, therefore, that in conditions in which the two syllables are of equal duration, the first will be perceived as longer and thus "more stressed" than the second, particularly if it is compounded by other acoustic correlates of stress, as high intensity. The results are summarized in Table 4.1 for the condition in which Fo is constant.

First Syllable	Pa Syllable 1	Pa Syllable 2
1. High I Long D	75	25
2. High I Short D	35	65
3. Weak I Long D	40	60

Table 4.1 Percent stress judgments for each syllable when the intensity and duration of the first syllable is varied. (The second syllable is always medium intensity and long duration).

As the final length judgment is made in relation to the first syllable, only in condition #2 is this expectation "compensated" for since in the other conditions, the durations are equivalent. We will note therefore that in condition #2, the shorter duration of the first syllable (while retaining the higher intensity) resulted in a 40% increase of final stress responses, relative to condition #1. Finally, as we predicted, the condition in which the greatest percentage of initial stresses is reported is condition #1; the condition in which the effect of longer duration and higher intensity are compounded. The result reported for this study (as discussed in Lehiste, 1970) is that fundamental frequency is the most important acoustic correlate of perceived stress while the relative importance of duration and intensity are approximately equivalent. I would suggest that, were there a correction factor for bias, one would find that duration is indeed highly correlated with perceived stress in French. Other results from other languages can be likewise reanalyzed; the majority of experiments reflect position biases rather than the relative acoustic importance of a particular cue to stress perception.

In summary, a language with a fixed location of stress does not necessarily have a position bias which coincides with the stress rule. This is because, it is not always the case that this position will be cued. We have seen in K'ekchi that the position is not cued by duration while in Cakchiquel (to some extent) it is. This, we argued is because duration is used to signal a vowel length contrast in K'ekchi and it is not used in this way, in Cakchiquel.

It is interesting to note that stress in Czech, a language with fixed initial stress and phonemic vowel length, has been described by Ondráčkova (1962) as being primarily indicated by fundamental frequency and intensity, while increases in duration are minimal. On the other hand, stress in Polish, a language with fixed penultimate stress and no phonemic length has been described by Jasem, Morton and Steffen-Batóg (1968) as being primarily indicated by fundamental frequency, secondarily by duration and finally by intensity.

Therefore, we will say that whether or not a position will be cued for stress in languages with a fixed stress rule is dependent upon the phonemic use of length (or pitch) in that language. The way in which a listener uses this information is dependent upon his/her expectation of stress properties - which may vary from language to language.

Bibliography

Allen, G. (1975) 'Speech rhythm: its relation to performance universals and articulatory timing.' *J. of Phonetics* 3, 75-86.

Berinstein, A. (1978a) 'A cross-linguistic study on the perception of stress.' *Journ. of the Acoust. Soc. Amer.* Vol. 63, V8.

Berinstein, A. (1978b) 'Acoustic correlates of stress in K'ekchi.' *Journ. of the Acoust. Soc. Amer.* Vol. 64, II23.

Berinstein, A. (1979) 'Timing differences and so-called 'isochrony.' (in progress).

Bolinger, D.L. (1958) 'A theory of pitch accent.' *World* 14, 109-149.

Bolinger, D.L. (1962) 'Intonation as a universal. *Proc. 9th Intern. Congress of Ling.* 1962. The Hague: Mouton, 1964.

Catford, J.C. (1977) *Fundamental Problems in Phonetics.* Bloomington: Indiana University Press.

Chuang, C-K. and W. S-Y, Wang. (1977) 'The time-order-error in judgment of prosodic features - the pitch, the loudness, and the duration.' *Journ. of the Acoust. Soc. Amer.* Vol. 62, Paper V7.

Delattre , P. (1966) 'A comparison of syllable length conditioning among languages.' *IRAL* Vol. IV/3, September.

Elert, C. (1964) *Phonologic Studies of Quantity in Swedish.* Almqvist and Wiksells, Uppsals.

Fraise, P. (1963) *The Psychology of Time.* New York: Harper and Row.

Fromkin, V. and Ladefoged, P. (1966) 'Electromyography in speech research.' *Phonetica* 15, 219-242.

Fry, D. (1955) 'Duration and intensity as physical correlates of linguistic stress.' *Journ. Acoust. Soc. Amer.* 27, 607-611.

Fry, D. (1958) 'Experiments in the perception of stress.' *Language and Speech* 1, 126-52.

Gandour, J. (1974) 'On the representation of tone in Siamese.' *UCLA Working Papers in Phonetics* 27, September.

Hudgins, C.V. and R.H. Stetson. (1937) 'Relative speed of articulatory movements.' *Archives neérlandaises de phonétique expérimentale* 13, 85-94.

Hyman, L. (1977) 'On the nature of linguistic stress,' In Hyman, L. (ed.) *Studies in Stress and Accent, So. Calif. Occasional Papers in Linguistics* 4, 37-82.

Jakobson, R., C.G.M. Fant and M. Halle (1952) *Preliminaries to Speech Analysis*. MIT, Acoustics Laboratory, Technical Report No. 13, Cambridge, Mass.

Jassem, W. (1959) 'The phonology of Polish stress.' *Word* 15, 252-269.

Jassem, W. J. Morton and M. Steffen-Batóg. (1968) 'The perception of stress in synthetic speech-like stimuli by Polish listeners.' *Speech Analysis and Synthesis* 1, 289-308.

Jones, D. (1969) *An Outline of English Phonetics*. (9th ed.) New York: Dutton.

Klatt, D.H. (1975) 'Vowel lengthening is syntactically determined in a connected discourse.' *J. of Phonetics* 3, 161-172.

Klatt, D.H. and W. Cooper (1975) 'Perception of segment duration in sentence contexts.' In Proceedings of the Symposium on Dynamic Aspects of Speech Perception, I.P.O. Eindhoven, Netherlands (August 1975) in *Structure and Process in Speech Perception*. ed. by A.Cohen and S.G. Nooteboom.

Klatt, D.H. (1976) 'Linguistic uses of segmental duration in English: acoustic and perceptual evidence.' *Journ. Acoust. Soc. of Amer.* 59, (5): 1208-1221.

Lea, W.A. (1977) 'Acoustic correlates of stress and juncture.' In Hyman, L. (ed.) *Studies in Stress and Accent, Southern Calif. Occasional Papers in Linguistics* 4, 83-119.

Lehiste, I. (1970) *Suprasegmentals*. Cambridge, Mass. MIT Press.

Lehiste, I. (1973) 'Rhythmic units and syntactic units in production and perception.' *Journ. Acoust. Soc. Amer.* 54, 1128-1234.

Lehiste, I. (1975) 'The perception of duration within sequences of four intervals.' Paper presented at the 8th Int. Congress of Phonetic Sciences, Leeds, Aug 21, 1975.

Lehiste, I. (1976) 'Influence of fundamental frequency pattern on the perception of duration.' *J. of Phonetics* 4, 113-117.

Lieberman, P. (1967) *Intonation, Perception and Language*. Cambridge Mass: MIT Press.

Lindblom, B. (1968) 'Temporal organization of syllable production.' *Speech Transmission Lab, Quarterly Progress and Status Report.* Stockholm, Sweden, Royal Inst. of Technology 2, 1-5.

Lindblom, B. and Rapp, K. (1973) 'Some temporal regularities of spoken Swedish.' Papers from the Inst. of Linguistics 21. Univ of Stockholm.

Malmberg, B. (1966) Analyse des faits prosodiques-problèmes et méthodes. *Linguist. Theór. appl.* 3: 99-104.

Medress, M.F.,T.E. Skinner and D.E. Anderson. (1971) 'Acoustic correlates of word stress'. Paper presented at the 82nd meeting of the Acoust. Soc. of Amer. Denver, Colorado. Paper K3.

Menzareth, P. and De Oleza, J.M. (1928) *Spanishe Lautdauer.* Berlin and Leipzig.

Meyer and Gombocz. (1909) Zur Phonetik der ungarischen Sprache, *Le monde oreintal.* 1907-8, 122-87.

Nakatani, L. and J.A. Schaffer. (1976) 'Hearing "words" wihtout words: speech prosody and word perception.' *Journ. of the Acoust. Soc. Amer.* 60, paper M13.

Nie, N.H., C.H. Holl, J. Jenkins, K. Steinbrenner and D. Brent. (1975) *Statistical Package for the Social Sciences.* (2nd ed.) McGraw Hill.

Nooteboom, S.G. (1972) Production and perception of vowel duration. Doct. Dissertation. Utrecht.

Nooteboom, S.G. (1973) 'The perceptual reality of some prosodic durations.' *J. of Phonetics* 1, 25-43.

Oller, D.K. (1973) 'The effect of the position in utterance on speech segment duration in English. *Journ. of the Acoust. Soc. Amer.* 54, 1235-1247.

Ondráčkova, J. (1962) 'Contribution to the question concerning the rhythmical units in Czech.' *Phonetica* 8, 55-72.

Peterson, G.E. and I. Lehiste (1960) 'Duration of syllable nucleii in English.' *Journ. Acoust. Soc. Amer.* 32, 693-703.

Pike, K. and G. Scott (1975) 'Pitch accent and non-accented phrases in Fore (New Guinea). (ed. Ruth Brend) in *Bibliotheca Phonetica,* No. 11, 173-186, Karger, Basel. (1st print.: *Z. Phonetik* 16: 179-189. (1963)).

Pisoni, D. (1976) 'Fundamental frequency and perceived vowel duration.' *Journ. of the Acoust. Soc. Amer.* 59, Paper T2.

Rigault, A. (1962) Rôle de la fréquence, de l'intensité et de la durée Vocalique dans la Percepcion de l'accent en Français. *Proc. 4th Intern. Congress Phonetic Sciences.* The Hague: Mouton, 735-748.

Rosen, S.M. (1977a) 'The effect of fundamental frequency patterns on perceived duration.' *STC-QPSR,* 17-30.

Rosen, S.M. (1977b) 'Fundamental frequency patterns and the long-short vowel distinction in Swedish. *STL-QPSR* 1/1977, 31-37.

Sargent, (1975) 'Computer algorithms for the extraction and application of stress contours from continuous speech sentences. Report No. TR-EE75-44: School of Electrical Engineering, Purdue Univ. West Lafayette, Indiana.

Sharf, D.J. (1962) 'Duration of post-stress intervocalic stops and preceding vowels. *Lang. and Speech* 5, 26-30.

Smith, M. (1978) 'Perception of word stress and syllable length. *Journ. of the Acoust. Soc. Amer.* Vol. 63, V5.

Stetson, R.H. (1951) *Motor Phonetics: A Study of Speech Movement in Action.* Amsterdam, North Holl. Publ. Co. 1951. (2nd pr.) (1st pr. *Archives Néerlandaises de Phonétique Expérimentale 3* (1928)).

Wang, W. S-Y., I. Lehiste, C-K Chuang and M. Darnovsky (1976) 'Perception of vowel duration.' *Journ. of the Acoust. Soc.Amer.* Vol. 60, paper LL12.

Woodrow, H. (1951) 'Time perception.' In S.S. Stevens (ed.) *Handbook of Experimental Psychology.* New York: John Wiley and Sons, 1224-1236.