INFORMATION TO USERS

The most advanced technology has been used to photograph and reproduce this manuscript from the microfilm master. UMI films the original text directly from the copy submitted. Thus, some dissertation copies are in typewriter face, while others may be from a computer printer.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyrighted material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is available as one exposure on a standard 35 mm slide or as a 17" × 23" black and white photographic print for an additional charge.

Photographs included in the original manuscript have been reproduced xerographically in this copy. 35 mm slides or 6" × 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I
Accessing the World's Information since 1938

300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA

Order Number 8801712

A simultaneous physiological and acoustic study of fundamental frequency declination

Gelfer, Carole Ellen, Ph.D.

City University of New York, 1987

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

A SIMULTANEOUS PHYSIOLOGICAL AND ACOUSTIC STUDY OF FUNDAMENTAL FREQUENCY DECLINATION

by

CAROLE E. GELFER

A dissertation submitted to the Graduate Faculty in
Speech and Hearing Sciences in partial fulfillment of
the requirements for the degree of Doctor of Philosophy,
The City University of New York.

1987

This manuscript has been read and accepted for the Graduate Faculty in Speech and Hearing Sciences in satisfaction of the dissertation requirement of Doctor of Philosophy.

September 25, 1987
Date

Katherine S. Harris
Katherine S. Harris
Chair of Examining Committee

September 29, 1987
Date

Irving Hochberg
Irving Hochberg
Executive Officer

Katherine S. Harris
Katherine S. Harris

Thomas Baer
Thomas Baer

Lawrence J. Raphael
Lawrence J. Raphael
Supervisory Committee

The City University of New York

ACKNOWLEDGMENTS

So many people deserve to be acknowledged for being instrumental in the completion of this thesis. I am deeply grately to them all.

To my family and friends, for their continued support even though they never understood exactly what I was doing, and for not asking too often when I would be finished.

To friends and colleagues at The Graduate Center and at Haskins Laboratories; their encouragement and support have provided the backbone of this endeavor.

To the staff of Haskins Laboratories for the unique opportunity to really learn about speech, and to those who fixed the hardware and wrote the software, for actually making this work possible. My fondest thanks as well to Alice Dadourian and Sabina Koruluk for all their help over the years.

To my subjects, without whom none of this would exist, and who thought this work important enough to have needles stuck in their necks and to suffer the indignities that collecting respiratory data entails. And to Dr. Kiyoshi Honda for performing the tracheal punctures and EMG insertions with the skill that allowed my subjects to tolerate them.

To Rene Collier, whose own work provided an inspiration for this thesis and who helped me to get it off the ground.

To J.A. Scott Kelso, for giving me a perspective that I otherwise would not have had.

To my committee: To Larry Raphael, for his invaluable insights and for providing humor at some very dark moments. To Tom Baer, for sharing his vast knowledge of the larynx and who, in helping me to understand, helped to make this a far more interesting thesis than it might have been. To Kathy Harris, who first inspired me to do speech research and who has been willing to struggle with me from the start, believing that what I was doing was worthwhile. My heartfelt gratitude to them all for investing so much of themselves in this work.

To my husband, Harry Katz, who has never known me not to be working on my thesis, but who has sustained himself on the knowledge that I would someday finish, even when I doubted that I would. He has given more than I ever could have hoped for in the way of encouragement, support and understanding, and I thank him from the bottom of my heart.

To my children, Rachel and Yale, who are too young to understand but who have helped in their own ways. I dedicate this thesis to them and to their futures.

This work was supported by NIH Grants NS-13617 and NS-13870 to Haskins Laboratories.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST of FIGURES

CHAPTER 1

INTRODUCTION

In describing the acoustic characteristics of sentence intonation, the terms downdrift and declination have been used to characterize the behavior of both the rapid variations in fundamental frequency (F0) corresponding to syllable prominences whose peaks comprise the envelope of an F0 contour (see, for example, Cooper and Sorenson, 1981), and the slower variation in F0 that defines a reference level upon which these local prominences are superimposed (see, for example, Cohen, Collier and 't Hart, 1982). In their simplest form, these units of F0 decline correspond to what has been referred to as the unmarked breath group (Lieberman, 1967) or the intonation group (Breckenridge, 1977), being produced on a single expiration and marked on either end by a pause and/or inspiration.

Because declination occurs over relatively large segmental and temporal domains, it provides an opportunity to study one of the more global speech behaviors, and to possibly gain some insights into the size of the units over which speech is organized. While other speech phenomena, such as the exchange of sublexical units that occur in slips of the tongue (eg., Shattuck-Huffnagel, 1983) may also provide similar insights into speech organization, a unique aspect of the study of declination is the opportunity to make meaningful measures of relevant physiological events.

Acoustic and perceptual analyses of declination have typically focused on the relative stability of initial and/or final F0 values, the total amount of declination and the rate of F0 decline in the face of changes in utterance length, within and across syntactic units. A common assumption has been that the manner and extent to which these attributes vary is indicative of the psychological reality of declination and its status in a speaker's mental representation of an utterance of a given length and/or grammatical form. Unfortunately, however, cognitive processes are not readily observable. Still, to the extent that they are expected to have some physical reality, examining the patterns of control of the physiological processes that ultimately bear on the acoustic aspects of sentence intonation should provide some insight into the psychological reality of declination.

The general purpose of the physiological studies of declination has been to discover the mechanism(s) by which the frequency of vocal fold vibration is determined over time, thus accounting for F0 declination. Attempts have been made to uncover physiological data to support a "passive" model of declination - passive in the sense that declination is, for the most part, the result of the expiratory forces which act upon the vocal folds (i.e., subglottal pressure). According to such a model, speakers exercise little or no active control over these variables apart from insuring adequate conditions for maintaining phonation over some period of time. Declination, then, would emerge from this minimal amount of control.

Such a formulation, however, is seen by some researchers as being incompatible with the results of acoustic/perceptual studies in that, to date, it has been unable to account for the variability of certain acoustic aspects of declination without invoking models of "high-level preplanning" and "look-ahead mechanisms." Furthermore, while some physiological studies have demonstrated positive correlations between F0 declination and certain physiological parameters, such as falling subglottal pressure (Atkinson, 1973; Collier, 1975; Lieberman, 1967) and decreasing glottal length (Maeda, 1976), a causal relationship remains elusive. Thus, a passive model of declination, in its simplest form at least, remains unsupported.

The acoustic and physiological literature are reviewed separately here, partly for practical reasons and partly because their separation represents a dichotomy in the way research in this area has been conducted. For example, for the decade from approximately 1960-1970, the focus of declination research was on the physiological mechanisms underlying both local and global changes in F0. This generated a considerable controversy as to whether F0 control was vested in the larynx or the lungs. Studies conducted in the first half of the 1970's (eg., Atkinson, 1973; Collier, 1975; Maeda, 1976) basically extended these earlier studies, as well as the controversies generated by them, by enlarging their data bases and the number of simultaneous physiological measures made. While a number of unanswered questions still remain, the more recent focus of the research in this area has shifted toward acoustic/linguistic investigations, raising questions as to the extent to which a speaker exercises conscious control over physiological mechanisms in order to effect a planned F0 contour.

The present study represents an attempt to resolve some of the unanswered questions of the earlier physiological studies as well as those raised more recently. In addition to addressing the question of what the controlled physiological variables in sentence intonation are (for example, the larynx or the lungs), this study also attempts to answer the question of whether fundamental frequency declination requires a conscious representation of the details of the declination and the physiological means of achieving it, or whether it represents a lower level regulation, the purpose of which is to maintain dynamic stability, with minimal cognitive representation or intervention.

CHAPTER 2

REVIEW OF THE LITERATURE

Fundamental frequency declination in sentence production has been observed in a number of the world's languages [see Breckenridge (1977) and Pierrehumbert (1979) for a review]. That listeners in turn expect, or normalize for, declination has been demonstrated for English (Pierrehumbert, 1979), Japanese (Shimizu & Dantsuji, 1981), and Dutch (Leroy, 1984). Moreover, speech from which an overall downdrift in F0 has been removed sounds highly unnatural (Cohen, Collier & 't Hart, 1982).

The purpose of most acoustic studies of declination has been to establish the influence of a variety of linguistic conditions on various aspects of the F0 contour, such as initial and final F0 values and the overall trajectory or slope. The difference between conditions for any of these factors is often interpreted as evidence for the planning of declination on the highest cognitive level.

The purpose of most physiological studies has been to determine the mechanism(s) by which either local or global aspects of intonation can be generated. In general, there are two related questions that have motivated such studies. The first is whether control of F0 is determined by the properties of the respiratory system or whether it is vested primarily in the muscles of the larynx. The second is whether a speaker exercises active control over these physiological variables in producing an intonation contour.

ACOUSTIC STUDIES

Topline vs. Baseline Declination

It is generally assumed that the fundamental frequency of the voice tends to decline over the course of major syntactic constituents or units corresponding to the size of a breath group in an utterance. This declination is presumed to occur for both F0 maxima corresponding to syllable prominences whose peak values comprise the envelope of the F0 contour, and for the F0 minima that lie in between. Characterizations of these acoustic maxima and minima have been referred to as topline and baseline declination, respectively.

Whether the two declination 'lines' are acoustically equivalent has been an issue for some debate. Breckenridge and Liberman (1977) noted that the topline declined faster than the bottom line, so that the two lines tended to converge. In fact, on the basis of the relationship of the upper and lower declination lines, Liberman (1975) defined declination as the gradual narrowing of pitch range which functions to mark syntactic boundaries. Thus, "it falls out automatically from the pitch range narrowing hypothesis that each instance of a tone is lower than the preceding instance of the same tone" (Breckenridge, 1977, p.51). However, Cohen, Collier and 't Hart (1982) found that, when plotted on a logarithmic scale, the convergence of the two lines virtually disappeared, suggesting that there is no decrease in pitch range over the declination domain.

Sorenson and Cooper (1980) found that the magnitude of local F0 changes decreased from the beginning to the end of a sentence. However, Maeda (1976) suggested that these local departures were greater than the baseline by a fixed value, so that only the reference level was actually exhibiting declination. Similarly, Fujisaki, Hirose and Ohta (1979) found the magnitude of pitch accents to be nearly constant, irrespective of sentence position, when expressed as a percentage increase in F0 relative to its baseline value.

It would seem that there has been no real resolution of the issue of acoustic equivalence of the two declination lines, so that justifications for studying one over the other in an attempt to characterize fundamental frequency declination over syntactic constituents vary among researchers. For example, linguistic investigations of declination approach it from a theory of intonation that attempts to relate abstract phonological representation to factors such as tune and prominence (e.g., Liberman, 1975; Liberman & Pierrehumbert, 1984; Pierrehumbert, 1979). The assignment of stress to syllables is important to such linguistic models, so that, in this case, the topline, defined by the relationship of syllable peaks, would be the declination line of choice for study. Other researchers, however, are motivated in favor of the topline by what appear to be more pragmatic considerations. For example, in choosing to analyze the topline, Sorenson and Cooper (1980) used the following rationale:

> If the topline and bottomline were parallel to one

> another, we would expect no such shrinkage between the values of the peaks and valleys to occur. The existence of shrinkage indicates that a single baseline may not be adequate to capture the F0 contour's declination and that separate descriptions of the topline and bottomline are required. For a number of reasons, including the fact that F0 peaks are easier to measure and are perceptually more salient than F0 valleys, we have focused our attention on the form of the F0 topline (p.409).

Apparently, Sorenson and Cooper consider the behavior of the F0 peaks to represent the most salient characteristics of the contour, so that, in their view, the inability of the baseline to predict these peaks eliminates it from further consideration (Cooper & Sorenson, 1981). They have, in fact, developed a mathematical formula, the Topline Rule, to characterize the form of the topline declination. This formula predicts the intermediate peak F0 values in single-clause English declarative sentences when the values of the first and last peaks are known. However, while initial peak F0 is presumed to influence the declination, it is not considered to be a part of it. Cooper and Sorenson's rationale for bypassing the initial peak is that the initial drop in F0 is quite substantial: half the total drop in F0 occurs by the time one quarter of the sentence is produced. The point at which F0 falls to half the declination range is referred to as the 'key point', although "we do not attach any special significance to the location of the key point in terms of the speaker's production of the sentence" (Sorenson & Cooper, 1980, p.410). The topline, then, is a line drawn between the keypoint and the last peak in a sentence, and is described as a line with a (negative) slope that is two-thirds the difference in frequency between the first and last peak divided by the difference in the times at which they occur.

Still, both acoustic and physiological arguments exist in favor of measuring the baseline. Fujisaki and his colleagues (Fujisaki and Hirose, 1982; Fujisaki, Hirose and Ohta, 1979) have developed a model of intonation that allows for two basic inputs - the phrase level and accent level commands - which are realized as the 'voicing' (baseline) and 'accent' (syllabic) components, respectively. According to this model, it is the voicing component that represents the declination, while the accent components vary independently. Cohen et al. (1982) argue that that the baseline reflects the 'natural reference level', with syllable peaks reflecting local prominences. They maintain that, when declarative utterances of similar lengths produced by a given speaker are superimposed, the baseline shows greater stability (i.e., overlap) between utterances than the topline. This may be the result of a variety of factors that influence peak height (and, therefore, the topline), including the intrinsic pitch of vowels, the amount of emphasis, semantic load, or the identity of a prevocalic consonant ('t Hart, in press, reported in Leroy, 1984). Additional evidence for the selective influence of semantic load on the topline comes from Bruce

(1982). Cohen et al. argue further that, when only a single peak exists, there can be no topline, "whereas the bottom line then shows up in its pristine form" (Cohen et al., 1982, p.270); that is, the baseline declination is always definable.

Acoustic Endpoints

A considerable amount of data exists to suggest that, for declarative sentences produced by a given speaker, final F0 values are relatively stable despite changes in the length of utterances (Boyce & Menn, 1979; Breckenridge, 1977; Bruce, 1982; Sorenson & Cooper, 1980; Maeda, 1976), initial starting frequency (Cooper & Sorenson, 1981), pitch range (Liberman & Pierrehumbert, 1982), or the insertion of dependent clauses such as parentheticals (Kutik, Cooper & Boyce, 1983). It has been suggested (Cooper & Sorenson, 1981) that the stability of endpoints reflects a cognitive use by a speaker of final F0 as a target, set in advance of the execution of a sentence. However, this notion must be evaluated further in the context of other aspects of the F0 contour.

Acoustic Onsets, Range and Slope

There is evidence to suggest that sentence category influences the declination. Thorsen (1979c) found that declarative utterances evidence the steepest decline in F0, with a shallower decline for syntactically marked questions. Most of the studies on declination, however, have examined the contours of declarative sentences, and the review that follows will also focus on this sentence type.

In a study of baseline declination, Maeda (1976) found that initial F0 and the total amount of declination remained essentially invariant over differences in sentence length, with an inverse change in the slope of the declination. A similar result has been found as well for word lists of increasing length (Sternberg, Wright, Knoll & Monsell, 1980). Cooper and Sorenson (1981) noted both a variable rate of topline decline and a systematic adjustment in utterance-initial peak frequency as a function of utterance length. Thus, Cooper and Sorenson (1981; Sorenson & Cooper, 1980) find variability in both the amount and rate of F0 decline, although they maintain that the former is less variable than the latter.

Cooper and Sorenson ascribe their results to a "look-ahead" mechanism whereby the slope of F0 is programmed once the length of an utterance is known to the speaker.

> At the beginning of an utterance, the speaker's look-ahead mechanism informs him of approximate sentence length, which is somehow used to generate the approximate F0 value of the first peak. The approximate value for the last peak is also known to the speaker, as evidenced by its constancy across sentences of different length....Once the speaker begins talking, feedback (auditory or otherwise) informs him of the value of the first peak, which together with the value of the last

> peak and approximate estimated sentence duration can be used to generate the Topline Rule. The speaker then endeavors to produce those peak values of F0 based on the rule (Sorenson & Cooper, 1980, p.419).

In a study of Danish intonation, Thorsen (1986) found variation in all topline measures with increasing utterance length, where length was defined by the number of stress groups (i.e., a stressed syllable and all subsequent unstressed syllables). As length increased, there was a tendency for initial F0 to rise slightly, for endpoints to fall, and, consequently, for the total range to increase. At the same time, however, slope decreased, so that Thorsen's results regarding changes in both range and rate of F0 decline are in accordance with those of Cooper and Sorenson. And, while she does not advocate complicated mathematical computations such as those involved in Cooper and Sorenson's Topline Rule, she does advocate a 'hierarchical theory,' whereby "certain gross aspects of the intonation of an utterance are anticipated and laid out at the moment the utterance is initiated" (Thorsen, 1986, p.1041).

In a study of Swedish intonation, Bruce (1982) examined both F0 onset and initial peak (i.e., F0 maximum of first phrase accent) frequencies and found the former to be stable while the latter increased as a function of utterance length, suggesting that they may be independently controlled aspects of the intonation contour. However, unlike Maeda, Cooper and Sorenson, or Thorsen, he found "no simple variation in the slope of either the baseline or the topline due to variations in utterance length" (p.78). Instead, where the first phrase accent was higher, the F0 values of succeeding accents were higher as well relative to those where the first pitch accent was lower. Ultimately, however, F0 endpoints were the same, regardless of utterance length. Moreover, Bruce found the same results whether length was manipulated within the subject noun phrase (left expansion) or within the verb phrase (right expansion). Thus, evidence for pre-planning in Bruce's data appears to be limited to the manipulation of accent peak frequencies with increasing utterance length.

In an attempt to characterize the form of the declination, Bruce noted that the size or range of the stepping of successive accent falls decreased throughout utterances, suggesting the exponential nature of their decline. This, he suggested, is consonant with Pierrehumbert's (1980) theory of the implementation of accent downdrift in English which "proposes a local implementation rule to account for the overall F0 course of an utterance containing a number of accentual downdrifts. The rule computes the value of a given pitch accent as a constant ratio of the value of the immediately preceding pitch accent. The overall downdrift is the result of a recursive application from left to right of this local rule. Look-ahead is presupposed, but the rule still generates the exponential character of the F0 course" (Bruce, 1982, p.80).

The idea that an exponential decay characterizes the shape of the declination contour finds additional support from the studies of Fujisaki (Fujisaki & Hirose, 1982; Fujisaki et al., 1979), who has modeled F0 declination in Japanese as a second-order linear system. In a study of Japanese declarative sentences, Fujisaki et al. (1979) found the baseline declination contour to be independent of sentence length, dropping steeply at the beginning and gradually approaching an (non-zero) asymptote. Similarly, Liberman and Pierrehumbert (1982) found that, with the exception of terminal F0, an exponential function could be fit accurately to a meaningless 'utterance' composed of a string of trisyllabic words. As an exponential function, consecutive values of F0 decrease toward the asymptote by some constant proportion relative to the immediately preceding value. Consequently, successive decrements are progressively smaller until the asymptote is approached. However, if slope is calculated simply as F/t, where F represents the total decrease in F0 over the duration (t) of the utterance, then the exponential nature of the decay is obscured. That is, given a relatively constant range of F0 decline, the slopes of utterances of increasing lengths will always appear to decrease.

Other researchers, while not describing an exponential decay, per se, have noted that there does not exist a simple linear relationship between declination and utterance length. For example, 't Hart (1979) has proposed rules for the quantification of the baseline slope whereby, for Dutch utterances of less than five seconds duration, declination is linear when plotted on a log scale. This formulation implies an exponential decay, since the logarithm of each point on an exponential curve will, by definition, yield a straight line. Thus, while perhaps not expressing this view directly, 't Hart has lent additional support to the notion of the nonlinear nature of F0 declination.

If an exponential decay characterizes F0 declination, then the rate of decline should be greatest for the earlier portions of utterances. This has been noted by Bruce (1982), Cooper and Sorenson (1981) and Fujisaki et al. (1979). Furthermore, even if initial F0 is adjusted according to utterance length, these values should be unrelated to final F0 values to the extent that, for an utterance of sufficient duration, terminal F0 will be determined by the asymptote, although this value may be reached some time prior to the end of the utterance (Bruce, 1982). Thus, there may be little or no declination in the latter portions of longer utterances (Fujisaki et al., 1979). Such a formulation predicts the relative invariance of final F0, rendering the notion of terminal F0 as a mental target unnecessary.

Sentence Prominence and Declination

Lexical items within a sentence are produced with different degrees of prominence, depending on their relative importance in conveying the meaning of that sentence. With the addition of emphatic or contrastive stress, the natural focus of a sentence can be emphasized or even moved to another location within the sentence.[1] By manipulating the amount of emphasis received by a lexical item within a sentence, the effect of the location of sentence prominence on the organization of the declination can be determined.

Thorsen (1979b) compared the topline F0 contours of sentences composed of three stress groups, where contrastive stress was placed on the first, second or third stressed syllable, against a neutral counterpart; that is, a sentence containing no contrastive stress. For sentences with emphasis on the first stressed syllable, F0 was, as expected, higher than the same syllable in the neutral sentence. However, for the rest of the first stress group and through the remaining two stress groups, the F0 contour was lower than for the neutral condition. For the conditions where emphasis was placed on the second or third stressed syllable, the results were similar. That is, F0 was always higher on the syllable receiving contrastive stress, but was lower through the remaining stress group(s). Moreover, the F0 contour was also lower throughout stress group(s) preceding the contrastive stress. Thorsen's theory is that, in addition to the extra prominence received by an emphatically stressed syllable, a "shrinking" of the surrounding contour helps to promote the impression of contrast.

Looking at the effects of neutral, early and late stress on topline declination, Eady and Cooper (1986) found an interaction of prominence and declination only in the case of an early stressed syllable, following which the F0 contour was lowered relative to the neutral and late stress conditions. The latter two conditions, on the other hand, showed no effect of the placement of prominence apart from the expected local syllable effects. Thus, Eady and Cooper's results support Thorsen's only for sentences where stress occurs early in a sentence.

Bruce's (1982) Swedish study of similar design to Thorsen's, however, supports neither of the above studies. He found no reorganization of the declination as a function of the position of stress placement. Rather, his results suggest only that stress peaks show the effects of sentence length, and that the subsequent declination reflects that interaction. Moreover, he found little or no declination through the stress groups preceding the emphasized syllable.

It is possible that these discrepant results are simply the result of studying the effects of sentence prominence in three separate languages (i.e., Danish, English and Swedish). However, there is another possible explanation. In the studies of both Thorsen and Eady and Cooper, attributes of the topline were measured, whereas Bruce measured the baseline. Thus, one might draw the tentative conclusion that, while syllable peaks are indeed sensitive to the placement of sentence prominence, the baseline is not.

Declination and Complex Sentences

While the study of declination over single clause sentences can provide insights into the nature of speech organization over individual syntactic units, it does not distinguish the sentence from the clause as an organizing unit. Moreover, it does not answer the question of how syntactic structure influences the encoding of F0 contours, both in the region of syntactic boundaries and across the syntactic units that an utterance comprises.

Maeda (1976) noted that, in long utterances, the declination was characterized by a number of baselines corresponding to the division of these sentences into syntactic units. Thus, the overall contour could not be predicted from sentence duration alone. Cooper and Sorenson (1981) used their Topline Rule in an attempt to establish the domain over which declination is organized. They hypothesized that, for a sentence composed of two main clauses, if the rule were to operate over the entire sentence, the declination line would resemble that of a single clause sentence. If, on the other hand, the "domain of application" were the individual clauses, each would evidence its own declination line. The stronger version of the latter hypothesis predicts that each clause will have an identical slope and starting frequency (complete resetting), while the weaker version predicts a new declination line for the second clause, but with a decreased slope and/or starting frequency (partial resetting) (Sorenson & Cooper, 1980).

Cooper and Sorenson's results provide no evidence for sentential declination across two syntactic units. At best, they have been able to partially support their hypothesis; that is, that declination is organized according to the constituent structure of an utterance, although not as two isolated sentences with the same declination line. However, their results regarding resetting are rather inconsistent. In one study (Cooper & Sorenson, 1977), they found no evidence of resetting for conjoined main clauses, suggesting that there may be significant fall-rise patterns at clause boundaries that serve to signal the syntactic structure, but which do not necessarily represent a resetting of the declination line. In a later set of studies (Cooper & Sorenson, 1981), however, they found significant (partial) resetting in sentences composed of both conjoined main clauses and subordinate-main clauses. Cooper and Sorenson ultimately conclude that, because the Topline Rule does not predict the slope of the declination of either a complex sentence or the individual clauses that it comprises, "the speaker does impose some form of declination over the individual clauses in a two-clause utterance, but that, in addition, the entire utterance is subject to a similar process" (Cooper & Sorenson, 1981, p.91).

Other researchers have also suggested that, to the extent that a single falling contour does not characterize the declination function of a complex sentence, F0 may be reset at clause boundaries in a way that is relevant to the syntactic structure of an utterance. However, the ability to quantify resetting may depend on what one defines as the relevant aspect of F0. According to Fujisaki's model of declination, for example, it is the 'voicing' component, or baseline, that may be reset at clause boundaries, while the accent components vary independently of the baseline. Assuming the independence of these two components, this model suggests that the resetting of the baseline does not necessarily imply the resetting of the topline. Conversely, the absence of topline resetting does not preclude the presence of baseline resetting. This point was made by Cohen et al. (1982) in a re-examinaton of Cooper and Sorenson's (1977) data. While Cooper and Sorenson found F0 peaks on either side of the clause boundary to be of roughly equal height, so that it was concluded that there was no resetting, Cohen et al. demonstrated as much as a 30 Hz resetting when the baseline was measured instead.

Collier (1987) has also presented data to suggest that whether one finds F0 resetting depends on whether one measures the topline or baseline declination. He found that, when unstressed syllables occurred on either side of the boundary of subordinate-main clause sentences (representing the baseline), F0 values were higher following that boundary. Resetting was observed whether measured as the difference of syllable average F0 values, or as the difference between extreme F0 values (i.e., the lowest F0 for syllables preceding the clause boundary and the highest F0 for those succeeding it). For utterances with stressed syllables on either side of the clause boundary (representing the topline), however, there was no evidence of F0 resetting, regardless of how it was measured (i.e., average or extreme values).

Comparing clause-initial baseline values, Collier found utterance-initial values to be an average of 4 Hz higher than second clause-initial values. However, he suggests that "perceptually this difference is so small (less than one semitone) that it is likely to result in an impression of equal pitch; i.e., of 'complete resetting'" (p.410).

These results support Fujisaki's contention that the topline and baseline may vary independently of each other. However, Collier goes on to suggest that resetting, or the lack of it, is planned for by the speaker. In discussing the consistent presence of baseline resetting in his data, he concludes that it "must be considered an integral feature of the speaker's intonational plan" (p.410). The absence of topline resetting, on the other hand, "may reflect the speaker's plan to suspend declination during the pause and to start the second clause at the same pitch as at the end of the first" (p.410). Thus, while Collier's data regarding top- and baseline declination would appear to lend support to Fujisaki's hypothesis, his interpretation of the data does not.

The evidence that syntactic structure influences declination by way of its resetting at clause boundaries argues against the sentence, per se, as the unit of organization. It is a different question, however, whether the start of a new declination 'line' is equivalent to the start of a new sentence. For the most part, the evidence for this equivalence is negative. Maeda (1976) found non-initial baselines to exhibit little or no declination. Cooper and Sorenson (1981), in comparing a (second) main clause to the same clause produced as an isolated sentence, found that "peak F0 was always lower when the clause appeared at the end of a two-clause utterance, and that this difference was greatest for the first peak, generally decreasing in magnitude for each successive peak" (p.87). In addition to differences in initial peak F0 values, Collier (1987) noted a tendency for second clauses to have narrower F0 ranges, and in particular lower F0 peaks, relative to comparable first clauses. Only Bruce (1982) has found declination in second clauses to be comparable to the same clauses produced as isolated sentences.

The results of the acoustic studies highlight the need to explain the variability in certain aspects of the declination that occur as a function of variations in utterance length, stress configuration and/or syntactic structure. A more basic question, however, is whether variability in the declination necessitates explanations involving extensive preplanning on the part of the speaker, or whether these variations might be explained in such a way as to obviate the need for such elaborate theories. Obviously, this question cannot be answered adequately until the presumed physiological underpinnings of F0 declination are examined in parallel with purely acoustic studies.

PHYSIOLOGICAL STUDIES

The mechanisms by which changes in F0 may be effected are numerous, although not necessarily of equal importance. In normal speech, there are continuous and often simultaneous adjustments in the activity of a number of physiological variables. However, while all of these mechanisms may effect changes in F0 to some degree, there is no reason to suppose that they are all necessarily involved in declination, per se. Furthermore, it is important to distinguish between those which contribute to top- and baseline declination, as these mechanisms cannot be physiologically equivalent. That is, the former can only be subserved by the forces that actively raise F0, while the latter are likely to be subserved by the forces which lower it, whether active or passive.

One of the challenges of research in intonation is to separate the respective contributions of physiological mechanisms to local adjustments of F0 from their contribution to the global intonation contour, or declination. Indeed, it may be asked whether declination is the composite result of continuous local adjustments, or whether there are independent global and local aspects of F0 control.

Subglottal Pressure

Subglottal pressure (Ps) is generated primarily by the combined action of passive and active respiratory forces (Draper, Ladefoged & Whitteridge, 1960), whose contributions may vary as a function of the requirements of the vocal task (Mead, Bouhuys & Proctor, 1968). Ps is the primary determiner of vocal intensity (Draper et al., 1960; Koyama, Kawasaki & Ogura, 1969; Mead et al., 1968). Intensities typical of conversational speech are commonly produced on subglottal pressures between 5 and 10 centimeters of water (cm-H2O) (Baer, 1979), although pressures between 4 and 40 cm-H2O have been reported for vocal intensities ranging from those for quiet conversation to shouting over considerable noise (Draper et al., 1960). Below 2-3 cm-H2O, voicing cannot be maintained (Lieberman, Knudson & Mead, 1969).

Stetson (1951) suggested that all syllables were accompanied by ballistic chest pulses produced by the contraction of the internal intercostal muscles. Ladefoged (1962) modified this theory somewhat, showing instead that only stressed syllables were accompanied by increased intercostal activity and, therefore, Ps. A number of later studies have provided further evidence that an increase in Ps accompanies stressed or accented syllables (Atkinson, 1973; Fromkin & Ohala, 1968; Lieberman, 1967). Thus, because stress is marked by increases in F0 (eg., Fry, 1955), a consistent relationship between Ps and F0 would appear to exist. However, the relationship is complicated by the fact that stress is also marked by increases in intensity (Fry, 1955), which is known to correlate directly with increased subglottal pressure (Baer, Gay & Niimi, 1976). In addition, increases in F0 are also accompanied by increased laryngeal muscle activity (Atkinson, 1973; Collier, 1974; Fromkin & Ohala, 1968; Ohala, Hirano & Vennard, 1968). Thus, the question is not only whether Ps actually contributes to raising F0, but, in addition, the extent to which it contributes.

A number of investigators have demonstrated, both in laboratory preparations of animal larynges (Koyama, Harvey & Ogura, 1971) and with human subjects (Baer, 1979; Fromkin & Ohala, 1968; Isshiki, 1959; Ladefoged, 1962; van den Berg, 1959), that sudden, unexpected increases in Ps result in increases in F0 during an interval too short for even reflexive laryngeal adjustments to occur. These studies suggest that Ps can account for a change in F0 at a rate of approximately 2-7

Hz/cm-H2O, with the lower portion of the range appropriate for mid frequencies (i.e., speech) and the higher portion of the range appropriate for higher frequencies (i.e., falsetto). Thus, the question of whether increased subglottal pressure contributes to raising F0 as well as intensity can be answered. Similarly, the degree to which it contributes can be reasonably well estimated. However, it has been shown to account for only a small fraction of the F0 increase over a syllable (Fromkin & Ohala, 1968; Ohala et al., 1968).

The mechanism by which changes in subglottal pressure affect fundamental frequency is not well understood. However, it is generally believed that Ps produces changes in the both the aerodynamic forces within the glottis and the nonlinear stiffness properties of the vocal folds, both of which influence their frequency of vibration (Ishizaka & Matsudaira, 1972; Titze, 1980; van den Berg, 1959).

As lung volume decreases over the course of an utterance, Ps will decline as well if there are no compensatory adjustments. Attempts to attribute F0 declination to the falling subglottal pressure have been founded on the assumption that, in a declarative utterance, declination is a manifestation of "minimal control throughout expiration so that changes in fundamental frequency...would follow from changes in transglottal air pressure" (Lieberman, Sawashima, Harris & Gay, 1970, p.313). According to this view, declination is a passive phenomenon in the sense that it results from the interaction of the mechanical properties of the vocal folds and the expiratory forces acting upon them. Beyond insuring conditions conducive to the maintenance of phonation over some period of time, declination should emerge from a minimal amount of speaker control.

Lieberman's (1967) work represents one of the earliest attempts to relate F0 declination to an 'archetypal' articulatory pattern; that is, "the simplest, or basic, state of muscular control that would produce the intonational signal" (Lieberman et al., 1970, p.313). In support of Lieberman's hypothesis, Atkinson (1973) and Collier (1975) have demonstrated a decline in F0 that parallels the subglottal pressure contour. On a smaller scale, the fall in F0 over 'declarative' words has also been attributed primarily to a decrease in Ps (Monsen, Engebretson & Vemula, 1978). Collier's data show a drop in F0 of 5 Hz per centimeter of water, and are thus within the range of 2-7 Hz/Cm-H2O that a purely passive model of declination would predict (Baer, 1979; Hixon, Klatt and Mead, 1971; Ladefoged, 1963; Ohala, 1978; Ohala et al., 1968). Others, however, have measured too great a decline in F0 to be accounted for by the falling subglottal pressure alone. For example, Lieberman's data suggest a ratio as large as 16 Hz/Cm-H2O. Maeda (1976), while only inferring a fall of 3 centimeters of water, suggests that this is insufficient to account for the 20-40 Hz declination that he measured. What he did demonstrate, however, was a high correlation between F0 and glottal length, positing a "tracheal pull" theory whereby the vocal folds are gradually shortened as the trachea and larynx descend with the sternum as a result of lung deflation. Thus, while Maeda finds a pure subglottal pressure explanation unacceptable, his tracheal pull theory still relies on the properties of the respiratory system as the underlying mechanism in declination.

Still, the notion that F0 declination is the passive consequence of a decreasing subglottal pressure has received considerable criticism (eg., Fromkin & Ohala, 1968; Vanderslice, 1967). Rather, the role of Ps, it is suggested, is a much more primitive one. "On this view, the subglottal pressure, far from being the 'archetypal correlate' of intonation, appears purely ancillary: it is programmed to be sufficient to enable the larynx to carry out is role" (Vanderslice, 1967, p.76).

Intrinsic Laryngeal Musculature

The primary determiner of increasing F0 is the cricothyroid (CT) muscle, the contraction of which rotates the thyroid and cricoid cartilages around the cricothyroid joint, thereby stretching and increasing the longitudinal tension of the vocal folds (van den Berg, 1958). Increased CT activity correlates highly with increases in F0 (Atkinson, 1973; Baer et al., 1976; Collier, 1974; Shipp, 1975; Shipp & McGlone, 1971), although not all F0 increases are necessarily preceded by increased CT activity (Atkinson, 1973). Furthermore, while some have found the magnitude of CT activity to correlate well with the amount of F0 rise (eg., Collier, 1974), others have found that CT activity could not "be related in any direct fashion to the extent of F0 rise" (Atkinson, 1973, p.88).

In general, the relaxation of the CT will cause F0 to fall (Atkinson, 1973; Collier, 1974), and the rate at which F0 falls will vary as a function of the rate at which the cricothyroid relaxes and/or is inhibited (Collier, 1974). However, Erickson, Baer and Harris (1983) have shown the effect of CT relaxation on F0 lowering to be frequency-dependent. That is, CT relaxation appears to initiate F0 lowering for syllables with high to low frequency falls, but not for those with falls from mid to low frequency. Thus, it would appear that the influence of CT relaxation on F0 may be a function of the magnitude of its initial activity. As for the effect of cricothyroid relaxation on the overall F0 contour, Titze and Dunham (1987) have suggested that "two passive mechanisms influencing tension in the vocal fold cover, dynamic tissue stretch and stretch relaxation, could possibly be the basis for a number of observed fundamental frequency contours in speech" (p.317). Collier (1975; 1987), on the other hand, has suggested that CT relaxation may act in concert with falling Ps to produce declination.

Other intrinsic laryngeal muscles, such as the thyroarytenoid (TA) and lateral cricoarytenoid (LCA) muscles may also effect increases in F0. The contraction of the TA produces an opposing force (i.e., shortening) to the 'active tension' applied by the contraction of the CT, serving to further increase the tension of the vocal folds. A general positive correlation between vocalis or thyroarytenoid activity and F0 has been demonstrated in singing (Baer et al., 1976), the production of glissando maneuvers (Shipp, 1975), sustained phonation

(Shipp & McGlone, 1971), and the production of sentences (Atkinson, 1973; Garding, Fujimura & Hirose, 1970; Simada & Hirose, 1970), although the latter studies did not find the relationship between TA and F0 to be as consistent as that for CT and F0. In addition to its contribution to increasing F0, some studies have found the TA to contribute to the control of register, being more active in chest than falsetto register (Baer et al., 1976; Hirano, Ohala & Vennard, 1969; Hirano, Vennard & Ohala, 1970), although other studies find no such contribution from this muscle (Shipp & McGlone, 1971). In addition, the TA may contribute to the control of voice quality and voicing distinctions (Hirano, 1974), showing greater activity during voiced segments and a decrease in activity when voicing ceases (Atkinson, 1973).

Contraction of the lateral cricoarytenoid, a laryngeal adductor, may influence F0 by applying medial compression to the vocal folds. It has been shown to increase in activity during the singing of ascending scales (Baer et al., 1976) and, like the TA, it may also contribute to voicing distinctions (Atkinson, 1973). However, its role in the regulation of F0 remains unclear.

Extrinsic Laryngeal Musculature

The extrinsic muscles of the larynx are those which attach either to the laryngeal cartilages directly or to the hyoid bone from which the larynx is suspended by ligaments. The action of these muscles is complicated by virtue of their attachment to structures outside the larynx (eg., tongue, jaw, sternum), but they are generally thought to affect F0 by altering the position of the hyoid bone and/or the relative geometry of the thyroid and cricoid cartilages.

The suprahyoid musculature (e.g., geniohyoid, genioglossus) may raise fundamental frequency by causing the forward translation of the hyoid bone, resulting in the tilting forward of the thyroid cartilage and an increase in the longitudinal tension of the vocal folds (Honda, 1983). The activity of these muscles has been shown to correlate directly with both F0 and CT activity and to be involved in F0 adjustments indirectly by virtue of their primary involvement in supraglottic articulatory gestures, such as jaw and tongue raising (Honda, 1983).

It has also been suggested that vertical movement of the larynx/hyoid complex may effect changes in the vertical tension or mass of the vocal folds and, therefore, affect F0 (Ohala, 1972), although some investigators have suggested that the muscles that influence larynx height also tilt the thyroid cartilege, so that longitudinal tension is increased indirectly (Sonninen, 1968). Vertical movements of the larynx have been found to correlate with both F0 raising (Honda, 1983; Shipp, 1975; Shipp & Izdebski, 1975; Vanderslice, 1967; Ohala, 1972) and lowering (Maeda, 1976; Shipp, 1975; Shipp & Izdebski, 1975; Vanderslice, 1967), although the changes in F0 that result from the superiorly-directed vertical larynx movements appear to be associated primarily with changes in vowel quality (Honda, 1983). In singing, however, the degree to which vertical movements of the larynx correlate with F0 appears to depend on voice training. Untrained singers show positive correlations, while trained singers often show no or even negative correlations with F0 (Shipp, 1975; Shipp & Izdebski, 1975).

The downward excursion of the larynx and the lowering of F0 have been associated with the activity of the strap (i.e., infrahyoid) musculature. The sternothyroid muscle (ST), for example, has been shown to be extremely active during maximum displacement of the larynx, although it does not appear to show significant activity until the larynx is lowered beyond a certain point (Shipp, 1975). Increased sternohyoid muscle (SH) activity has been noted prior to the onset of phonation, the function of which may be to stabilize the larynx (Atkinson, 1973; Garding et al., 1970). In general, the activity of the sternohyoid is inversely correlated with F0 (Atkinson, 1973). Studies of its activity during singing show it to be most active at the lowest part of the range, gradually decreasing until its activity ceases as F0 increases (Baer et al., 1976). It has been associated with F0 lowering following a stressed syllable (Erickson & Atkinson, 1976; Ohala et al., 1968), although the timing of this activity appears to be related to the tonal structure of the syllable. For example, in falls from high to low frequencies, increased sternohyoid activity does not occur until after F0 has begun to fall, following the relaxation of CT activity (Erickson & Atkinson, 1976; Erickson et al., 1983). In syllables with falls from mid to low frequencies, on the other hand, where significant CT activity is absent, the SH shows an increase in activity prior to the onset of the fall in F0 (Erickson et al., 1983).

Atkinson (1973) has suggested that, towards the end of an utterance, F0 lowering may be assisted by extrinsic laryngeal muscles so that Ps will not fall below a level at which voicing cannot be maintained efficiently. The contribution of the sternohyoid to the overall lowering of F0 over an utterance has been suggested by observations of consistent activity throughout the latter portions of declarative utterances where F0 is falling (Fromkin & Ohala, 1968; Ohala et al., 1968), although later studies have failed to demonstrate a consistent relationship between SH activity and the global F0 contour (Atkinson, 1973; Collier, 1974; Garding et al., 1970). It has further been suggested that the role of the sternohyoid muscle in regulating F0 is an ancillary one, being related primarily to segmental gestures (Ohala and Hirose, 1970; Simada and Hirose, 1970), or "any gesture of speech or nonspeech that would most likely require a lowering or fixation of the hyoid bone..." (Ohala & Hirose, 1970, p.41).

Physiological Correlates of F0 Resetting

The studies reported above have considered the physiological variables underlying the production of linguistically (and physiologically) simple sentences; that is, those comprising a single major syntactic unit and produced on a single expiration, with an intonation contour appropriate for a declarative sentence. However, there is another class of utterances that warrants physiological description, although almost none exists to date. These are utterances whose acoustic declination may be discontinuous, such as two-clause sentences where the F0 contour is partially or completely reset following a major boundary (Cooper & Sorenson, 1981; Fujisaki & Hirose, 1982; Maeda, 1976).

While acoustic studies suggest that there is likely to be F0 resetting following a syntactic boundary, it is not clear to what physiological mechanism(s) the resetting is related. Fujisaki and Hirose (1982) have suggested that resetting is related to whether or not a speaker pauses, thus implying that it is the pause, or the accompanying physiological events, that trigger the resetting. Cooper and Sorenson (1981), on the other hand, have stated that, while resetting at clause boundaries may be accompanied by increased subglottal pressure (although they do not actually measure it), it may also occur following pauses of such short durations as to preclude a new inspiration. In the absence of a pause, they suggest (following Maeda, 1976) that the auxilliary expiratory muscles serve as the mechanism by which Ps, and therefore F0, may be raised.

To date, only Collier (1987) has looked at multiple physiological variables accompanying F0 resetting. He analyzed subglottal pressure and cricothyroid activity in relation to top- and baseline resetting and as a function of respiratory activity at the clause boundary in two-clause Dutch utterances.

With respect to partial resetting (i.e., the relationship of second clause-initial to first clause-final values), he finds an increase in Ps of approximately .81 and 1.03 cm-H20 for the baseline and topline, respectively. These increases correspond to an increase in F0 of 9 Hz for the baseline and a decrease in F0 of 9 Hz for the topline. Collier concludes that "the speaker habitually increases the Psg level after the pause, but that this increase does not explain the observed F0 differences, either in individual tokens or overall" (p.412). There are no cricothyroid data presented for these comparisons.

Unfortunately, Collier does not consider the effects of respiratory activity at the clause boundary in relation to partial resetting. With respect to the relationship of first and second clause-initial F0 values, however, he finds an almost complete resetting of F0 when there is no inspiration, but no evidence of it when an inspiration does occur. In addition, a comparison of second clause-initial values shows F0 to be lower following an inspiration. For subglottal pressure, there is no evidence of complete resetting in either respiratory condition, with second clause-initial values also being slightly lower in the presence of an inspiration. Cricothyroid values, on the other hand, correlate well with those for F0. However, the CT data in Collier's study must be interpreted cautiously (by his own admission), as the EMG signal is an extremely weak one, with the microvolt values barely exceeding what might be considered baseline noise.

Collier thus suggests a paradoxical effect of respiratory activity on resetting. That is, both F0 and CT activity are lower following an inspiration at the clause boundary than they are at utterance initiation, while Ps activity shows almost no effect at all. Collier's data show a significant correlation of fundamental frequency with CT activity and suggest to him that "the onset frequency of the declination line is dependent upon the level of CT activity" (p.415), although the amount of CT activity may be related to the point in the speaker's F0 range at which he chooses to begin a second clause. Subglottal pressure, on the other hand, shows no correlation with F0. Collier thus suggests that "the speaker raises Psg to a level that will allow for adequate phonation, without taking into account the syntactic make-up of the utterance..." and that "Psg primarily provides the driving force for phonation and cannot explain the momentary F0 values in any detail (p.413). It is thus Collier's conclusion that subglottal pressure plays a small role in local intonational phenomena like the resetting of F0, and that, to the extent that inspiring at clause boundaries has an effect on resetting, it is a negative one.

THE PLANNING AND CONTROL OF DECLINATION AND ITS RESETTING

In addition to some of the more straightforward acoustic and physiological issues in declination, the studies reviewed herein also raise the question of whether the control of fundamental frequency declination is the result of a conscious representation of the details of the declination and the physiological means of achieving it, or whether it represents a lower level regulation, the purpose of which is to maintain dynamic stability, with minimal cognitive representation or intervention.

Liberman and Pierrehumbert (1982) have considered variations in peak F0 as indices of phrasal preplanning, although they refer to these variations as differences in pitch range, by distinguishing between what they call 'hard' and 'soft' preplanning. 'Hard' preplanning is defined as processing that is essential to the realization of an utterance and which must be accomplished before its execution. 'Soft' preplanning, on the other hand, implies "the sorts of preparation that a speaker may freely choose to make, out of rational calculation, ritual observance or any other cause, and that might well be omitted for a linguistically equivalent utterance under other circumstances" (p.43).

Cooper and Sorenson's Topline Rule perhaps represents the strongest version of a 'hard' preplanning hypothesis in that it proposes that almost every aspect of a declination contour (i.e., the initial peak, the endpoint and the difference between them) for an utterance of a given length must be known to a speaker in advance of the execution of that utterance, so that rather elaborate calculations of the declination line, or slope, can be made. In addition, the Topline Rule is language-specific; that is, it seems only to be able to characterize English utterances (Cooper & Sorenson, 1981).

Perhaps the Topline Rule would not seem so overspecific were it simply suggested as a way of modelling or characterizing fundamental frequency declination, however accurate. Certainly others have employed formulae to that end (eg., Fujisaki & Hirose, 1982; Liberman & Pierrehumbert, 1982; 't Hart, 1979). But Cooper and Sorenson go a step further and suggest that the Topline Rule has a place in the mental representation of an utterance, thus elevating the specification of the declination to the realm of higher cognitive processing.

By assuming that a speaker has a desired declination contour in mind, the concept of planning implies that the speaker also has some notion of the necessary means of achieving it, either by exercising control over laryngeal muscle activity or by generating specific subglottal pressures. This point of view is promoted by Collier (1987) in his discussion of both F0 declination and resetting.

> ...some declination features are determined by linguistic variables and the speaker has to incorporate the appropriate declination values into his mental plan, when programming the production of an intonation pattern....Having computed these melodic targets, the speaker knows which physiological means will meet his perceptual ends (p.419).

The alternative to this point of view is that there is no planning or programming on the part of the speaker; that certain properties of the declination emerge, not from conscious speaker control, but from the interaction of events that conspire to produce either variability or stability in the F0 contour. For example, the amount of F0 resetting at a clause boundary may be the result of whether or not a speaker takes a new inspiration at that boundary, thus 'resetting' cricothyroid activity and/or subglottal pressure to a level appropriate for a new sentence. Although the evidence for this to date has been negative (Collier, 1987), the data exist only for one subject for whom the quality of the cricothyroid signal is questionable at best. Thus, resetting as a function of respiratory activity at the clause boundary is a subject that still warrants acoustic and physiological investigation.

If active speaker control is not involved in the production of fundamental frequency declination, the question then arises as to whether declination is simply the passive consequence, or byproduct, of uncontrolled forces. For example, the claim that declination results from "minimal control throughout expiration" (Lieberman et al., 1970, p.313) is not particularly illuminating regarding the behavior of subglottal pressure over time. In fact, none of the work that subsequently supported Lieberman's original 1967 hypothesis concerning the origins of declination in unmarked breath groups (i.e., Atkinson, 1973; Collier, 1974) has gone much further in characterizing the global Ps contour in any detail.

Some of the earliest studies of subglottal pressure in speech (e.g., Draper et al., 1960; Mead et al., 1968) found it to be stable during voice production, thus suggesting that the muscles of the respiratory system are marshalled in such a way as to maintain Ps. However, these studies examined only sustained phonations of constant amplitudes that also required constant pressures. Subglottal pressures during normal speech, on the other hand, are known to vary over time. Thus, it is not known whether the decline in subglottal pressure associated with F0 is the passive consequence of unchecked expiratory forces (i.e., lung deflation) or, alternatively, a dynamically stable variable. If it is the latter, then does this automatically imply that it is a planned (i.e., cognitively-generated) behavior?

Fowler (1984) has suggested that "it is not always justified to infer that a particular variable is explicitly controlled in a skilled activity just because it has regular properties, and it may not be wise to assume explicit control as a first hypothesis" (p.224). In characterizing the behavior of physiological systems, Yates (1982) distinguishes between what he calls 'control' and 'regulation'. By his definition, 'control' pertains to those "circumstances in which a given system receives additional inputs from...a controller, in such a manner as to cause the system of interest to show some 'desired' behavior, specified in advance in that external controller...." (p.244). 'Regulation,' on the other hand, "arises from the physical structure [of the system] and does not require the intervention of an external, active agency such as that of a controller....Regulation is intrinsic to a [physical system]; control is extrinsic...." (p.245).

While it may not be desirable to make too much of the semantic distinction between Yates' terms, there is a relevant conceptual distinction between something that is self-regulating and something that is externally regulated or controlled. If subglottal pressure is a self-regulated variable, then it should evidence properties that are relatively insensitive to external events and not predicated on speaker control. Moreover, to the extent that fundamental frequency declination is a function of the declining subglottal pressure (eg., Atkinson, 1973; Collier, 1975; Lieberman, 1967; Lieberman et al., 1970), then it could indeed be considered a passive consequence of a regulated physiological variable.

RESEARCH QUESTIONS

The present study is designed to address the following questions:

1. Is anticipated utterance length reflected in the height of the initial F0 peak of an intonation contour? A substantial length effect would implicate planning.

2. Do subglottal pressure and cricothyroid muscle activity evidence the same length effect (or lack of it) as F0? If so, does one play a larger role in determining the height of the initial F0 peak than the other?

3. Is there a reorganization of the shape (i.e., slope) of the non-initial part of the declination to the extent that active speaker control or preplanning is implicated in the realization of the declination contour?

4. What physiological mechanism(s) determines the shape of the non-initial portion of the declination contours? That is, what are the relative roles of the subglottal pressure, cricothyroid muscle and strap muscles in determining its ultimate shape or trajectory?

5. If the F0 declination is the result of a declining subglottal pressure over the course of an utterance, does it reflect a passive response of the Ps to unchecked expiratory forces? If so, then both Ps and F0 should show variability in their declination as a function of utterances whose airflow characteristics differ and which produce significant variations in the rate at which lung volume decreases over time.

6. In two-clause utterances, is the declination organized over the entire sentence or on a clause-by-clause basis? If the latter, does each clause exhibit a declination typical of a single clause utterance so that initial F0 for a second clause is reset to a level equivalent to that of utterance-initial F0? Or, alternatively, is the extent to which there is a resetting of F0 following a clause boundary a function of whether a speaker inspires at that boundary and thus 'resets' the physiological variables responsible for increasing F0 (i.e., Ps or CT)?

FOOTNOTES

[1] In English, sentence focus is usually predicted by the Nuclear Stress Rule. Thus, a linguistic distinction exists between sentence focus and contrastive stress. In Danish, however, there are no such rules for assigning sentence focus, so that "wherever such a slightly heavier stress is introduced, it invariably invokes the impression of contrast" (Thorsen, 1979b, p.65).

With the exception of Thorsen, studies of the interaction between declination and the location of prominence within a sentence have used the terms focus and contrastive stress interchangeably (Bruce, 1982; Eady & Cooper, 1986), although in all cases they are actually investigations of the effects of the latter. In the context of the present discussion, the term "stressed syllable" is used to mean constrastively stressed syllable and not sentence focus.

CHAPTER 3

METHODS

EXPERIMENT 1

Experiment 1 is an investigation of F0 declination over single clause utterances and the physiological mechanisms presumed to underly it.

Stimuli

The stimuli for Part 1 of this experiment were Dutch utterances of three lengths, ranging from six to twenty syllables (App. A1). For each length condition, emphatic stress was placed either on the first syllable receiving lexical stress (i.e., the second syllable in the utterance), the last syllable receiving lexical stress (i.e., the penultimate syllable), or both. These will be referred to as the Early, Late and Double stress conditions, respectively. In addition, each utterance type was produced in a reiterant form, using the syllables /ma/ and /fa/. The general purpose of employing reiterant speech was to neutralize segmental effects while preserving overall intonation and syllable timing (Larkey, 1983; Liberman and Streeter, 1978). By using these specific CV syllables, where the consonant was either a voiced continuant (i.e., /m/) or a voiceless fricative (i.e., /f/), the effect of varying airflow requirements on subglottal pressure could be determined. For this part of the experiment there were 27 utterance types in all (three length conditions (L1-3) X three stress conditions (Early, Late, Double) X three phonetic conditions (Dutch, /ma/, /fa/)).

Part 2 employed English utterances of three lengths, ranging from five to seventeen syllables (App. A1). The placement of stress was varied as in Part 1, with the addition of a Neutral form (i.e., no emphatic stress placement) for one subject (see below).

Two native speakers of Dutch (RC and LB) served as subjects for Part 1 of this experiment. One of these (RC), who was fluent in English, also served as a subject in Part 2. Another subject, a native speaker of American English (EB), served as the second subject in Part 2.

Subject EB also produced utterances without emphatic stress (i.e., Neutral). In addition, he attempted to produce all utterance types in reiterant form, using the syllables /ma/ and /fa/, as was done by the subjects participating in the first part of this experiment. However, the reiterant task proved to be extremely difficult for him. His reiterant productions deviated to such an extent from the utterances they were intended to mimic, in terms of the intonational pattern, the number of syllables, and the insertion of inappropriate pauses, that they were removed from any subsequent analyses. Differences in the skill with which individual speakers perform on reiterant speech tasks have been reported previously (Kelso, V.-Bateson, Saltzman and Kay, 1985; Larkey, 1983).

Data Collection and Procedures

For these subjects, standard techniques were used (Harris, 1981) to record electromyographic (EMG) data from intrinsic (cricothyroid) and extrinsic (sternohyoid) laryngeal musculature active during the raising and lowering of fundamental frequency, respectively. Insertions were verified by having the subjects perform maneuvers (vocal or nonvocal) known to involve the contraction of the intended musculature (Hirose, 1971).

Subglottal pressure was measured directly, using different methods for the different subjects. For RC and EB, a pressure transducer (Setra Systems 236L) was inserted percutaneously into the subglottal space through the cricothyroid membrane. For the second Dutch speaker (LB), a miniature pressure transducer (Millar SPC-350) was introduced pernasally through the posterior glottis into the trachea. The tracheal puncture procedure has the disadvantage of being a highly invasive procedure. In addition, the tracheal catheter is prone to clogging by tracheal secretions, which proved to be an intermittent problem for subject EB. On the other hand, it provides a signal that is quite accurate and relatively easy to calibrate. Using a miniature pressure transducer is less invasive, but calibration is more difficult, which proved to be a problem encountered in this study. In addition to being unable to calibrate the Ps signal, there was evidence of significant drift for which there could be no correction. The miniature transducer also evidenced significant sensitivity to changes in temperature that occured within the trachea upon inspiration (Cranen

& Boves, 1985) which resulted in spurious peaks of positive Ps activity for every inspiration where the Ps value should have been negative. Thus, subject LB's Ps data were unusable. Because this experiment is largely an investigation of the physiological mechanisms underlying F0 declination, this subject was removed from this study.

Respiratory activity was monitored noninvasively using a Respitrace inductive plethysmograph. Two Respibands, each containing a wire coil that acts as a transducer, are placed around a subject's chest and abdomen. The changes in the cross sectional area enclosed by each band that occur during respiration result in changes in the inductance of the coils, which are then converted into voltages by a variable frequency oscillator and an FM demodulator. These voltages are routed through DC amplifiers and are recorded on FM channels of an instrumentation tape recorder. Calibration is achieved by having a subject inhale and exhale a known quantity of air (1280 cc) in at least two of three positions (sitting, standing and supine). In this way, the relative contributions of the rib cage and abdomen can be assessed. The calibrated output of the individual bands is then summed and used as the index of changes in lung volume and airflow (rate of expiration). For subject EB, the plethysmographic data were unusable due to significant drift and sporadic discontinuities in the signal. Thus, plethysmographic data exist only for subject RC.

During the experiment, the subjects were seated in a dental chair, with their heads supported and stabilized by a head rest. Although the Respibands are presumed to be insensitive to movement because of their widths, extraneous movements were nonetheless discouraged in order to prevent movement artifact in the Respitrace signal (Bless, Hunker & Weismer, 1981).

The stimuli were read from lists held by the subjects at a comfortable distance. They were instructed to return to a resting respiratory level following each repetition in order to insure that no two consecutive items were produced on the same expiration, and that token-initial inspirations were not influenced by the final level of the previous expiration.

Fundamental frequency was derived from the output of a miniature accelerometer attached to the pretracheal skin surface. For subjects undergoing a tracheal puncture, this device was used because it provided the least amount of interference with the placement of the catheter. The accelerometer measures F0 by responding to the vibration of the neck surface, driven by the fluctuation of air pressure within the trachea due to the opening and closing of the glottis. The output signal is a relatively simple one, with the effects of vocal tract resonances absent or considerably attenuated (Stevens, Kalikow & Willemain, 1975). It has the disadvantage, however, of evidencing baseline drift, although this was not considered to be significant enough to interfere with the extraction of fundamental frequency from the signal.

As was the case in producing reiterant speech, the skill with which individual speakers can successfully produce utterances with different stress configurations appears also to vary. Subject RC was extremely successful in producing both the reiterant forms of Dutch utterances and in manipulating stress configuration, whether in Dutch or in English. Subject EB, on the other hand, had some difficulty in producing the desired stress patterns, so that judgments of produced (as opposed to intended) stress configuration were ultimately made by obtaining listener judgments.

Using the accelerometer waveform, tokens were extracted, duplicated, computer-randomized and put on audio tape. The accelerometer waveform was used because it contains the acoustic cues for stress (i.e., amplitude, duration and F0), with vocal tract resonances absent. Ten listeners heard all the tokens of each length condition twice, making judgments on the first listening as to whether emphatic stress occurred early in the utterance, and then making judgments on second listening as to whether emphatic stress occurred late in the utterance. Listeners were not asked to judge if both early and late stress occurred in an utterance, as this proved to be an extremely difficult task, even though the utterances were known to the listeners ahead of time. Judgments of the presence or absence of early and late stress were then combined to derive judgments of stress configuration over an entire token.

The criterion for inclusion of a token in the data analysis was that at least 55% (i.e., 11) of the twenty listener judgments agreed as to stress configuration (i.e., Early, Double, Late or Neutral). Only four tokens were so ambiguous as to fail to receive a majority of responses for any one stress pattern, and these were removed from subsequent analyses. In fact, only a minority of tokens failed to receive at least two-thirds of the total judgments for a particular stress pattern.

Tokens were resorted on the basis of listener judgments and analyzed accordingly. A final analysis revealed that, for 75% of all tokens, the perceived stress patterns were the same as those (presumably) intended by the speaker. In other words, only one quarter of all tokens were ultimately included in utterances with stress configurations different from the originally intended pattern.

EXPERIMENT 2

Experiment 2 is an investigation of fundamental frequency resetting.

Stimuli

The stimuli in this experiment were two-clause English sentences, variations of which were produced by five subjects. Two of the five (RC and EB) also served as subjects in Experiment 1.

Subject RC produced three utterance types. Two of these differed only in the placement of the syntactic boundary, which served to alter slightly the length of each clause. The third sentence conjoined two clauses similar to those the first two sentences comprised (App. A2).

Subject EB produced sentences similar to those produced by RC, but whose lengths were varied systematically by adding lexical items to either the first or second clause. Specifically, first clauses of three lengths were combined with second clauses of three lengths for each of two syntactic forms (subordinate-main clause and conjoined clauses) to form eighteen utterance types. In addition, both the three main first and second clauses were produced as isolated sentences, for six additional utterances.

Subjects BK, JS and KM produced a subset of the same subordinate-main clause sentences produced by RC and EB, with the addition of different sentences of similar form. The purpose of adding these utterances was to allow the comparison of two groups of utterances differing in lexical content, but whose syntactic structure would presumably produce equivalent acoustic results for F0. In both groups of sentences, lexical items were added to the second (main) clauses in order to produce utterances of three lengths for each sentence type (App. A2).

Data Collection and Procedures

In addition to fundamental frequency data, physiological data (i.e., Respitrace, CT and Ps) were collected for the two subjects (RC and EB) who participated in Experiment 1. (For these subjects, Experiments 1 and 2 were run in succession on the same day.) However, for EB, the tracheal catheter through which the pressure transducer was inserted became permanently clogged, rendering his subglottal pressure data unusable for this experiment. In addition, the EMG signal from the cricothyroid muscle changed during Experiment 2, suggesting that its placement had shifted following the completion of Experiment 1. As a result, only fundamental frequency data were usable for this subject for Experiment 2. He was retained as a subject for this experiment, however, because, in addition to being a physiological study, it was also largely an acoustic study of F0 resetting.

Plethysmographic and F0 data were recorded for the three additional subjects (BK, JS, KM). Respiratory data were collected using the Respitrace as described under Experiment 1. F0 was derived from the output of an electroglottograph (EGG) which measures changes in impedance through the pretracheal skin as a function of varying vocal fold contact (Baer, Lofqvist & McGarr, 1983; Fourcin, 1974; Scherer, Druker & Titze, 1987). This device has the advantages of evidencing less baseline shift than the laryngeal accelerometer used with the first two subjects, and of being insensitive to vocal tract resonances. For this reason, it is the device of choice, when it can be used (Askenfelt, Gauffin, Sundberg & Kitzing, 1980). For practical purposes, however, the output signals of the two devices are very similar.

Subject RC produced five repetitions of each of the three utterance types in block form under each of three conditions. In the first condition, the subject was instructed to produce each sentence on a single expiration, without a pause at the clause boundary (-pause/-inspiration). In the second condition, he was instructed to produce each sentence with a pause but no inspiration at the clause boundary (+pause/-inspiration). In the third condition, he was instructed to produce each sentence with both a pause and an inspiration at the clause boundary (+pause/+inspiration). Apparently, however, Condition 2 (+pause/-inspiration) represents a rather unnatural production strategy, for, while the subject believed he was following instructions, he actually produced a small inspiration during the pause on almost every token of two of the three utterance types. As a result, only the data for Conditions 1 and 3 (hereafter Conditions 1 and 2), for which all tokens were produced as intended, were analyzed.

Subject EB produced ten repetitions (five randomized and five non-randomized) of each of the eighteen two-clause utterances and the six single-clause utterances. Unlike RC, he received no instructions as to producing specific respiratory patterns (i.e., whether or not to breathe at clause boundaries). It was hoped that this would elicit a variety of more natural respiratory patterns over utterances of this type. The intent became moot, however, because his plethysmographic data were unuseable.

Subjects BK, JS and KM produced the experimental stimuli in two parts. In the first part, they each produced ten randomized repetitions of each of the six two-clause utterances and each of the six second main clauses as isolated sentences. They received no instructions regarding respiratory patterns except to attempt to return to a resting position following each repetition. It was hoped that, for the longest two-clause utterances at least, they would produce some tokens on a single expiration and others with an inspiration at the clause boundary, so that the effects of the different respiratory behaviors on F0 resetting could be assessed. However, no subject produced enough of both patterns to make a statistical analysis meaningful. As a result, F0 resetting data for these utterances were analyzed only with respect to utterance (i.e., Clause 2) length.

The second part of the experiment for these subjects was similar to that for RC. That is, they each produced utterances under two respiratory conditions: first, on a single expiration and second, with an inspiration at the clause boundary. A subset of the utterances produced by these subjects in the first part of this experiment were used. Five tokens of the Length 1 two-clause utterances for each of the two sentence types were produced for each condition. These tokens were analyzed separately from those in the first part of the experiment.

DATA ANALYSIS

For both experiments, the speech, glottographic, electromyographic and plethysmographic data were recorded simultaneously and analyzed using the Physiological Signal Processing (PSP) software library and either a PDP 11/45 or VAX 11/780 computer at Haskins Laboratories. The PSP library contains routines for sampling, editing, analyzing and displaying speech, EMG and movement data.

Signals were sampled from instrumentation tape. A timing track recorded on tape and read by the computer was used to synchronize signals that were not input simultaneously, such as those requiring different sampling rates and/or those requiring sampling as AC or DC signals. Both speech and laryngographic data were sampled at a rate of 10,000 samples per second. Speech data were used only for purposes of aligning tokens to specific acoustic events and determining the windows on either side of these 'line-up points' that mark the boundaries of a given utterance type. Tokens were then assigned to an appropriate utterance set and labelled for subsequent analyses.

A cepstral analysis, part of the Interactive Laboratory System (ILS) routine, was used to derive F0 from both the EGG and accelerometer signals. In very general terms, this process involves performing an FFT on the acoustic signal, converting it to a dB scale and performing an inverse FFT from which the fundamental can be more easily identified. For the reset data of one subject (EB), the output of the accelerometer was sampled directly through a Visi-pitch, a period-by-period F0 extractor. While the latter technique has the advantage of on-line performance, it also appears to have rather low threshold levels, so that tape noise is sampled as well. As a result, it was useful for only a small portion of the total F0 analysis in this study, and the ILS routine was used to analyze the remainder of the F0 data.

The two techniques also differ with regard to resolution. That is, the ILS routine is set to use a fairly broad sliding window for calculations, with a 10 ms interval between successive calculations (effective sampling rate of 100 Hz). On-line sampling through the Visi-pitch, on the other hand, yields virtually instantaneous F0 values at 5 ms intervals (effective sampling rate of 200 Hz). Thus, the ILS-calculated contours are more highly smoothed than those obtained using the Visi-pitch. However, the purposes of this study do not require resolution greater than that provide by the ILS routine.

Rectified and integrated EMG signals were sampled at a rate of 200 samples per second. A separate processing routine for sampling DC signals was used to sample the plethysmographic and pressure data, also at 200 Hz.

CHAPTER 4

RESULTS

EXPERIMENT 1: INITIAL CONDITIONS

1. The Effect of Utterance Length

Duration: Table 1 shows the averaged token durations for both subjects at each length. For subject RC, the data are collapsed across all stress (i.e., Early, Double, Late) and phonetic (i.e., Dutch, /ma/, /fa/) conditions. For RC's English utterances (hereafter referred to as RC English) and for subject EB, the data are collapsed across all stress conditions. The averaged token data for the individual stress conditions are shown in Appendix B. The numbers in parentheses indicate the number of tokens produced by each subject for each condition.

For RC, when durations are collapsed across all phonetic and stress conditions, each successive length condition is significantly longer (L2-L1=.900 sec; $t=19.92$, $p<.001$: L3-L2=1.260 sec; $t=15.37$, $p<.001$). Collapsed across stress conditions, the same results is found for RC English (L2-L1=.596 sec; $t=19.57$, $p<.001$: L3-L2=1.049 sec; $t=33.57$, $p<.001$) and EB (L2-L1=.591 sec; $t=33.76$, $p<.001$: L3-L2=.837 sec; $t=28.10$, $p<.001$).

Not suprisingly, these results show that the subjects took longer to produce sentences with a greater number of words, whatever the stress condition.[2] It thus becomes appropriate to ask whether acoustic

Table 1

Mean Utterance Durations (in Milliseconds)

	Length 1	Length 2	Length 3	
Mean	1291	2191	3451	
SD	82.38	284.79	461.43	RC
N	43	44	43	
Mean	945	1541	2599	
SD	52.32	106.43	61.60	RC English
N	15	16	15	
Mean	819	1402	2246	
SD	71.82	82.65	164.92	EB
N	43	37	41	

or physiological variables might reflect these length differences, either at the outset of a sentence or during its execution.

Fundamental Frequency: Table 2 shows the range of F0 peaks and their means for all tokens, collapsed across all conditions, for both subjects. It can be seen that EB's F0 values tend to be higher than those for both RC and RC English. In addition, RC's English utterances have significantly higher peak F0's than the Dutch, /ma/ and /fa/ utterances (i.e., RC) ($t=6.35$, $p<.001$), representing, perhaps, a difference in the strategies used by this speaker in manipulating initial stress configuration when speaking native (i.e., Dutch) and non-native (i.e., English) languages.

Figure 1 shows the averaged F0 contours in the region of the initial peaks for the three utterance lengths for the Double stress condition for RC, RC English and EB. Both the range of F0 peaks and the averaged peak F0 values for each length condition are shown in Table 3. For RC and RC English, initial F0 peak values increase as utterance length increases, although the magnitude of this increase from one length to the next is variable. In fact, while the increase in peak F0 is significant for each length increment for RC (L1 vs. L2: $t=2.284$, $p<.05$; L2 vs. L3: $t=6.546$, $p<.001$), no length increment for RC English yields a significant increase in peak F0 (L1 vs. L2: $t=.832$, $p>.1$; L2 vs. L3: $t=.328$, $p>.1$; L1 vs. L3: $t=1.117$, $p>.1$). This negative result may be due, at least in part, to the greater variability for peak F0 in RC's production of English utterances than

Table 2

Ranges and Mean Token Values for F0, Ps and CT Activity

Fundamental Frequency

(Hertz)

	Range	Mean	SD	N
RC	104-179	142	16.74	131
RC English	127-196	162	22.45	46
EB	154-323	226	49.46	121

Subglottal Pressure

(Cm-H20)

	Range	Mean	SD	N
RC	6.83-13.10	9.87	1.29	131
RC English	7.90-14.25	11.22	1.94	46
EB	7.11-14.26	10.57	2.04	106

Cricothyroid Muscle Activity

(Microvolts)

	Range	Mean	SD	N
RC	157-505	319	76.00	131
RC English	152-535	363	104.16	46
EB	36-248	143	47.65	121

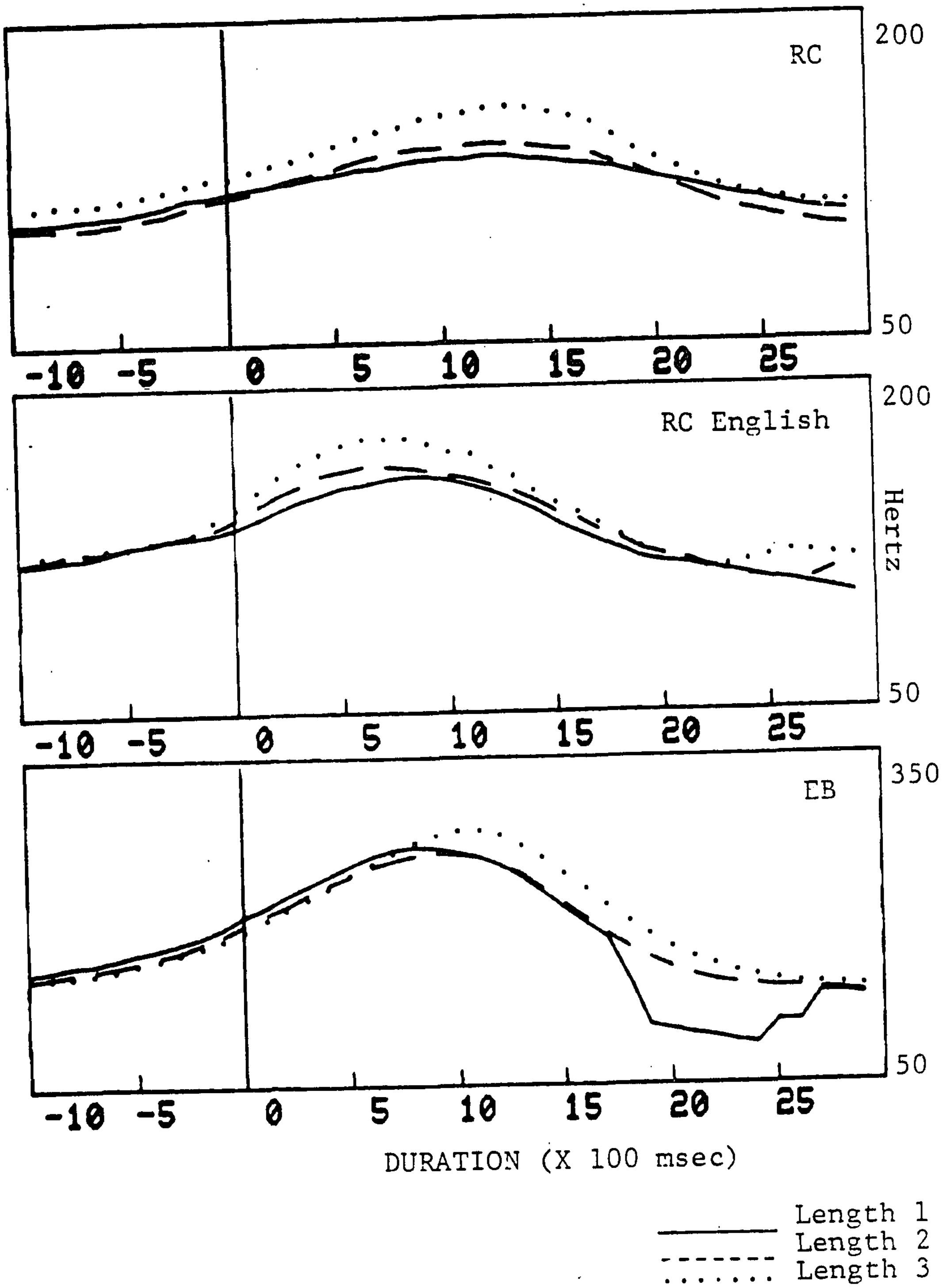


Figure 1. Initial peak F0 across utterance lengths.

Table 3

Initial Peak Fundamental Frequency (in Hz) as a Function of Utterance Length

	Range	Mean	SD	N	
Length 1	104-149	132	10.74	43	
Length 2	116-164	138	13.56	44	RC
Length 3	132-179	157	13.67	44	

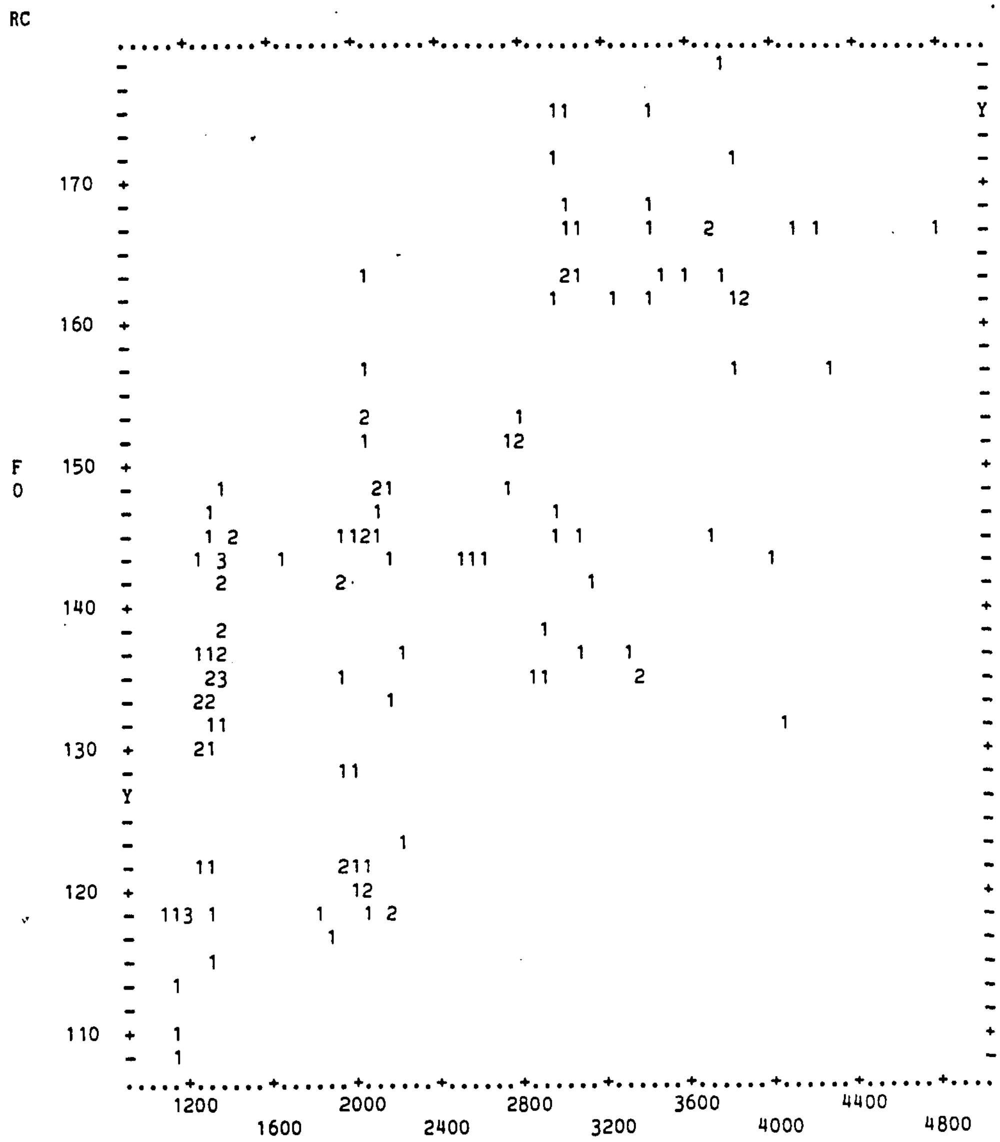
	Range	Mean	SD	N	
Length 1	132-179	157	15.33	15	
Length 2	132-192	163	23.63	16	RC English
Length 3	127-196	166	27.19	15	

	Range	Mean	SD	N	
Length 1	154-323	230	55.72	43	
Length 2	154-323	217	42.94	37	EB
Length 3	169-323	230	48.30	41	

for the production of the Dutch utterances and their reiterant counterparts, as is apparent from the standard deviations in Table 3. F0 token averages for the individual stress conditions are shown in Appendix C.

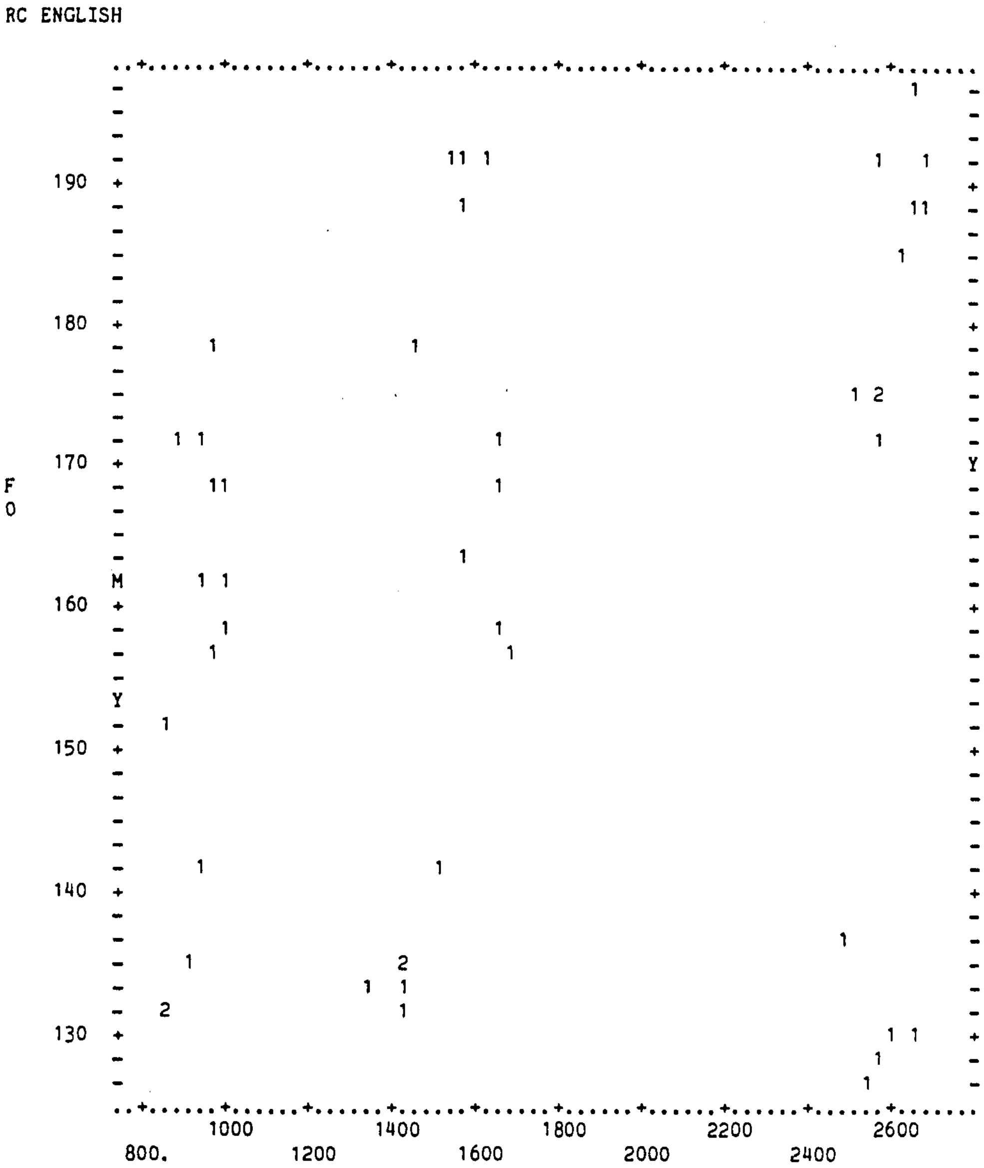
Subject EB shows no systematic increase for initial peak F0 as a function of utterance length. It can be seen from Table 3 that the average peak values for Length 1 and 3 tokens are the same, while Length 2 tokens are an average of 13 Hz lower. However, between Lengths 2 and 3, this difference is not significant $(t=1.251, p>.1)$. Within each of the four stress conditions (App. C), the pattern for peak F0 is also erratic from one length to the next.

Analyses of variance show a significant interaction between peak F0 and sentence length for RC $(F=363.95, p<.0001)$ and RC English $(F=11.2, p<.001)$, but not for EB $(F=1.19, p>.1)$. Correlations of F0 and utterance length are highly significant for RC $(r=.678, p<.001)$, but are low and nonsignificant for both RC English $(r=.232, p>.1)$ and EB $(r=.04, p>.1)$ (Figs. 2-4). Within individual phonetic and stress conditions (App. D), the same general relationship between length and peak F0 prevails for RC and EB, although, for RC English, the absence of an overall positive correlation between peak F0 and length can be attributed to a significant <u>negative</u> correlation for the Late stress utterances which obscures the significant positive correlations for the Early and Double stress utterances.



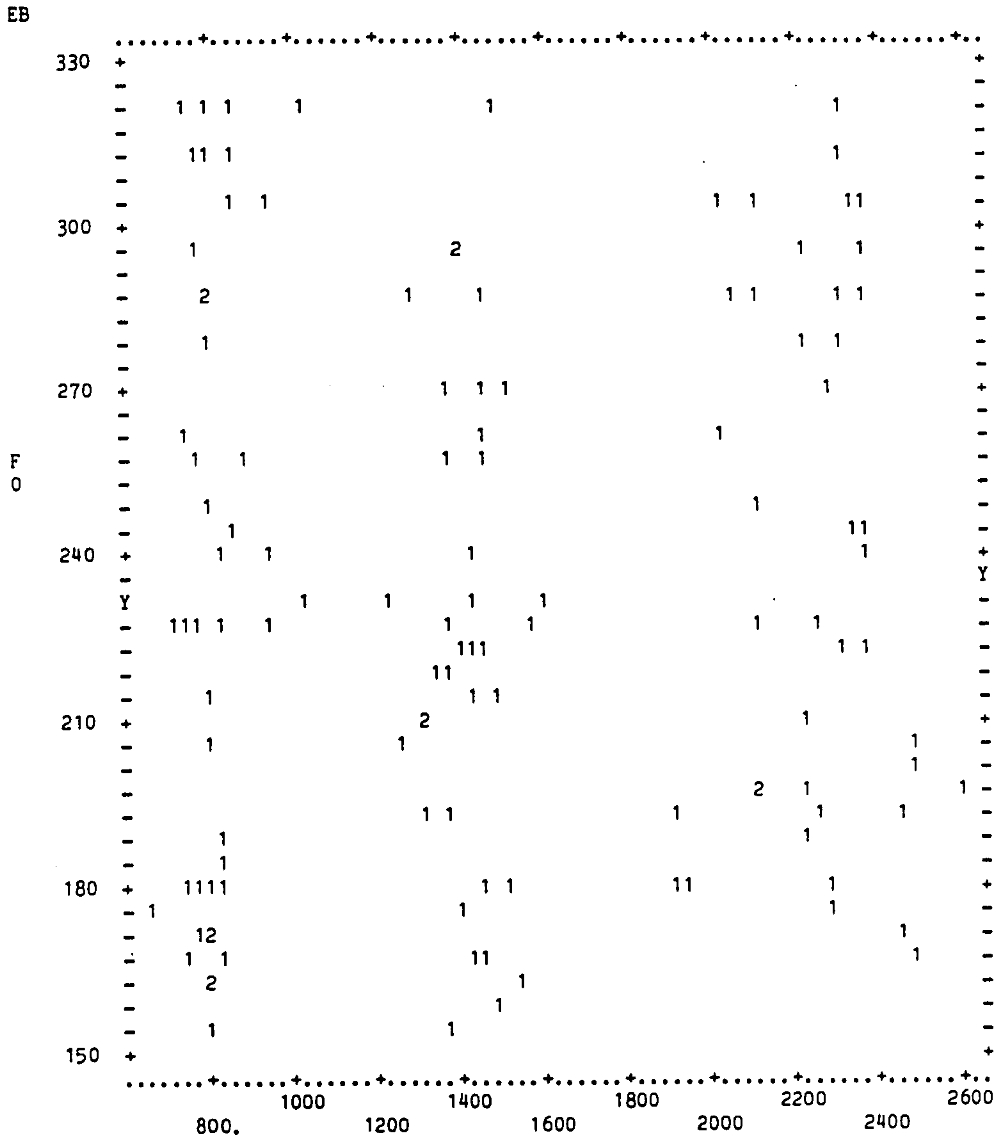
```
N =  131
R = .678                          DURATION (MSEC)
P < .001

                                     MEAN     S.D.
--REGRESSION LINE--  -RES.MS-    X  2306.9   940.22
Y= 114.80 +.01207*X    152.74    Y  142.65   16.744
```

Figure 2. Initial peak F0 (in Hertz) as a function of utterance length for RC.



```
N =   45
R = .232                        DURATION (MSEC)
P = .125
                                MEAN    S.D.
--REGRESSION LINE--  -RES.MS-  X  1704.6  688.58
Y= 148.17 +.00766*X    498.95   Y  161.22  22.703
```

Figure 3. Initial peak F0 (in Hertz) as a function of utterance length for RC English.



```
N =  121
R = .040                         DURATION (MSEC)
P = .664
                                    MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X  1486.0  613.96
Y= 228.39 +.00322*X    2463.0    Y  233.17  49.461
```

Figure 4. Initial peak F0 (in Hertz) as a function of utterance length for EB.

RC's peak F0 data thus show a systematic effect of utterance length both within and across all conditions for the Dutch, /ma/ and /fa/ utterances, while his English data show a less consistent relationship between peak F0 and length due to the differential influence of the various stress conditions. EB, on the other hand, shows no evidence of adjusting initial peak F0 in any systematic way as a function of utterance length.

Subglottal Pressure: The range and means for peak subglottal pressure for all tokens, collapsed across all conditions, are shown in Table 2. For both subjects, Ps values fall within the range of pressures acceptable for producing speech at varying intensities (eg., Baer, 1979; Draper et al., 1960). Moreover, the range of peak pressures is similar for both subjects (6.27, 6.35 and 7.15 for RC, RC English and EB, respectively). For RC, the mean peak Ps is significantly higher for the English as compared to the Dutch, /ma/ and /fa/ utterances ($t=5.306$, $p<.001$).

Figure 5 shows the averaged Ps contours in the region of the initial peaks for the three lengths, collapsed across all conditions, for RC, RC English and EB. The subglottal pressure data corresponding to the F0 data presented above are shown in Table 4 for both subjects. For EB, the asterisks indicate that there are a smaller number of tokens included in these analyses than were originally produced (see App. B) due to the intermittent clogging of the catheter through which the pressure transducer was inserted. Token averages for the

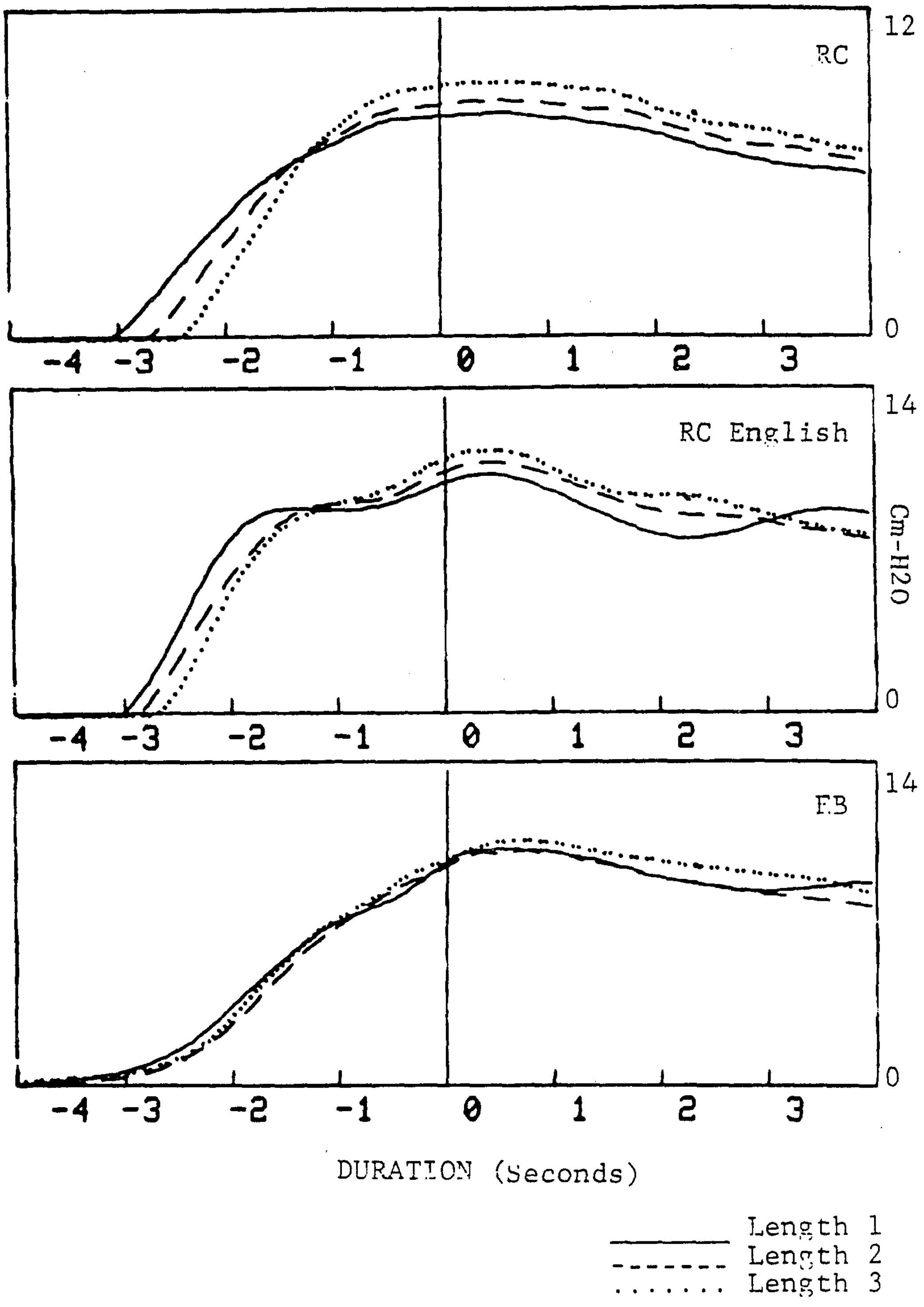


Figure 5. Initial peak Ps across utterance lengths.

individual stress conditions are shown in Appendix E. The numbers in parentheses for EB show the actual number of tokens for which there were usable Ps data. Note that no Length 2 Late stress tokens could be included in this Ps analysis.

For subject RC, the Ps data show trends that are similar to those for the F0 data (Table 4) in that there is a systematic and significant increase in peak Ps as utterance length increases (Length 1 vs. 2: t=2.596, p<.02; Length 2 vs. 3: t=2.788, p<.01). For the English utterances produced by this speaker, there is also a systematic increase in peak Ps as a function of increasing length. However, none of these increases is significant (L1 vs. L2, t=.995, p>.1; L2 vs. L3, t=.676, p>.1; L1 vs. L3, t=1.793, p>.05), again probably due to the subject's greater variability in producing English utterances.

Compared with RC and RC English, subject EB's peak subglottal pressure data show almost no effect of increasing utterance length. The greatest increase between any two conditions (i.e., .13 cm-H2O) is nonsignificant (t=.247, p>.1).

Analyses of variance reveal a significant interaction of peak Ps with utterance length for RC (F=47.28, p<.0001) and RC English (F=15.17, p<.0001), but not for EB (F=2.28, p>.1). Predictably, token-by-token correlations of peak Ps and length are significant for RC (r=.454, p<.001) (Fig. 6) and nonsignificant for EB (r=.093, p>.1) (Fig. 8). For RC English, the correlation does not reach significance (r=.291, p>.05) (Fig. 7), despite the significant interaction between

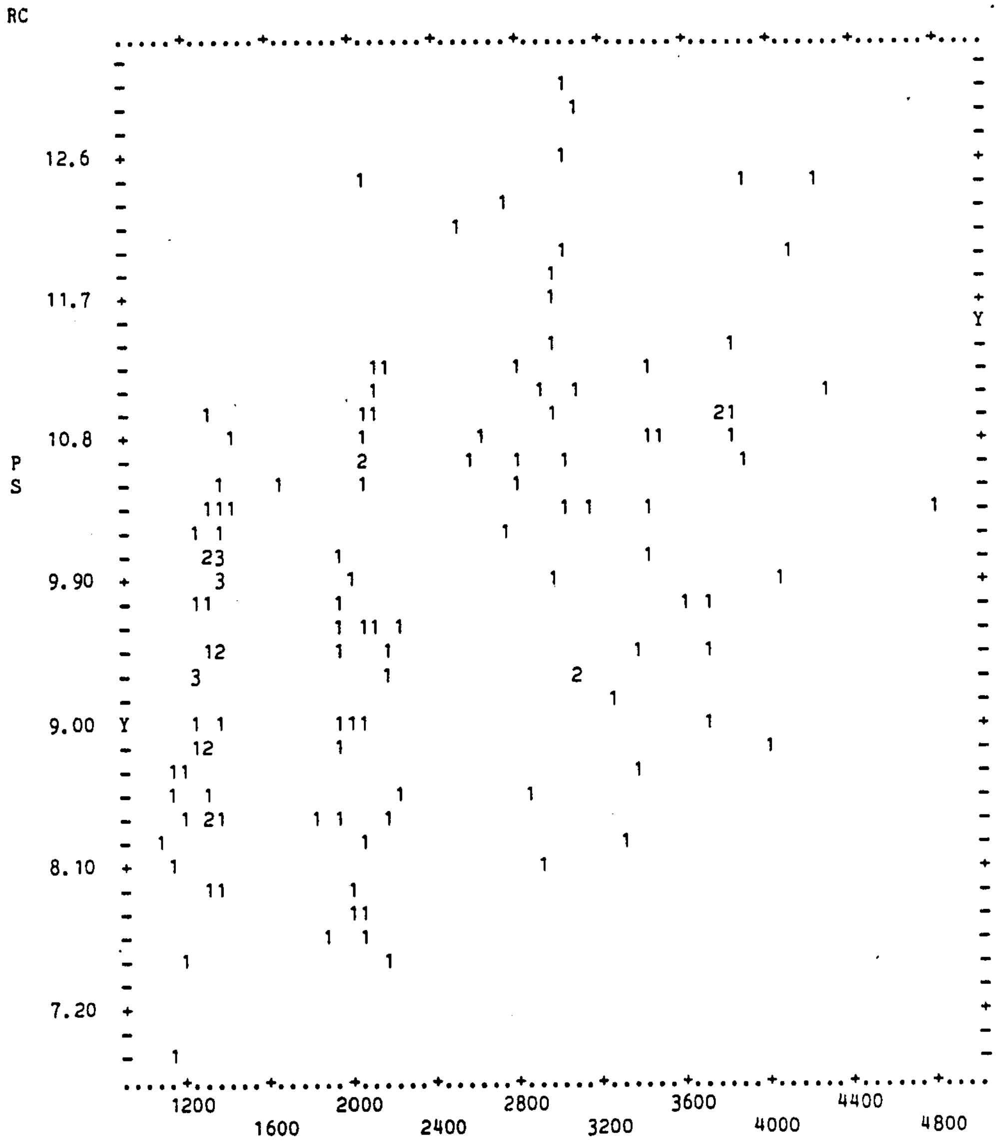
Table 4

Initial Peak Subglottal Pressure (in Cm-H2O) as a Function of Utterace Length

	Range	Mean	SD	N	
Length 1	6.83-10.88	9.17	.93	43	
Length 2	7.59-12.38	9.81	1.33	44	RC
Length 3	8.14-13.10	10.58	1.26	44	

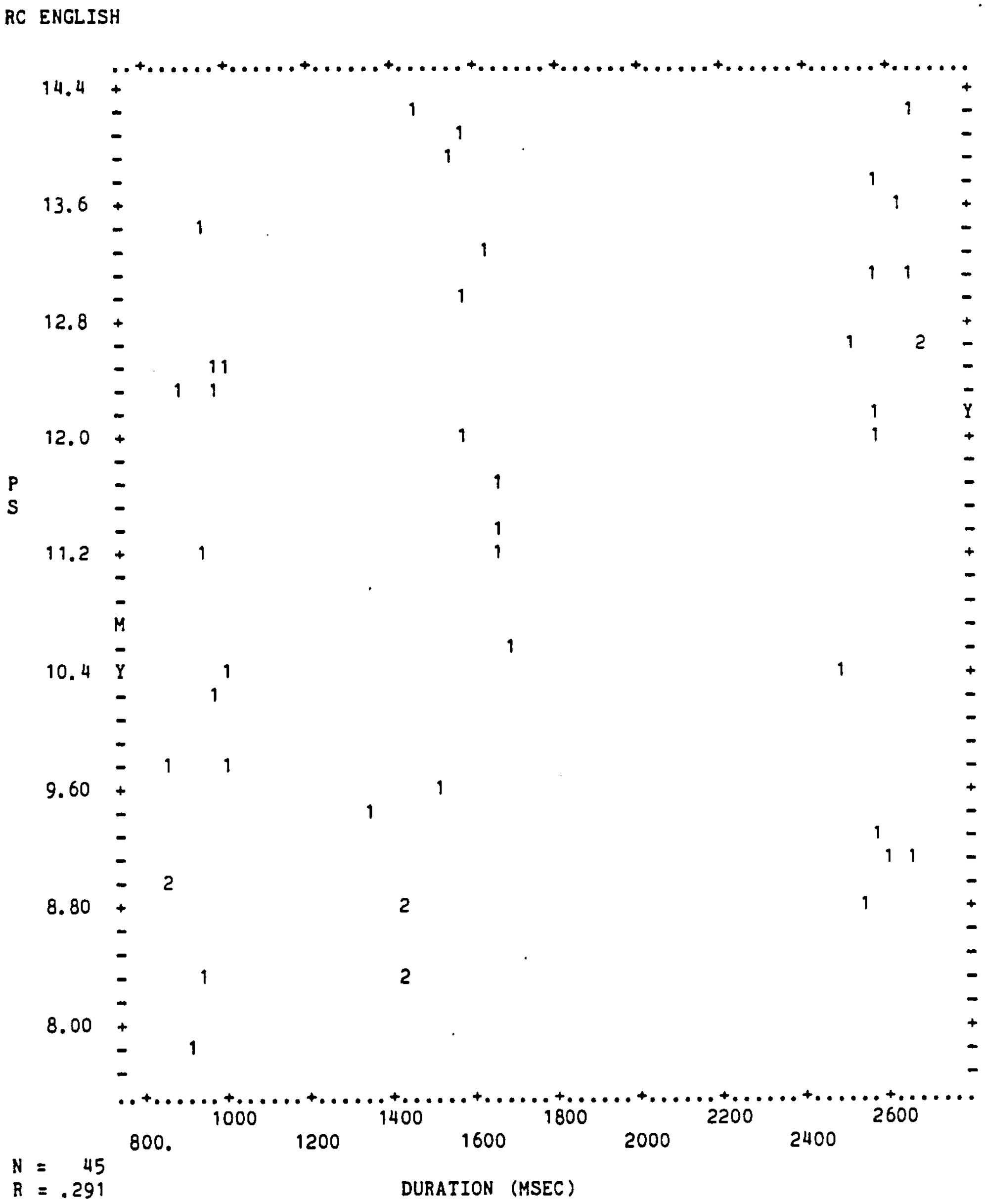
	Range	Mean	SD	N	
Length 1	7.90-13.38	10.60	1.70	15	
Length 2	8.30-14.18	11.29	2.12	16	RC English
Length 3	8.83-14.25	11.78	1.90	15	

	Range	Mean	SD	N	
Length 1	7.11-14.26	10.62	2.17	37*	
Length 2	7.61-14.08	10.75	2.20	32*	EB
Length 3	7.50-14.15	10.73	1.88	37*	



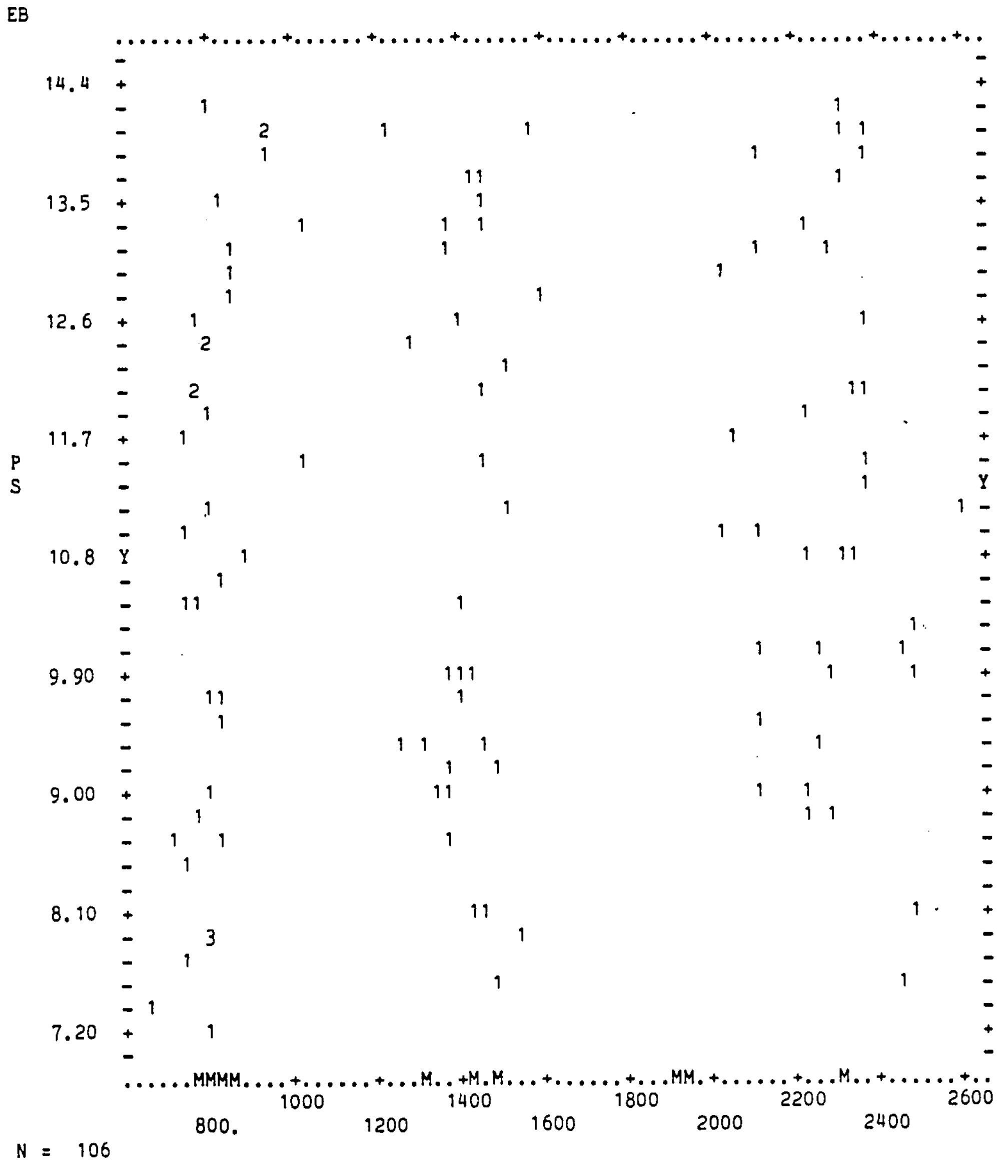
```
N =  131
R = .454                              DURATION  (MSEC)
P < .001

                                       MEAN     S.D.
--REGRESSION LINE--   -RES.MS-     X  2306.9   940.22
Y= 8.4325 +626E-6*X    1.3409      Y  9.8755   1.2948
```

Figure 6. Initial peak Ps (in Cm-H2O) as a function of utterance length for RC.



```
N =   45
R = .291
P = .052
                                          MEAN     S.D.
--REGRESSION LINE--  -RES.MS-      X   1704.6   688.58
Y= 9.7730 +828E-6*X     3.5903     Y   11.184   1.9579
```

Figure 7. Initial peak Ps (in Cm-H2O) as a function of utterance length for RC English.



```
N =  106
R = .093                        DURATION (MSEC)
P = .343
                                MEAN     S.D.
--REGRESSION LINE--  -RES.MS-   X  1509.2   625.51
Y= 10.529 +304E-6*X    4.1633   Y  10.987   2.0395
```

Figure 8. Initial peak Ps (in Cm-H2O) as a function of utterance length for EB.

peak Ps and utterance length. However, the reason for this result can be determined by examining the correlations for the individual stress conditions (App. F). For the individual stress conditions, all correlations are positive, but significance is reached only for the Double stress utterances.

Cricothyroid Muscle Activity: Both the range and mean cricothyroid activity for all tokens, collapsed across all conditions, are shown for RC and EB in Table 2. As was the case with the F0 and Ps data, the averaged peak CT data show RC's English utterances to have significantly higher values than the Dutch, /ma/ and /fa/ utterances ($t=3.051$, $p<.01$).

It is also apparent from this table that EB's CT values are substantially lower relative to those for RC, both in terms of the range of activity (i.e., 212 uV for EB as compared to 348 and 383 uv for RC and RC English, respectively) and absolute minimum and maximum levels of activity. While it is possible that the difference in CT activity for the two subjects reflects individual speaker differences, a more likely explanation is that these differences represent a placement of the electrode more proximal to the largest field of active motor units in the cricothyroid muscle for RC, resulting in a stronger signal.

Figure 9 shows the averaged CT traces in the vicinity of the initial peaks for each length condition for RC, RC English and EB, collapsed across all phonetic and length conditions. The corresponding averaged peak cricothyroid values and their ranges are presented in Table 5. Token averages for the individual stress conditions are shown in Appendix G.

For RC, CT activity increases systematically with utterance length (Table 5), although, unlike the F0 and Ps data, the difference between the means does not reach significance at every length increment (L1 vs. L2: t=.585, p>.1; L2 vs. L3: t=4.774, p<.001). Moreover, within individual stress conditions (App. G), the pattern of CT activity with increasing utterance length is less systematic than for the other two variables.

Cricothyroid activity for RC English shows a very small effect of increasing utterance length. Within stress conditions, the pattern of activity is again erratic from one length to the next (App. G). Similarly, subject EB's averaged CT data fail to show a consistent pattern of activity associated with utterance length, either within or across stress conditions.

Analyses of variance reveal a significant interaction of CT activity with utterance length for RC (F=26.14, p<.0001), but none for RC English (F=.06, p>.1) and EB (F=.06, p>.1). Token-by-token correlations of peak CT activity and length are significant for RC (r=.48, p<.001) (Fig. 10). For RC English and EB, coefficients are low

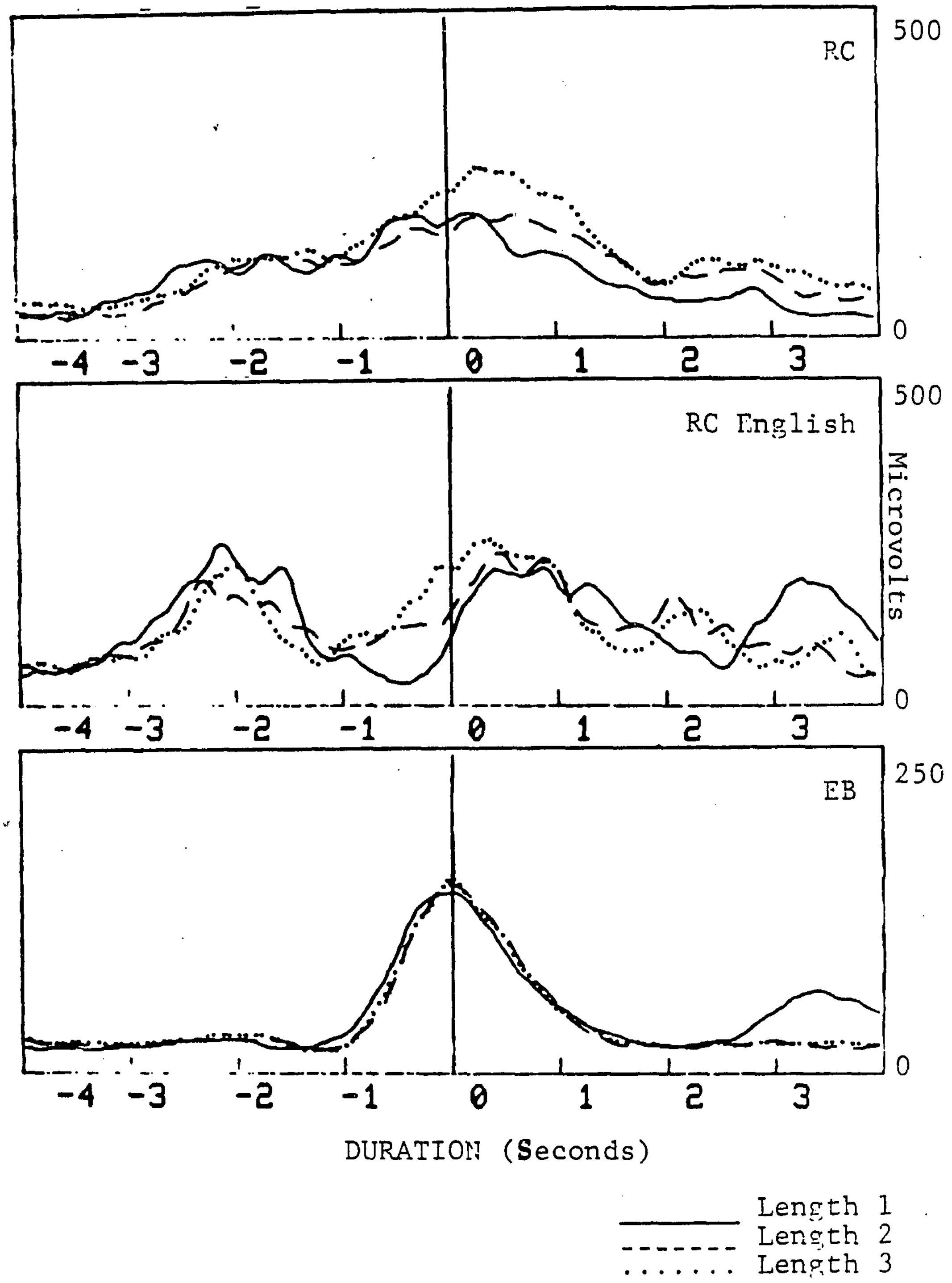


Figure 9. Initial peak CT activity across utterance lengths.

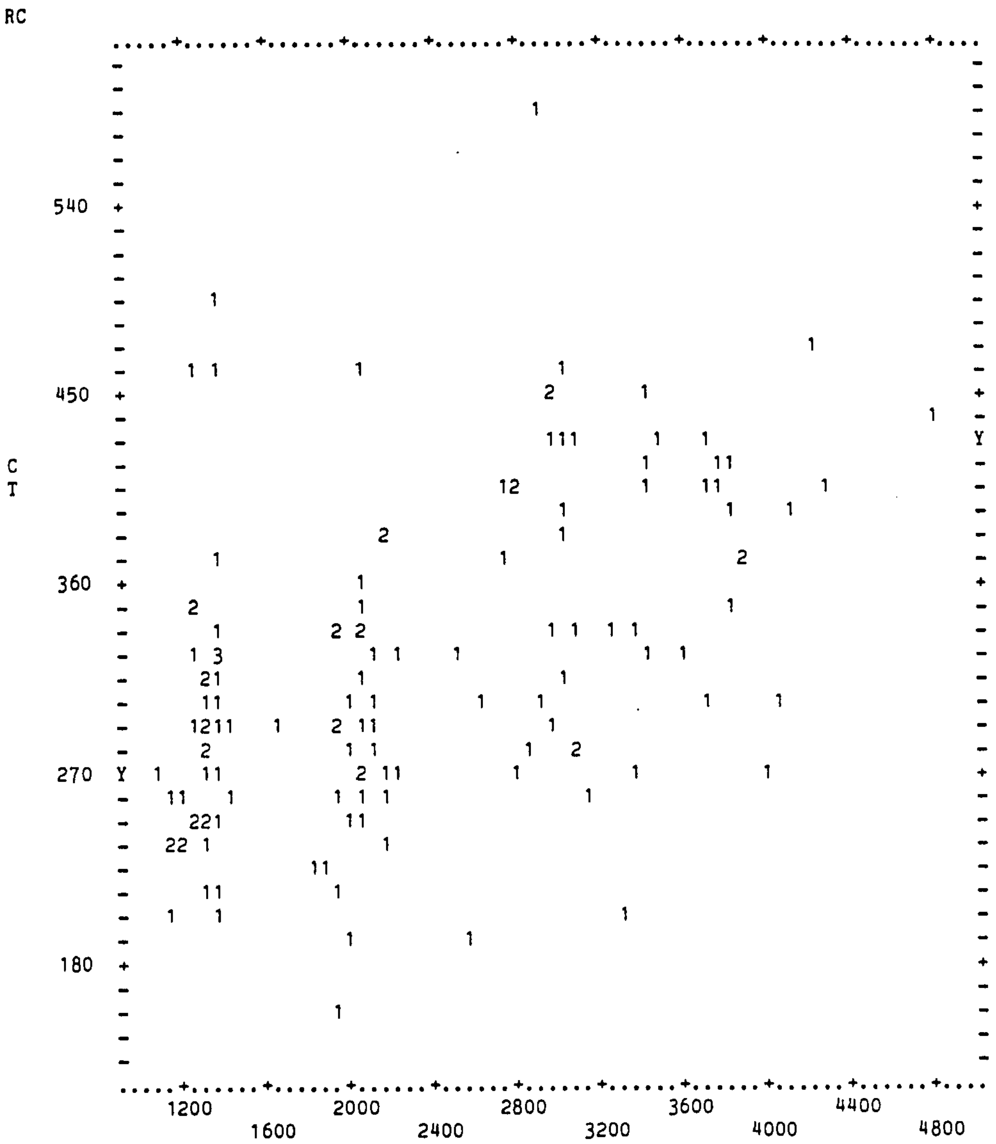
Table 5

Initial Peak Cricothyroid Muscle Activity (in Microvolts) as a Function of Utterance Length

	Range	Mean	SD	N	
Length 1	208-500	290	64.44	43	
Length 2	157-461	298	63.15	44	RC
Length 3	202-585	368	73.98	44	

	Range	Mean	SD	N	
Length 1	166-469	358	84.12	15	
Length 2	172-492	364	100.00	16	RC English
Length 3	152-535	366	130.50	15	

	Range	Mean	SD	N	
Length 1	47-222	144	54.74	43	
Length 2	36-215	142	45.75	37	EB
Length 3	49-248	142	42.38	41	



```
N =  131
R = .480
P < .001
                                        MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X   2306.9   940.22
Y= 231.30 +.03880*X    4479.1    Y   320.81   75.996
```

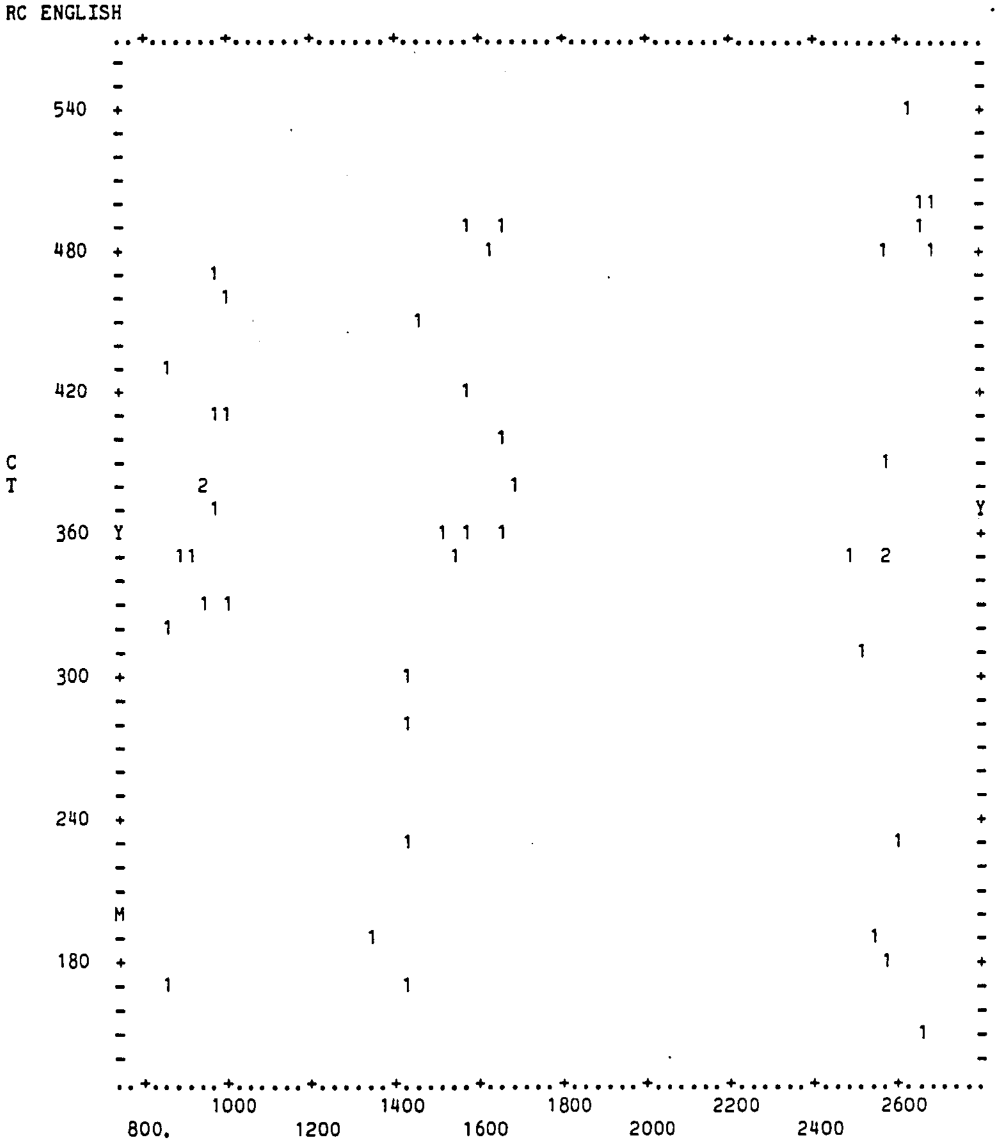
Figure 10. Initial peak CT activity (in microvolts) as a function of utterance length for RC.

and nonsignificant (RC English : $r=.058$, $p>.1$; EB: $r=.006$, $p>.1$) (Figs. 11 & 12). Correlation coefficients within conditions are shown in Appendix H.

The pattern of CT activity as a function of increasing utterance length is thus generally, but not entirely, similar to that for the fundamental frequency and subglottal pressure data for these subjects. That is, RC shows the most systematic effect, with RC English showing a less consistent effect and EB showing the smallest effect of utterance length.

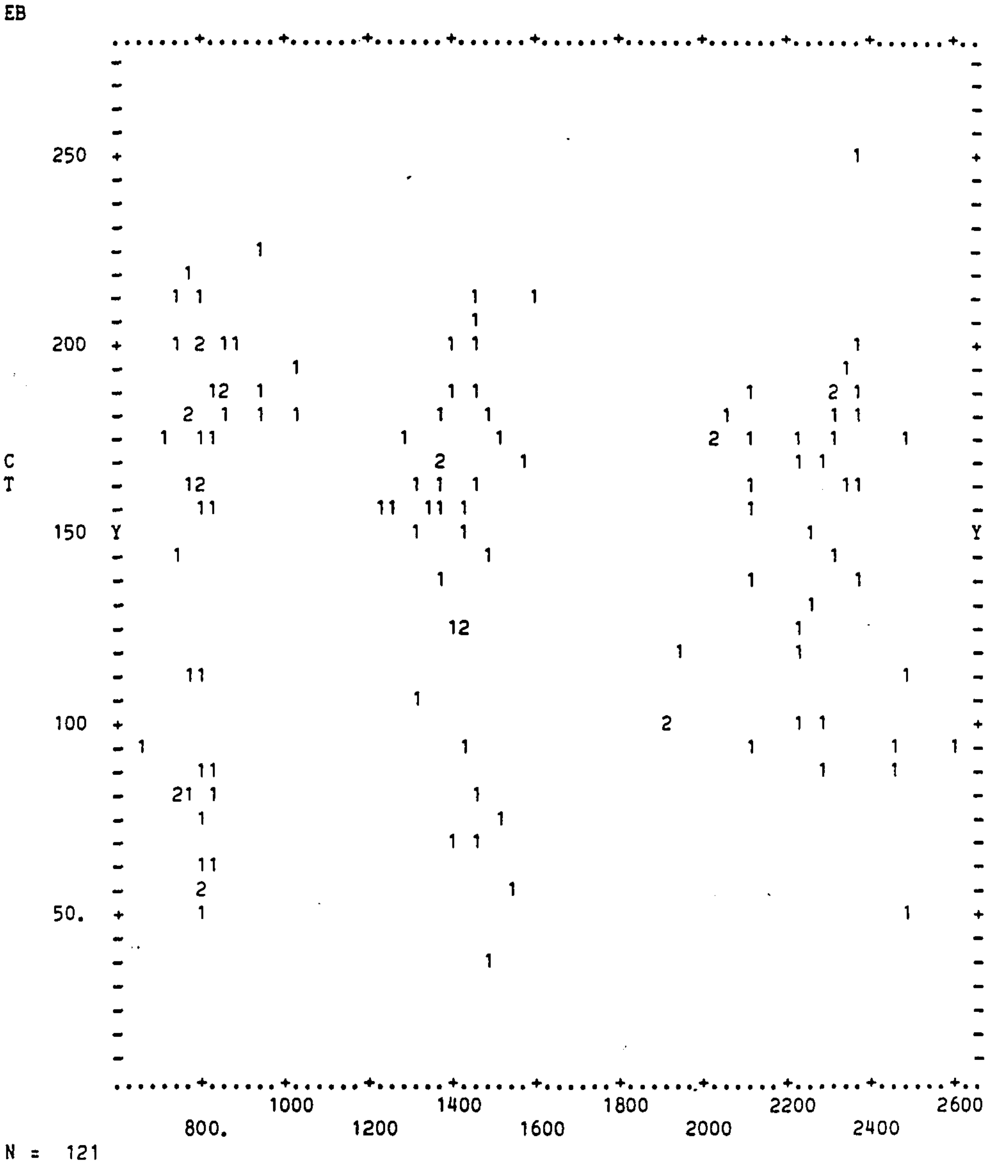
Respiratory Activity: In order to assess the effects of anticipated utterance length on respiratory events, values for depth of inspiration (calculated by subtracting the preceding trough from the peak) were derived from the summed Respitrace data for all tokens. These values were obtained only for subject RC, as the Respitrace data for EB were unusable.

Figure 13 shows the averaged Respitrace curves in the region of the initial inspiration for the three length conditions for RC and RC English. The corresponding values for depth of inspiration are shown in Table 6. Values for the individual stress conditions are shown in Appendix I.



```
N =   45
R = .058                              DURATION (MSEC)
P = .705
                                          MEAN    S.D.
--REGRESSION LINE--  -RES.MS-     X   1704.6  688.58
Y= 348.99 +.00868*X    10715.     Y   363.78  102.51
```

Figure 11. Initial peak CT activity (in microvolts) as a function of utterance length for RC English.



```
N =  121
R = .006                               DURATION (MSEC)
P = .949
                                           MEAN    S.D.
--REGRESSION LINE--   -RES.MS-     X   1486.0   613.96
Y= 146.85 +460E-6*X     2289.1     Y   147.54   47.645
```

Figure 12. Initial peak CT activity (in microvolts) as a function of utterance length for EB.

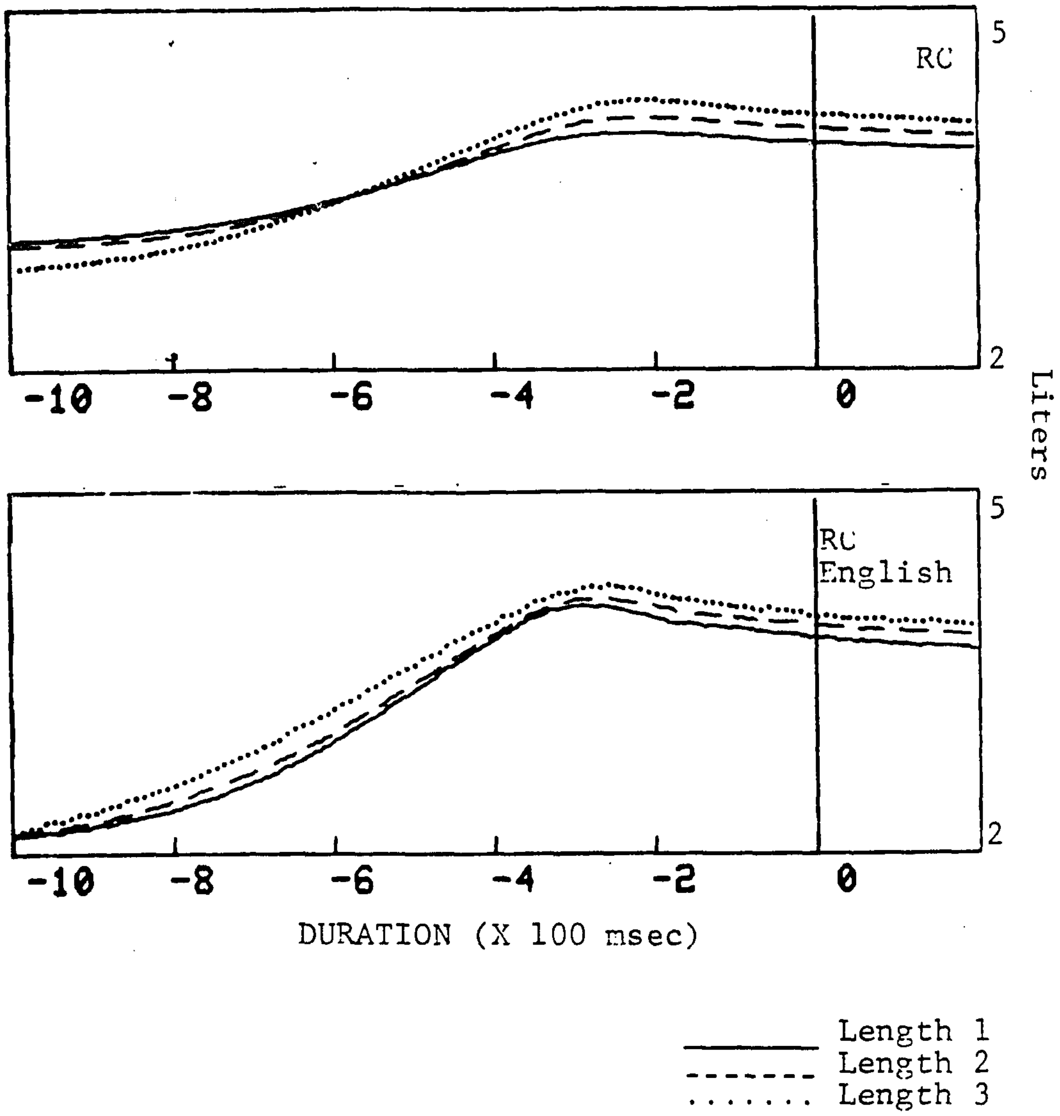


Figure 13. Depth of initial inspiration across utterance lengths.

Table 6

Depth of Initial Inspiration (in Liters) as a Function of Utterance Length

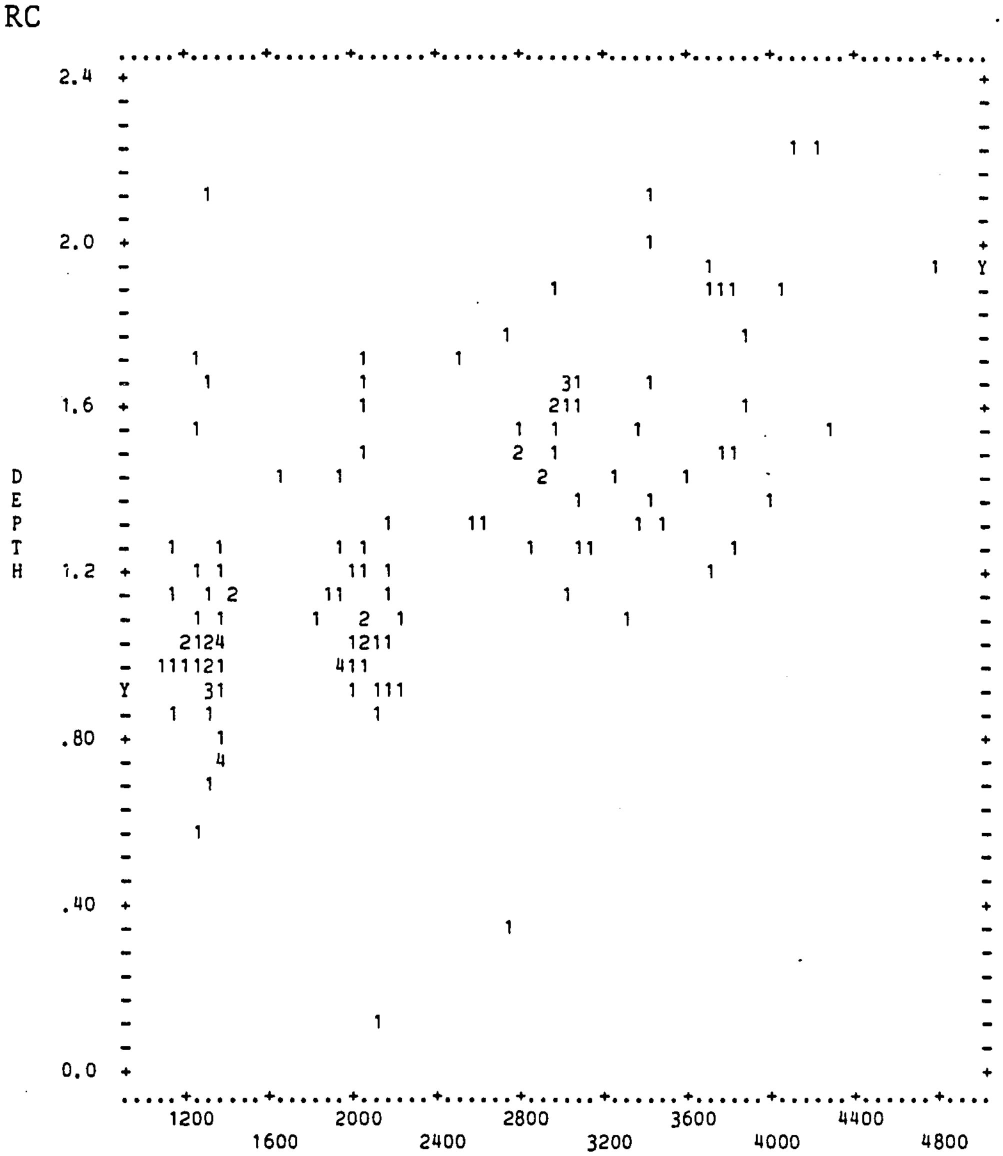
	Range	Mean	SD	N	
Length 1	.59-1.64	1.05	.28	43	
Length 2	.86-1.77	1.22	.25	44	RC
Length 3	1.09-2.24	1.59	.28	44	

	Range	Mean	SD	N	
Length 1	.87-1.27	1.10	.13	15	
Length 2	.79-1.89	1.17	.29	16	RC English
Length 3	.89-1.52	1.26	.21	15	

For both RC and RC English, the collapsed data show a systematic increase in depth of inspiration as a function of utterance length (Table 6), although, again, the magnitude of this increase is greater for the Dutch utterances and their reiterant forms (i.e, RC). For RC, the increase is significant at each length increment (L1 vs. L2: $t=2.989$, $p<.01$; L2 vs. L3: $t=6.538$, $p<.001$) while, for RC English, depth of inspiration shows a significant increase only between extreme lengths (L1 vs. L2: $t=.857$, $p>.1$; L2 vs. L3: $t=.984$, $p>.1$; L1 vs. L3: $t=2.509$, $p<.02$).

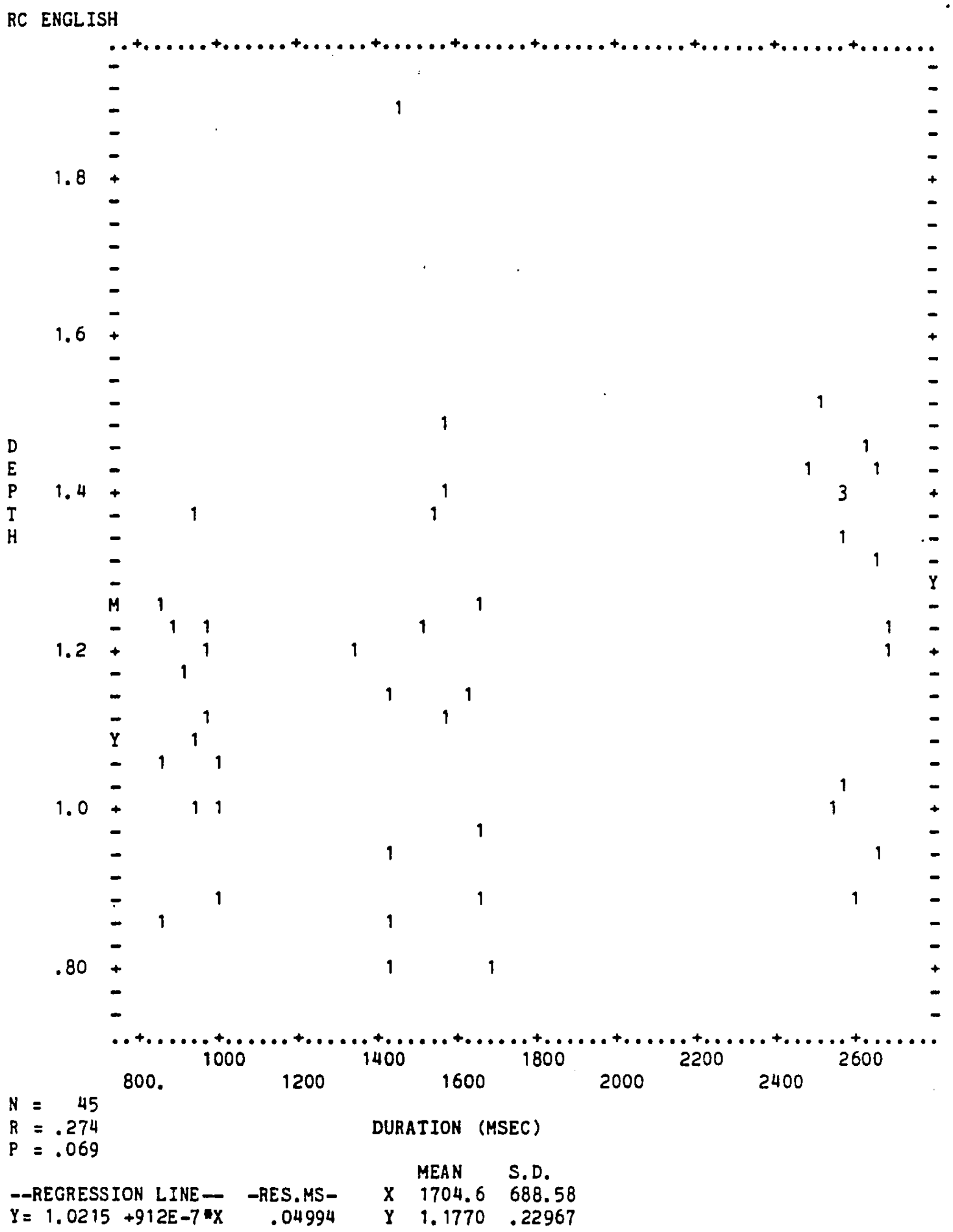
An analysis of variance reveals a significant interaction of depth of inspiration with utterance length for RC ($F=43.9$, $p<.0001$) but not for RC English ($F=2.95$, $p>.05$). This result can be predicted from the values for the individual stress conditions (Appendix I), which show a fairly systematic increase in inspiratory depth as a function of utterance length for RC, but far less consistent length variations for RC English.

It is also not surprising that, for RC, depth of inspiration shows a significant overall correlation with utterance length ($r=.639$, $p<.001$) while, for RC English, depth of inspiration fails to correlate with utterance length ($r=.274$, $p>.05$) (Figs. 14 & 15). Again, these overall correlations can be predicted from those for the individual conditions for both RC and RC English (App. J).



```
N =  131
R = .639
P < .001
                                     MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X  2306.9  940.22
Y= .67861 +255E-6*X    .08427    Y  1.2674  .37578
```

Figure 14. Initial depth of inspiration (in liters) as a function of utterance length for RC.



```
N =   45
R = .274
P = .069
                                        MEAN    S.D.
--REGRESSION LINE--  -RES.MS-      X  1704.6  688.58
Y= 1.0215 +912E-7*X    .04994      Y  1.1770  .22967
```

Figure 15. Initial depth of inspiration (in liters) as a function of utterance length for RC English.

The results in this section have shown that subjects may or may not make adjustments in initial conditions on the basis of utterance length, even when the utterances are known to them in advance. Subject RC, for example, showed systematic (and usually significant) increases in initial values for F0, Ps, CT and depth of inspiration with increasing utterance length. For RC English, length-dependent adjustments, while often present for these variables, were far less systematic. Subject EB, on the other hand, failed to show any systematic manipulation of initial F0, CT and Ps values in the face of significant variations in utterance length.

The acoustic results on initial peak F0 are thus in keeping with a literature that finds both the presence (i.e., Bruce, 1982; Cooper & Sorenson, 1981) and absence (i.e., Maeda, 1976) of a systematic length effect. To date, there are no studies that have looked at the effect of utterance length on initial Ps and CT activity, although, since they are both mechanisms by which F0 may be raised, it is reasonable to assume that there should be at least a correlational, if not a causal, relationship between these variables and F0. A more extensive analysis of the nature of this relationship is discussed in the sections that follow.

The finding that the depth of the initial inspiration does not always serve as a consistent index of anticipated utterance length has also been found by Horii and Cooke (1978) for subjects participating in an oral paragraph-reading task. They conclude that "typically, oral reading is done well within a respiratory capability (near equilibrium) and does not usually require special modification of respiratory maneuvers that are dependent on the length of the subsequent utterance. These data thus support a notion of semiindependence of the respiratory system to speech production such that depth of inspiration is unrelated to the subsequent utterance length, at least in the oral reading task" (p.477).

2. The Effect of Stress Configuration

Because increased F0 and amplitude are acoustic correlates of stress (Fry, 1955), it is expected that the early portions of utterances receiving contrastive stress (i.e., Early and Double) will show higher values for fundamental frequency, subglottal pressure and cricothyroid activity than will the same syllables receiving only lexical stress (i.e., Neutral and/or Late).

For both subjects, when the data are collapsed across tokens having the same initial stress configuration in common (i.e., Early stress or -Early stress) (Tables 7-9), peaks for Early and Double stress tokens (i.e., Early) are significantly higher than those for Neutral and/or Late stress tokens (i.e., -Early) for F0, Ps and CT (Table 10). Figures 16-18 contrast the initial F0, Ps and CT peaks for the Early and -Early utterances for RC, RC English and EB, respectively. For reasons related to difficulties involved in averaging extracted tokens of unequal lengths, the F0 data in these figures show the Early, Double and Late peaks separately. For the Ps

Table 7

The Effect of Emphatic Stress on Peak Fundamental Frequency

	+Early			-Early			
	Mean	**SD**	**N**	**Mean**	**SD**	**N**	
Length 1	138	5.32	29	118	6.56	14	
Length 2	147	6.66	29	121	3.32	15	RC
Length 3	166	5.59	30	139	4.89	14	
Mean	150	13.13	88	126	10.41	43	
Length 1	166	7.32	10	138	8.44	5	
Length 2	177	14.31	10	135	3.25	6	RC English
Length 3	184	8.88	10	130	3.91	5	
Mean	176	12.71	30	134	6.11	16	
Length 1	275	35.83	24	186	20.72	19	
Length 2	243	36.49	25	192	18.56	12	EB
Length 3	272	33.10	25	189	10.79	16	
Mean	263	17.67	74	189	17.12	47	

Table 8

The Effect of Emphatic Stress on Peak Subglottal Pressure

	+Early			-Early			
	Mean	SD	N	Mean	SD	N	
Length 1	9.61	.77	29	8.44	.70	14	
Length 2	10.57	.86	29	8.29	.56	15	RC
Length 3	11.09	1.02	30	9.55	.99	14	
Mean	10.42	1.06	88	8.76	.93	43	
Length 1	11.53	1.20	10	8.76	.71	5	
Length 2	12.51	1.31	10	8.85	.53	6	RC English
Length 3	13.00	.73	10	9.34	.57	5	
Mean	12.35	1.24	30	8.98	.62	16	
Length 1	12.22	1.46	22	9.02	1.60	15	
Length 2	11.69	1.86	23	8.88	1.13	9	EB
Length 3	12.31	1.38	24	9.15	.97	13	
Mean	12.07	1.58	69	9.02	1.27	37	

Table 9

The Effect of Emphatic Stress on Peak Cricothyroid Muscle Activity

	+Early			-Early			
	Mean	SD	N	Mean	SD	N	
Length 1	311	67.53	29	249	26.34	14	
Length 2	325	55.14	29	245	38.05	15	RC
Length 3	394	59.82	30	315	97.71	14	
Mean	343	72.70	88	270	70.32	43	
Length 1	372	76.66	10	331	100.99	5	
Length 2	417	55.78	10	257	71.85	6	RC English
Length 3	438	79.77	10	222	79.37	5	
Mean	409	74.60	30	270	90.62	16	
Length 1	188	17.81	24	100	43.48	19	
Length 2	170	25.11	25	115	41.95	12	EB
Length 3	175	25.05	25	109	29.74	16	
Mean	178	23.84	74	108	38.30	47	

Table 10

Significance of Differences Between +Early and -Early Stress Conditions

	F0		*Ps*		*CT*	
	t	*p*	*t*	*p*	*t*	*p*
RC	10.478	<.001	8.751	<.001	5.454	<.001
RC English	12.427	<.001	10.176	<.001	5.583	<.001
EB	22.723	<.001	10.113	<.001	12.402	<.001

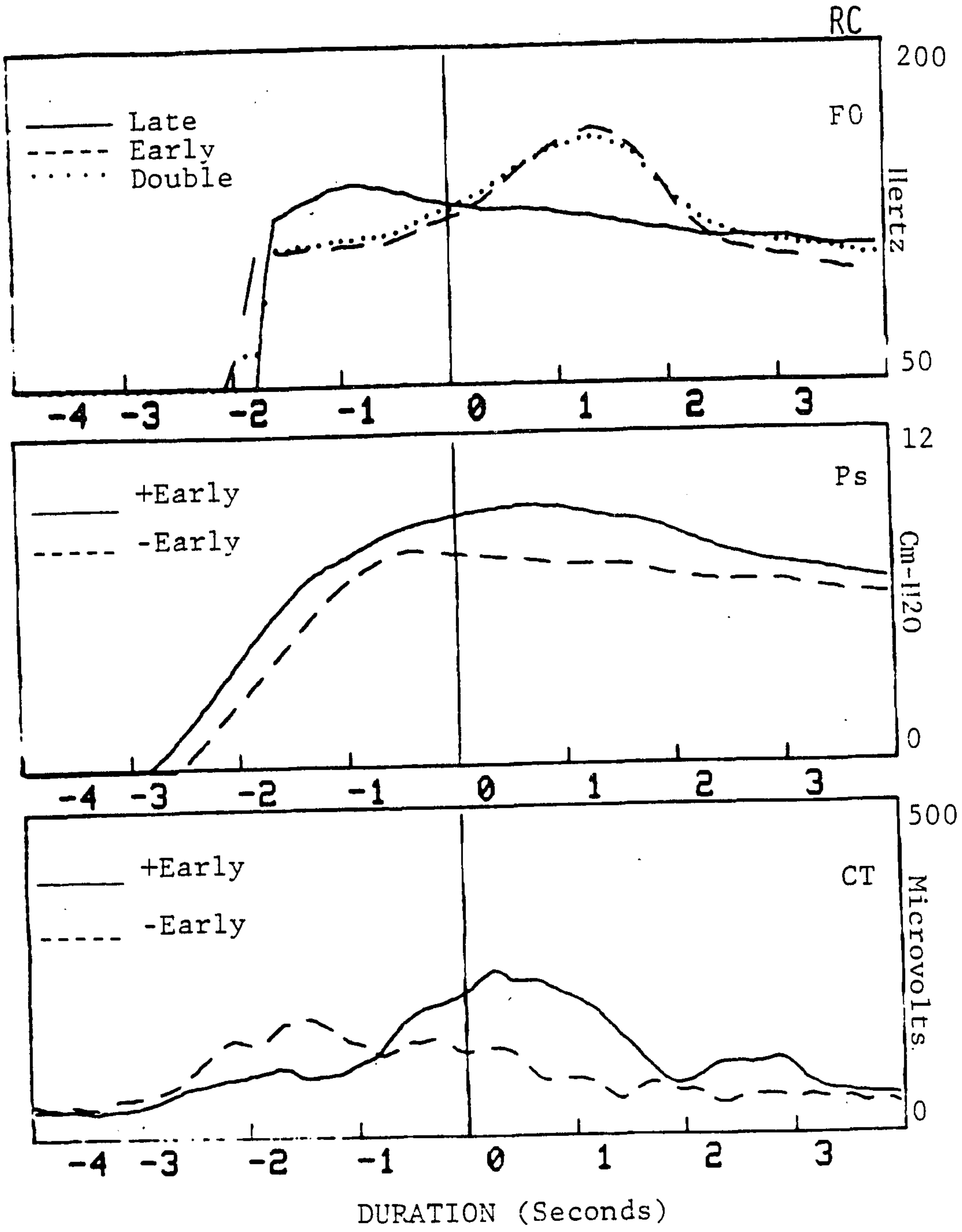


Figure 16. Comparison of initial stressed vs. unstressed syllables for RC.

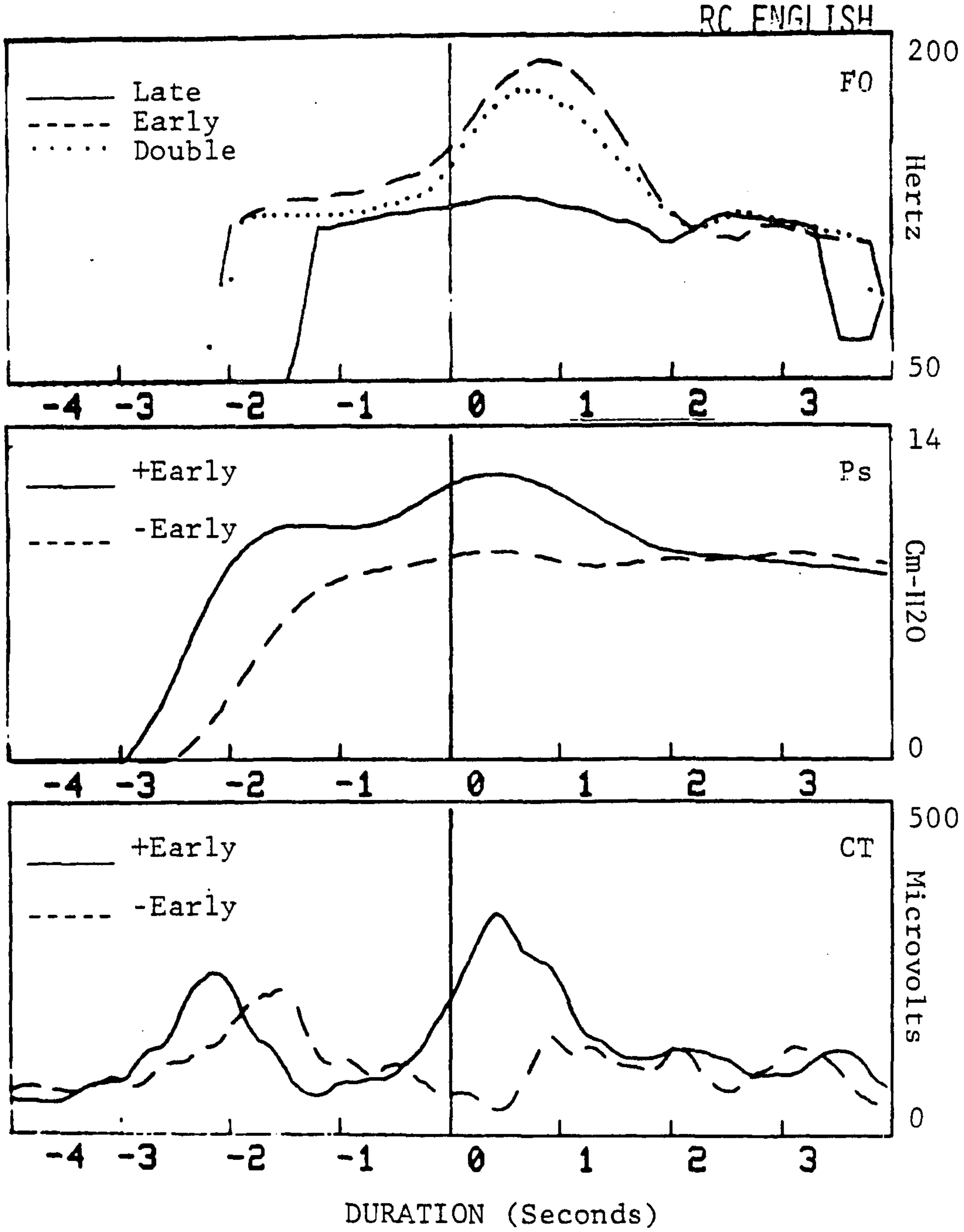


Figure 17. Comparison of initial stressed vs. unstressed syllables for RC English.

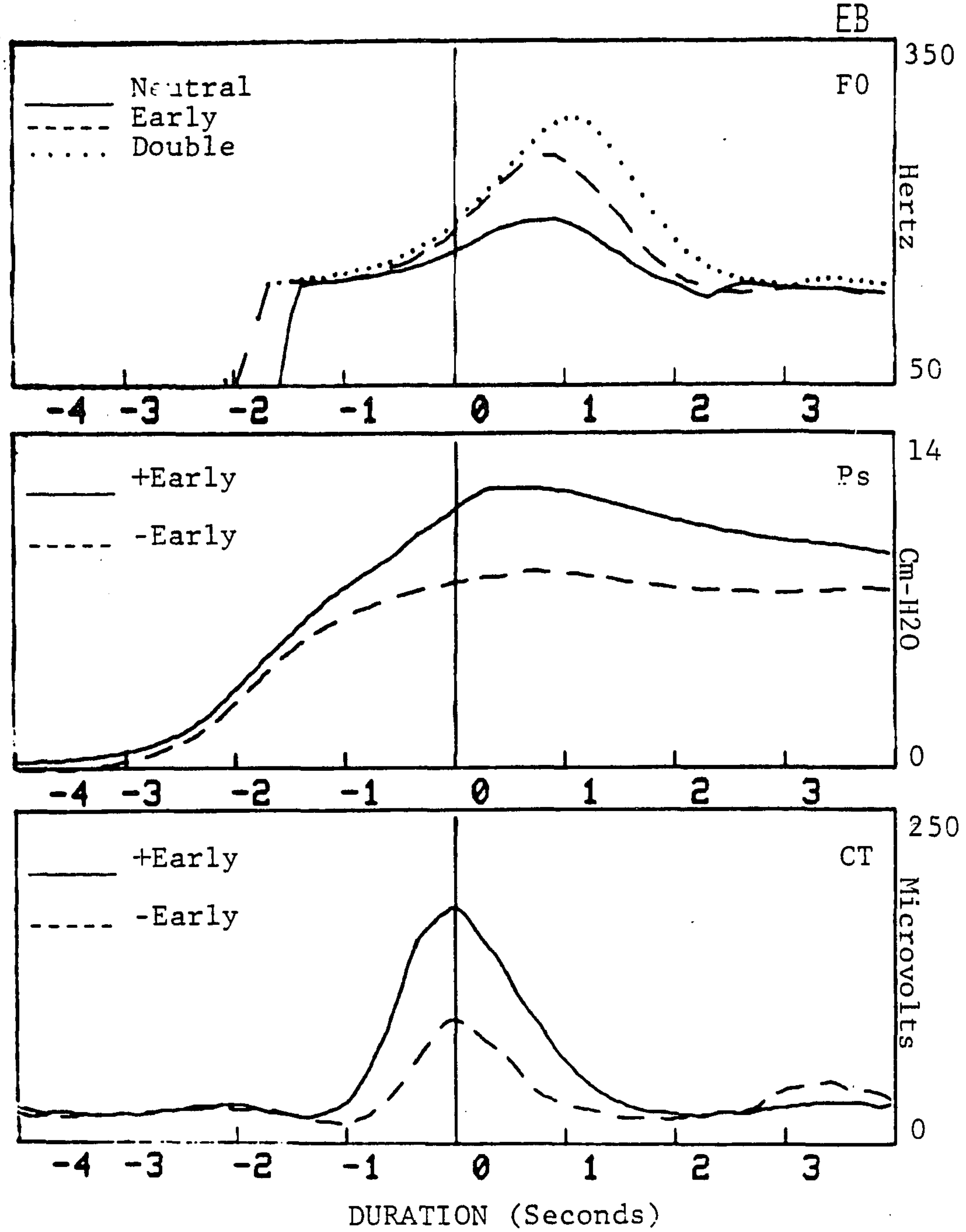


Figure 18. Comparison of initial stressed vs. unstressed syllables for EB.

and CT data, all tokens with the same initial stress configuration have been averaged together.

Analyses of variance show significant effects of stress condition for the three variables as well (Table 11). These interactions can be predicted from Tables 7 through 9 and Figures 19 through 27, which show that, for each variable, the lowest part of the range is occupied by the Late (and Neutral for EB) stress tokens, while the Early and Double stress tokens are distributed through the upper part of the range.

3. Correlating Fundamental Frequency with Subglottal Pressure and Cricothyroid Muscle Activity

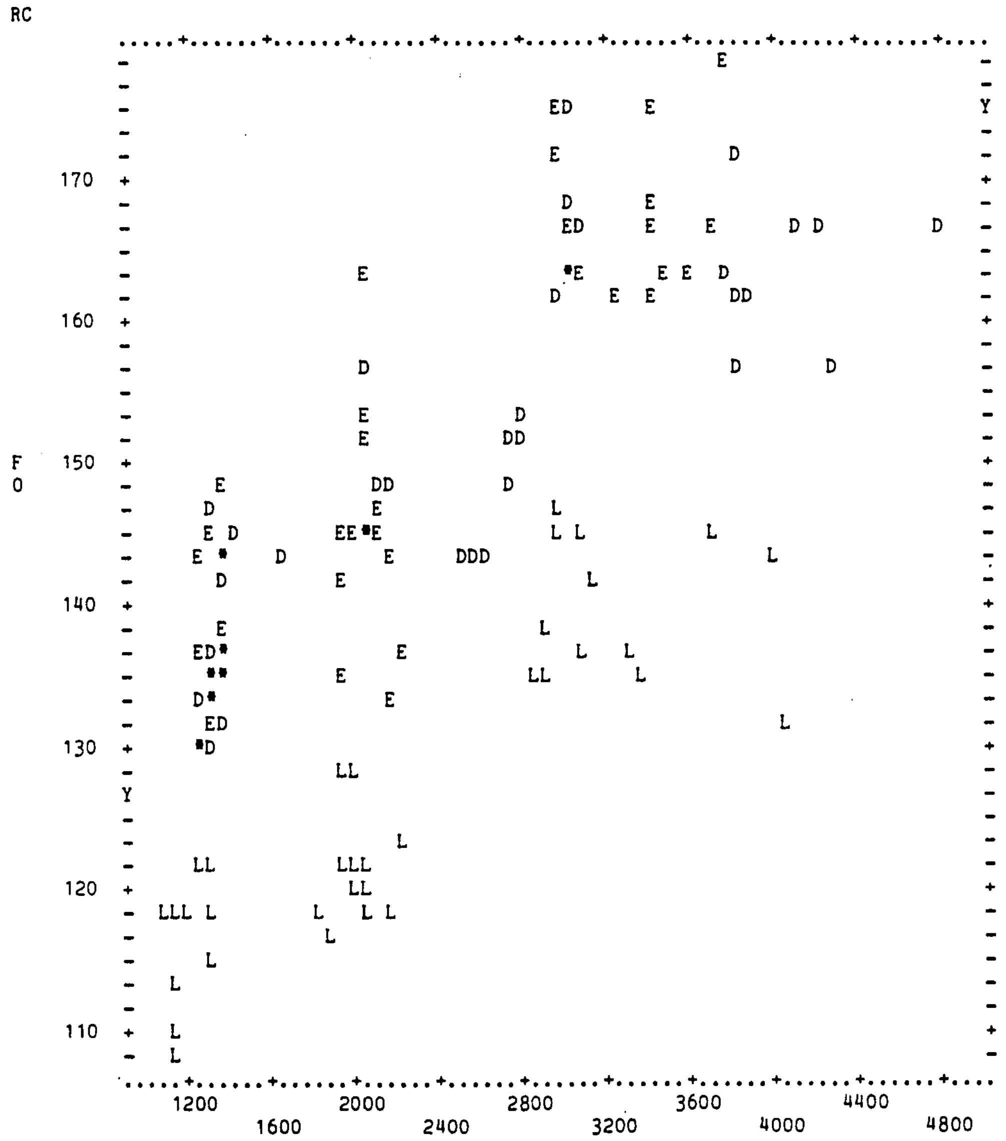
Up to this point, individual dependent variables have been considered in relation to their interaction with the independent variables of length and stress configuration. In this section, additional analyses were performed in order to relate the dependent variables to each other; in particular, F0 with Ps and CT activity.

F0 vs. Ps: Plots of F0 versus Ps for all tokens of all utterances are shown in Figures 28-30 for RC, RC English and EB, respectively. Correlation coefficients for these comparisons are highly significant (RC: $r=.786$, $p<.001$; RC English: $r=.937$, $p<.001$; EB: $r=.690$, $p<.001$). Correlations for both subjects for the individual conditions are shown in Appendix K. For RC, all correlations between Ps and F0 are

Table 11

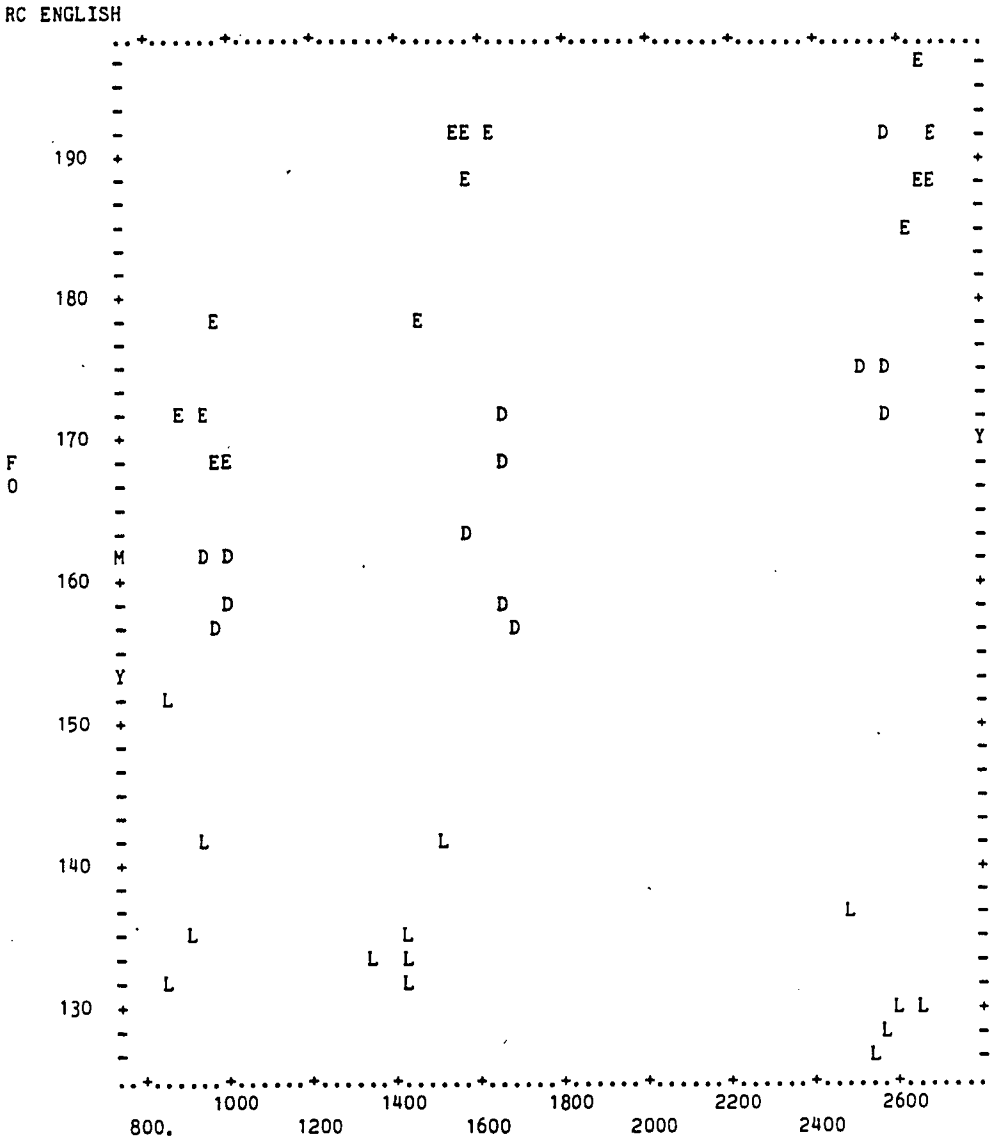
Significance of Interaction of Stress Condition and Initial Peak Values

	F0		*Ps*		*CT*	
	F	*p*	*F*	*p*	*F*	*p*
RC	365.88	<.001	135.74	<.001	23.41	<.001
RC English	322.18	<.001	200.36	<.001	27.77	<.001
EB	48.25	<.001	33.39	<.001	53.43	<.001



```
N =  131
R = .678
P < .001
                                                 MEAN    S.D.
--REGRESSION LINE--  -RES.MS-     X  2306.9  940.22
Y= 114.80 +.01207*X     152.74     Y  142.65  16.744

GROUP=EARLY    , SYMBOL=E
GROUP=DOUBLE   , SYMBOL=D
GROUP=LATE     , SYMBOL=L
```

Figure 19. Initial peak F0 (in Hertz) as a function of stress configuration for RC.

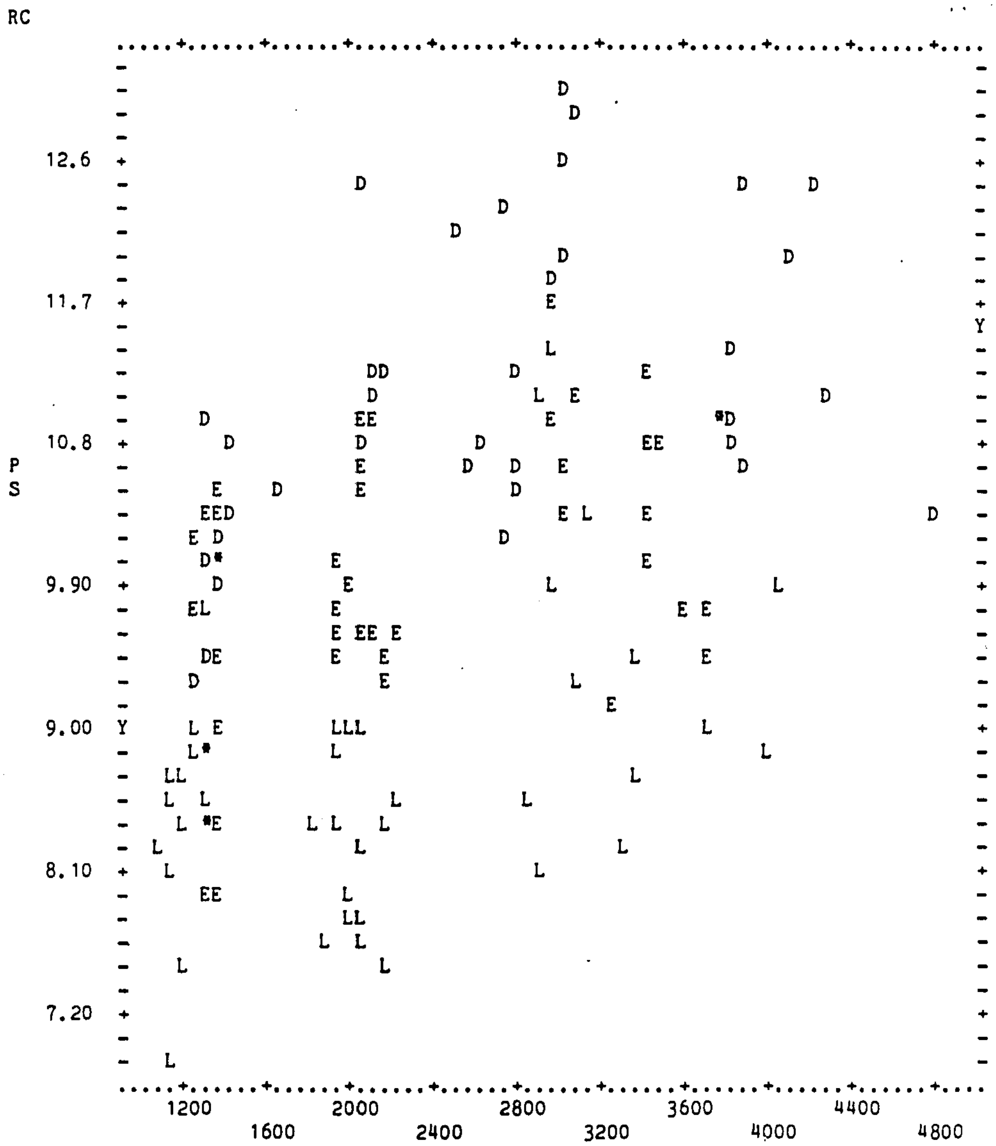


```
N =   45
R = .232                          DURATION (MSEC)
P = .125
                                  MEAN     S.D.
--REGRESSION LINE--  -RES.MS-   X  1704.6   688.58
Y= 148.17 +.00766*X    498.95   Y  161.22   22.703

GROUP=EARLY    , SYMBOL=E
GROUP=DOUBLE   , SYMBOL=D
GROUP=LATE     , SYMBOL=L
```

<u>Figure 20</u>. Initial peak F0 (in Hertz) as a function of stress configuration for RC English.

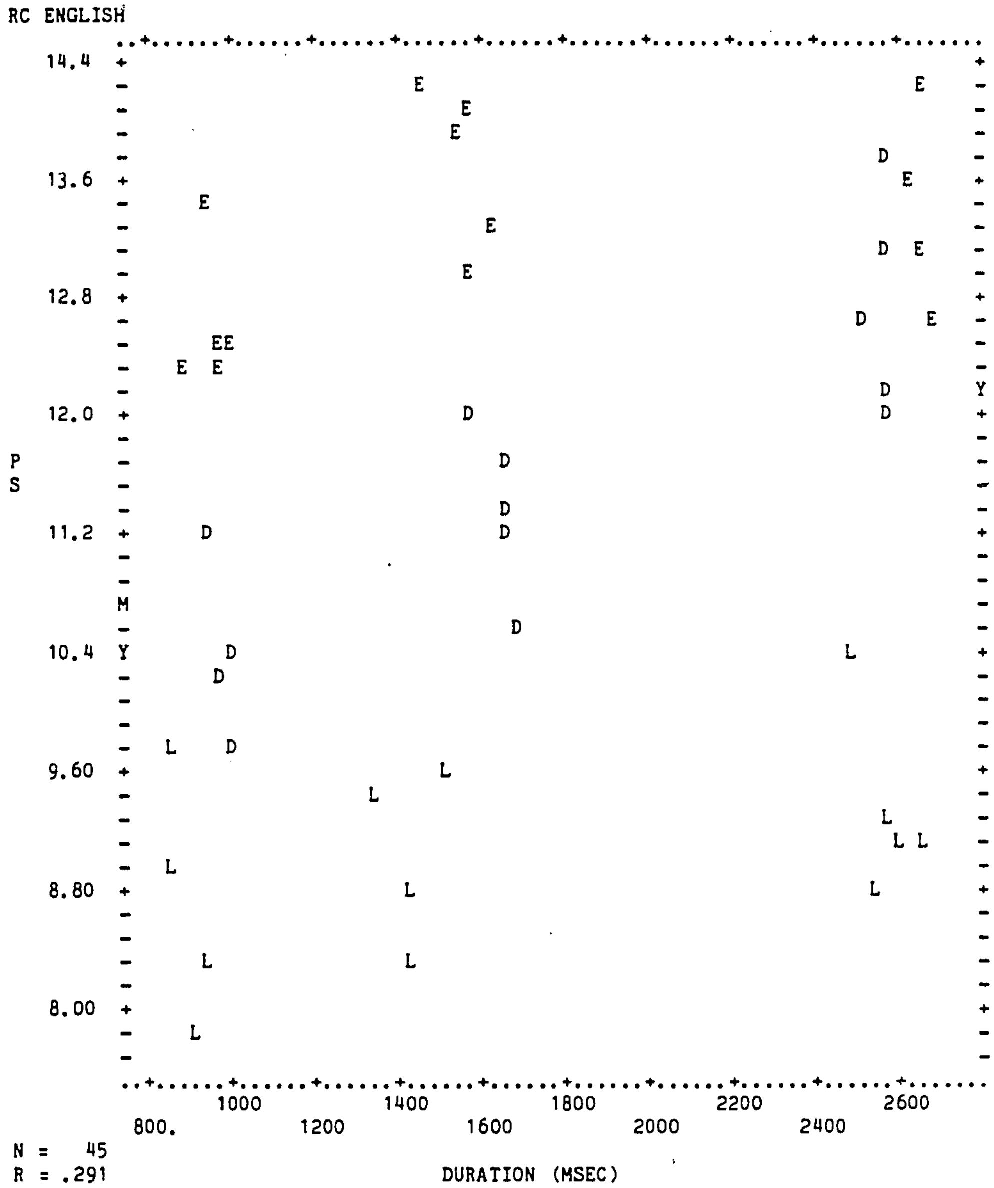
```
EB
          .......+......+......+......+......+......+......+......+......+......+..
     330  +                                                                       +
          -                                                                       -
          -    E E D     D               D                            D           -
          -                                                                       -
          -     EE E                                                  D           -
          -                                                                       -
          -        D  E                                     D  D       DE         -
     300  +                                                                       +
          -     E                     *                            D    E         -
          -                                                                       -
          -      E                E     D                    D D      D D         -
          -                                                                       -
          -      D                                                 E  E           -
          -                                                                       -
     270  +                           E  D D                          E           +
          -                                                                       -
          -    E                        E                   D                     -
F         -     E   D                D  E                                         -
O         -                                                                       -
          -      E                                             E                  -
          -        D                                                    ED        -
     240  +       E   E                D                                D         +
          -                                                                       Y
          Y              D      E      E     D                                    -
          -   END L   E              D      D                  D    D             -
          -                           EDD                             D D         -
          -                         EE                                            -
          -      N                     L D                                        -
     210  +                        L                               D              +
          -      L               D                                          N     -
          -                                                                 N     -
          -                                                    L   L            N -
          -                        N N                  L           N      N      -
          -       L                                                N              -
          -       L                                                               -
     180  +    LLLN                      N N             NL           N           +
          - N                         N                               N           -
          -     NN                                                          L     -
          -    N  L                    NN                                    N    -
          -      *                         N                                      -
          -                              N                                        -
          -      N                    D                                           -
     150  +                                                                       +
          .......+......+......+......+......+......+......+......+......+......+..
                       1000          1400          1800          2200          2600
                800.          1200          1600          2000          2400
N =  121
R = .040                          DURATION (MSEC)
P = .664
                                     MEAN    S.D.
--REGRESSION LINE--  -RES.MS-     X  1486.0  613.96
Y= 228.39 +.00322*X     2463.0    Y  233.17  49.461

GROUP=NEUTRAL , SYMBOL=N
GROUP=EARLY   , SYMBOL=E
GROUP=DOUBLE  , SYMBOL=D
GROUP=LATE    , SYMBOL=L
```

<u>Figure 21</u>. Initial peak F0 (in Hertz) as a function of stress configuration for EB.



```
N =  131
R = .454
P < .001
                                      DURATION (MSEC)

                                    MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X  2306.9  940.22
Y= 8.4325 +626E-6*X    1.3409    Y  9.8755  1.2948

GROUP=EARLY   , SYMBOL=E
GROUP=DOUBLE  , SYMBOL=D
GROUP=LATE    , SYMBOL=L
```

Figure 22. Initial peak Ps (in Cm-H2O) as a function of stress configuration for RC.

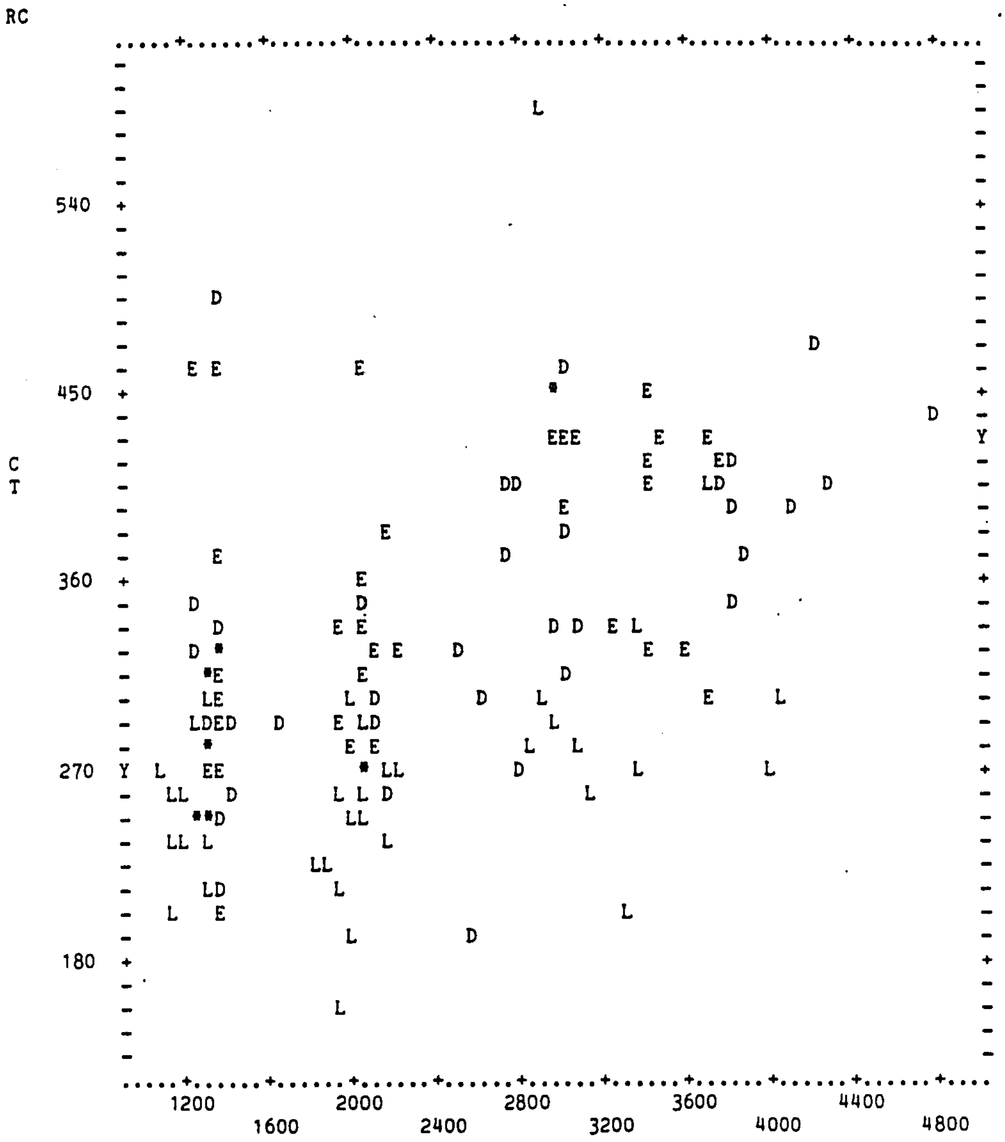


```
N =   45
R = .291
P = .052

                                   MEAN     S.D.
--REGRESSION LINE--  -RES.MS-    X  1704.6   688.58
Y= 9.7730 +828E-6*X    3.5903    Y  11.184   1.9579

GROUP=EARLY    , SYMBOL=E
GROUP=DOUBLE   , SYMBOL=D
GROUP=LATE     , SYMBOL=L
```

<u>Figure 23</u>. Initial peak Ps (in Cm-H2O) as a function of stress configuration for RC English.

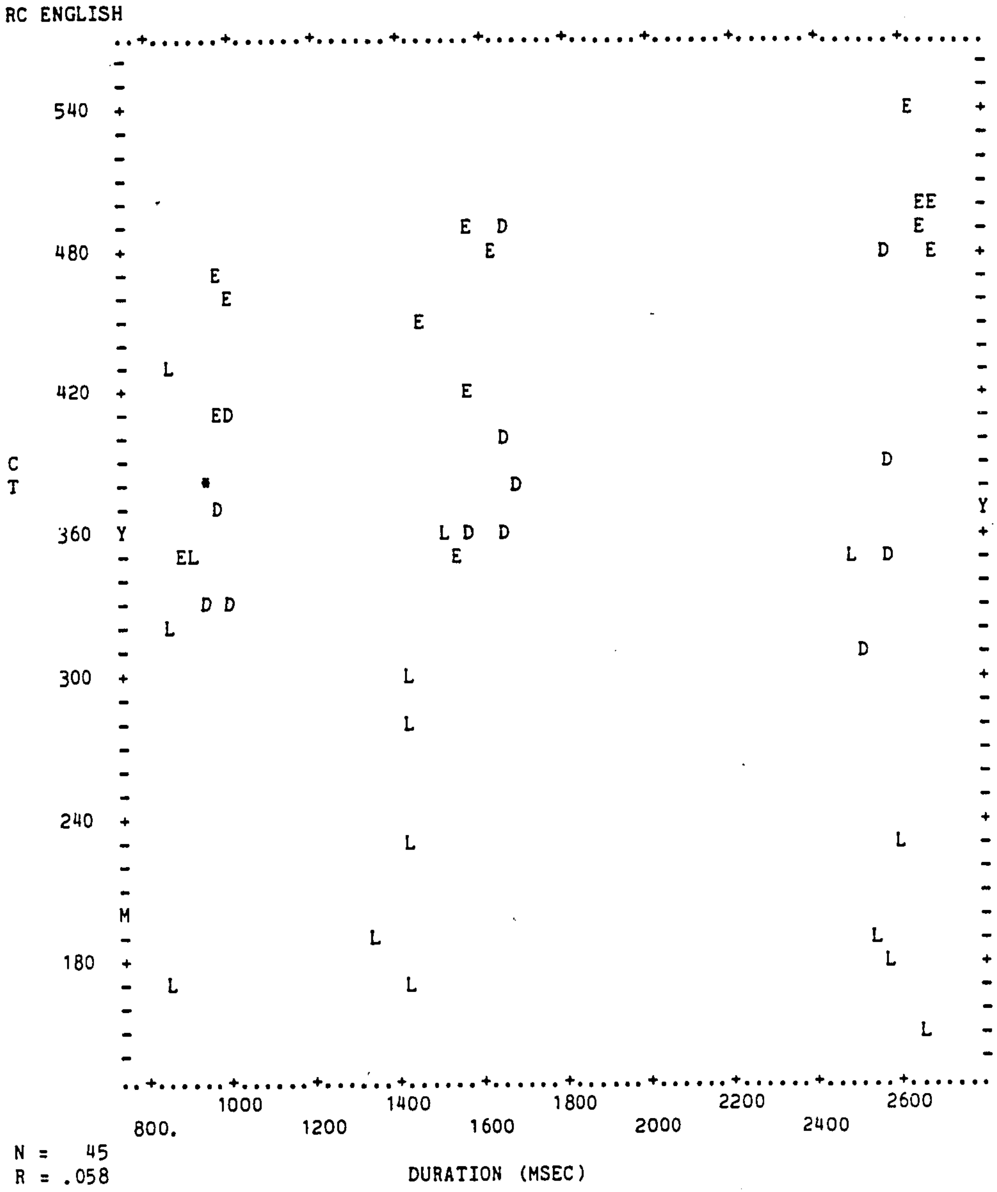
```
EB
         .......+......+......+......+......+......+......+......+......+......+..
         -                                                                       -
   14.4  +                                                                       +
         -      E                                                    D           -
         -           E         E           D                         D E         -
         -           E                                          D        D         -
         -                            ED                             D           -
   13.5  +       L                     D                                         +
         -              D           E  E                           E              -
         -        D                 D                         E     E            -
         -        D                                        D                     -
         -        D                         D                                    -
   12.6  +     E                     E                                 D         +
         -      E                E                                               -
         -                               D                                       -
         -     *                       D                              DE         -
         -      E                                                D              -
   11.7  +    E                                             D                    +
P        -              D              E                               D         -
S        -                                                             D         Y
         -      E                        N                                     N -
         -    E                                            D  D                  -
   10.8  Y         D                                              D  EE          +
         -       L                                                               -
         -    NE                     D                                           -
         -                                                                 N     -
         -                                                    D    D      N      -
    9.90 +                          DND                             N      N     +
         -      NE                   E                                           -
         -       L                                            L                  -
         -                      D N    N                           N             -
         -                          D   D                                        -
    9.00 +      L                  EE                         L   L              +
         -     N                                                  N N            -
         -   E   L                  N                                            -
         -    N                                                                  -
         -                                                                       -
    8.10 +                            NN                                   N     +
         -      *                         N                                      -
         -    L                                                                  -
         -                              N                                 L      -
         - N                                                                     -
    7.20 +      N                                                                +
         -
         ......MMMM....+......+...M..+M.M...+......+...MM.+......+...M..+......+..
                      1000          1400          1800          2200          2600
               800.          1200          1600          2000          2400
N =  106
R = .093                             DURATION (MSEC)
P = .343
                                    MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X  1509.2  625.51
Y= 10.529 +304E-6*X    4.1633    Y  10.987  2.0395

GROUP=NEUTRAL , SYMBOL=N
GROUP=EARLY   , SYMBOL=E
GROUP=DOUBLE  , SYMBOL=D
GROUP=LATE    , SYMBOL=L
```

Figure 24. Initial peak Ps (in Cm-H_2O) as a function of stress configuration for EB.



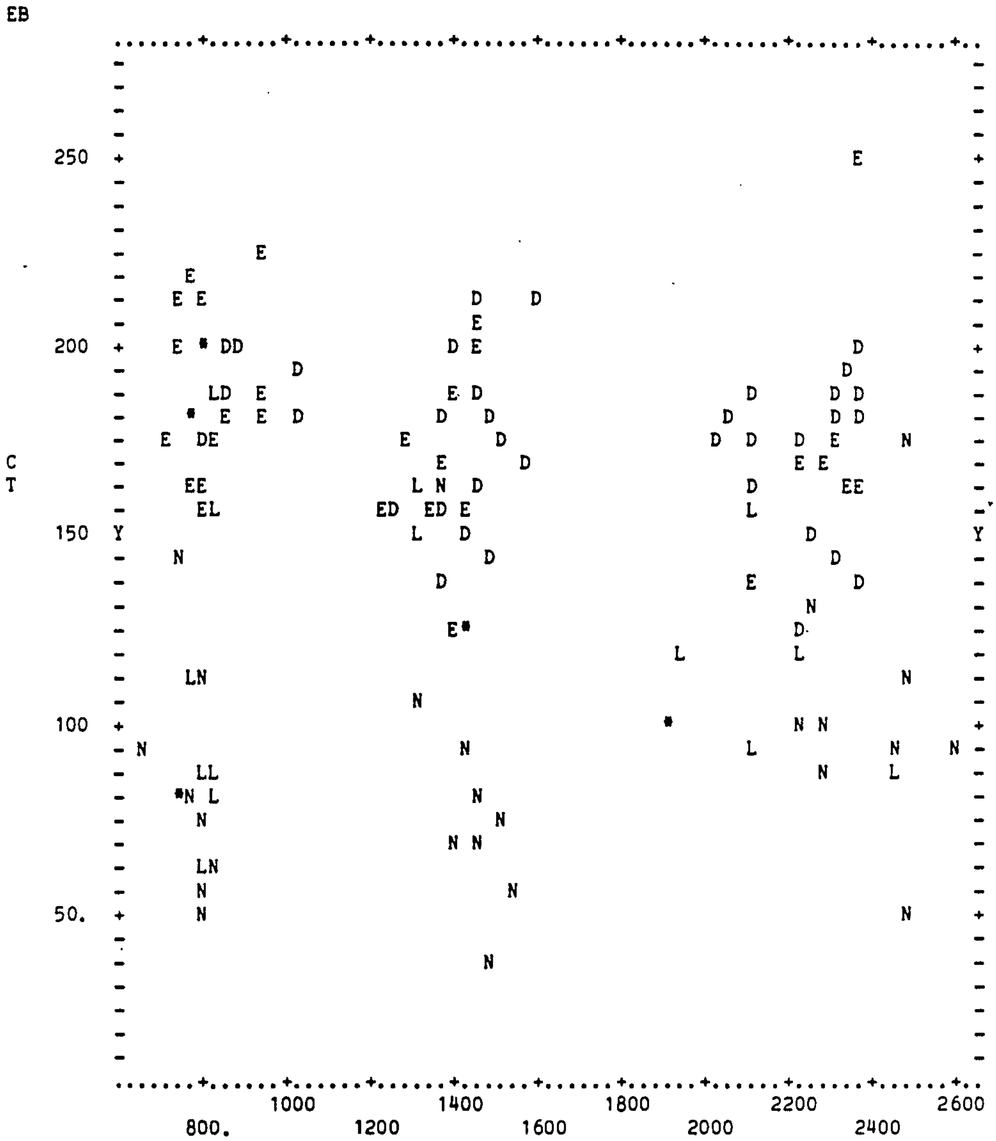
```
N =  131
R = .480                          DURATION (MSEC)
P < .001
                                     MEAN     S.D.
--REGRESSION LINE--  -RES.MS-    X  2306.9   940.22
Y= 231.30 +.03880*X    4479.1    Y  320.81   75.996

GROUP=EARLY    , SYMBOL=E
GROUP=DOUBLE   , SYMBOL=D
GROUP=LATE     , SYMBOL=L
```

Figure 25. Initial peak CT activity (in microvolts) as a function of stress configuration for RC.



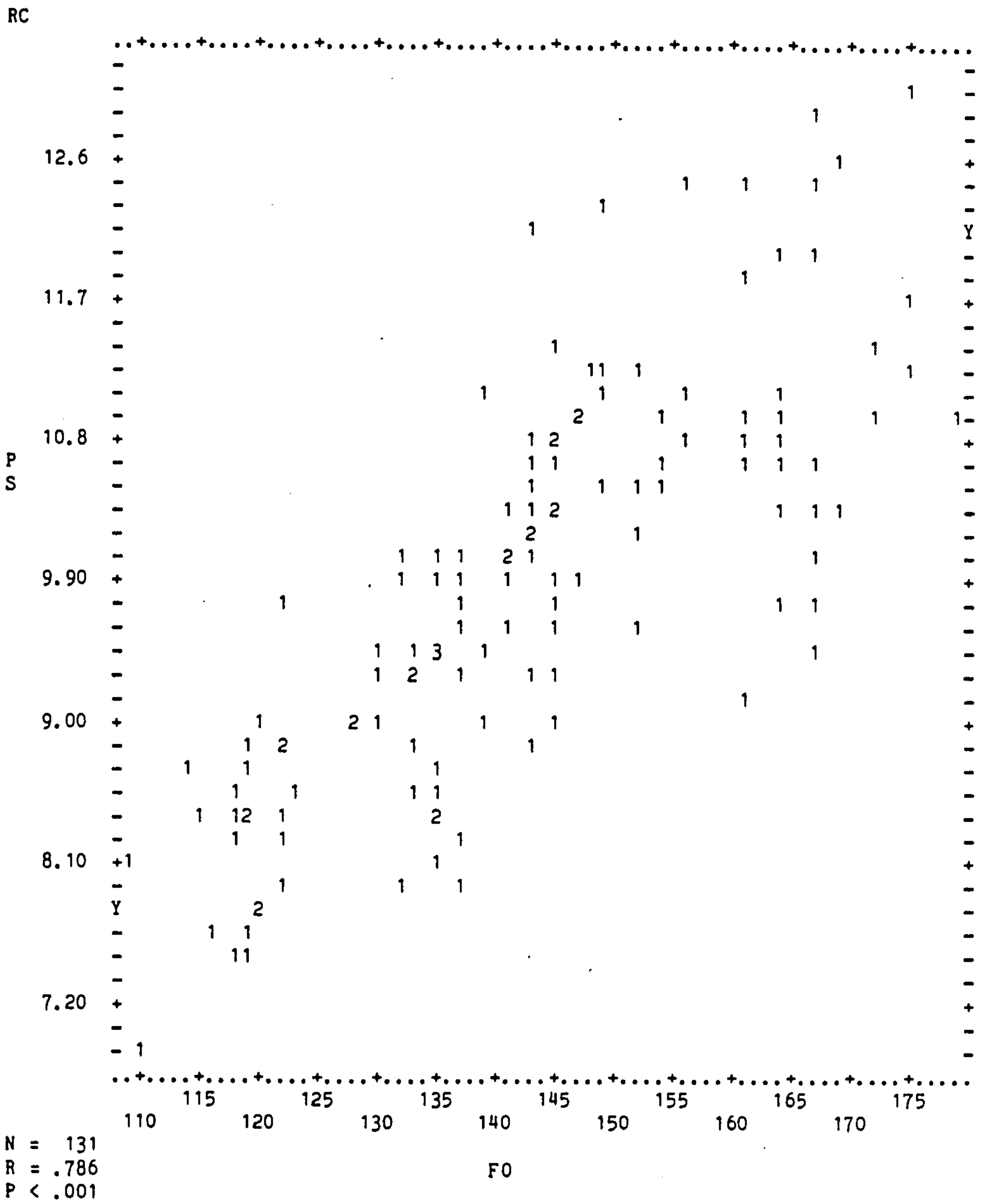
```
N =   45
R = .058
P = .705
                                          MEAN     S.D.
--REGRESSION LINE--  -RES.MS-       X   1704.6   688.58
Y= 348.99 +.00868*X    10715.       Y   363.78   102.51

GROUP=EARLY    , SYMBOL=E
GROUP=DOUBLE   , SYMBOL=D
GROUP=LATE     , SYMBOL=L
```

Figure 26. Initial peak CT activity (in microvolts) as a function of stress configuration for RC English.



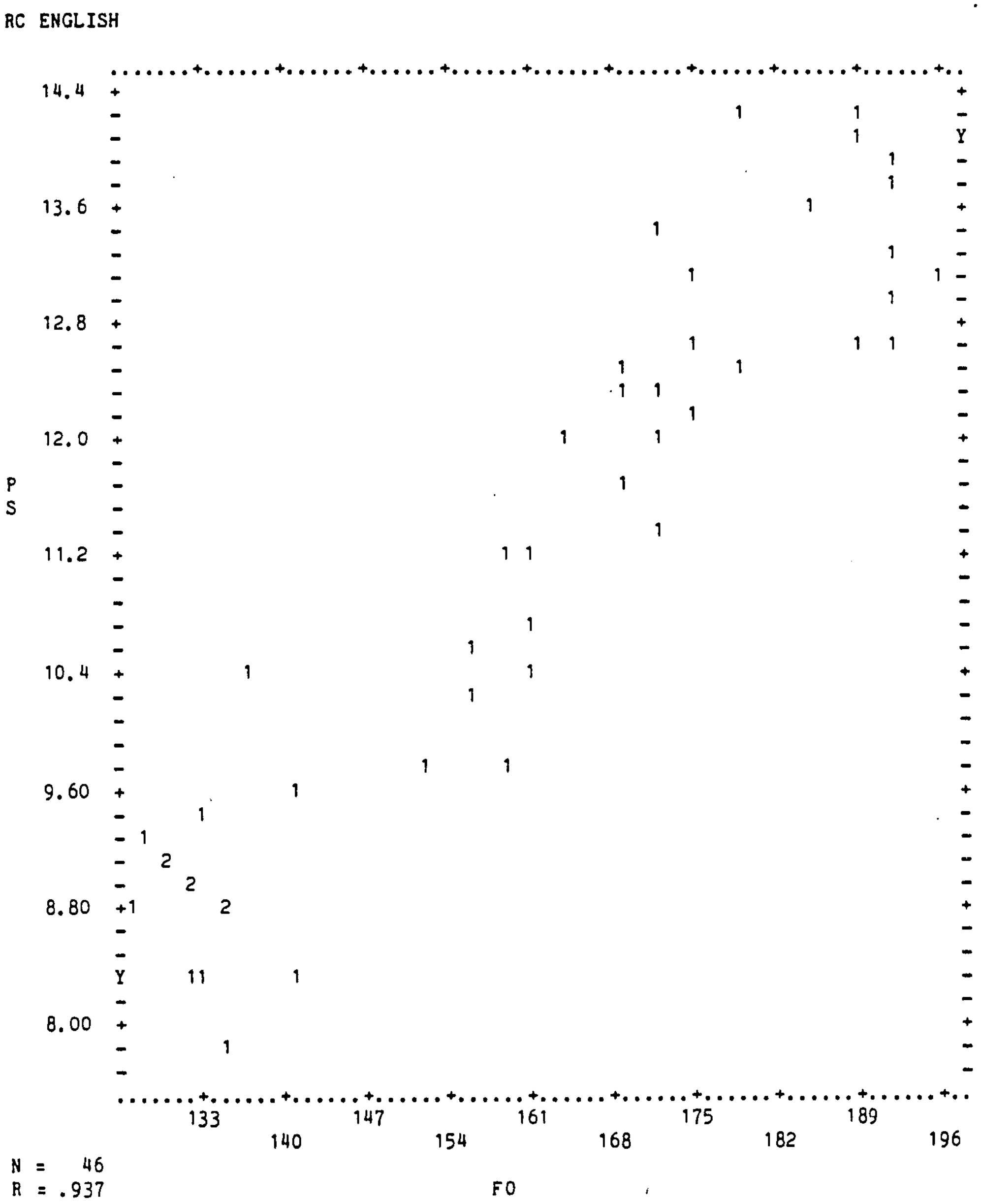
```
N =  121
R = .006
P = .949                              DURATION (MSEC)

                                      MEAN     S.D.
--REGRESSION LINE--  -RES.MS-    X  1486.0   613.96
Y= 146.85 +460E-6*X    2289.1    Y  147.54   47.645

GROUP=NEUTRAL , SYMBOL=N
GROUP=EARLY   , SYMBOL=E
GROUP=DOUBLE  , SYMBOL=D
GROUP=LATE    , SYMBOL=L
```

Figure 27. Initial peak CT activity (in microvolts) as a function of stress configuration for EB.



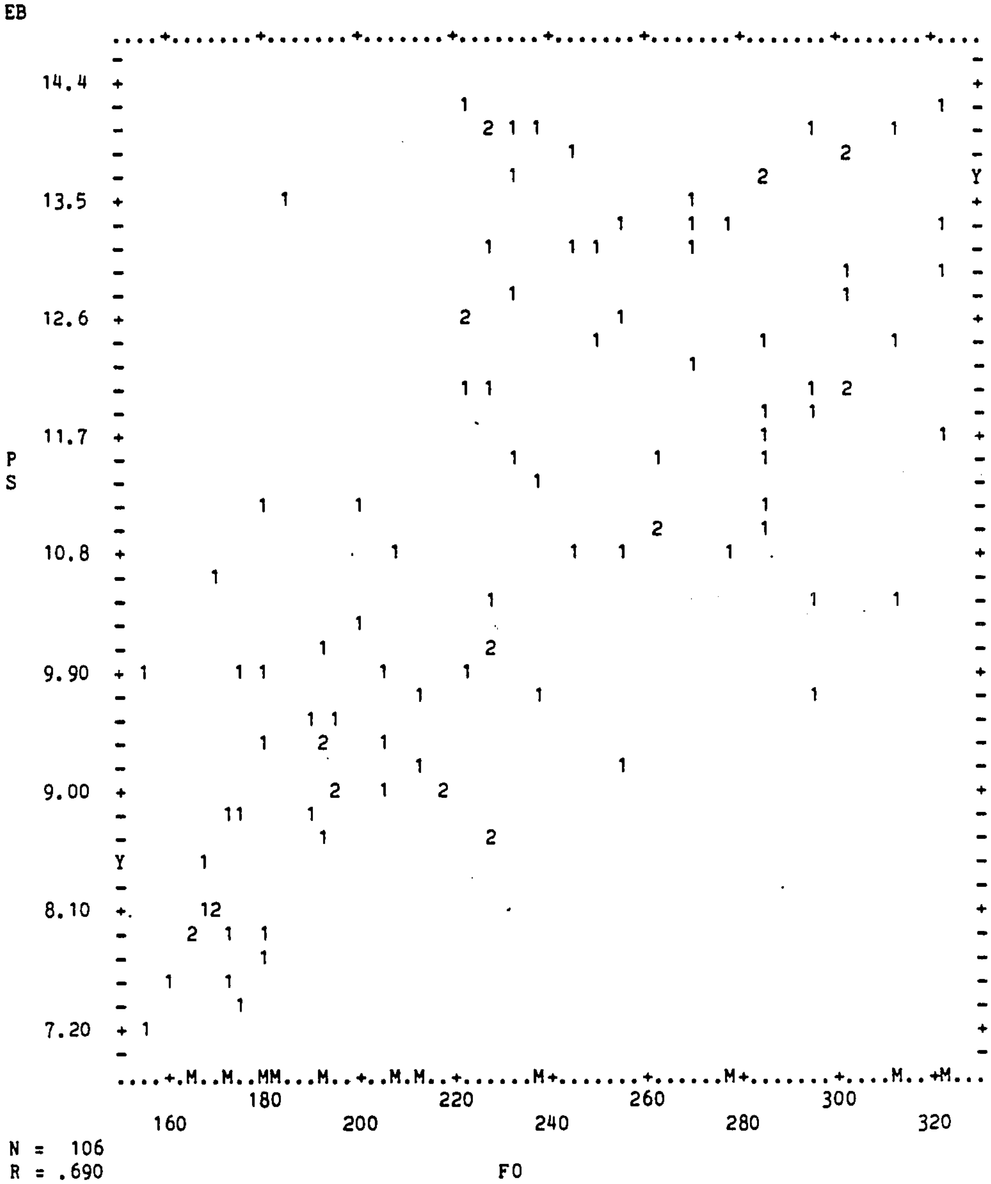
```
N =  131
R = .786
P < .001
                                          MEAN     S.D.
--REGRESSION LINE--   -RES.MS-       X   142.65   16.744
Y= 1.1999 +.06082*X     .64441       Y   9.8755   1.2948
```

Figure 28. Correlation of subglottal pressure and fundamental frequency for RC.



```
N =   46
R = .937
P < .001
                                           MEAN     S.D.
--REGRESSION LINE--  -RES.MS-     X   161.22   22.449
Y=-1.8649 +.08088*X    .46644     Y   11.174   1.9371
```

Figure 29. Correlation of subglottal pressure and fundamental frequency for RC English.



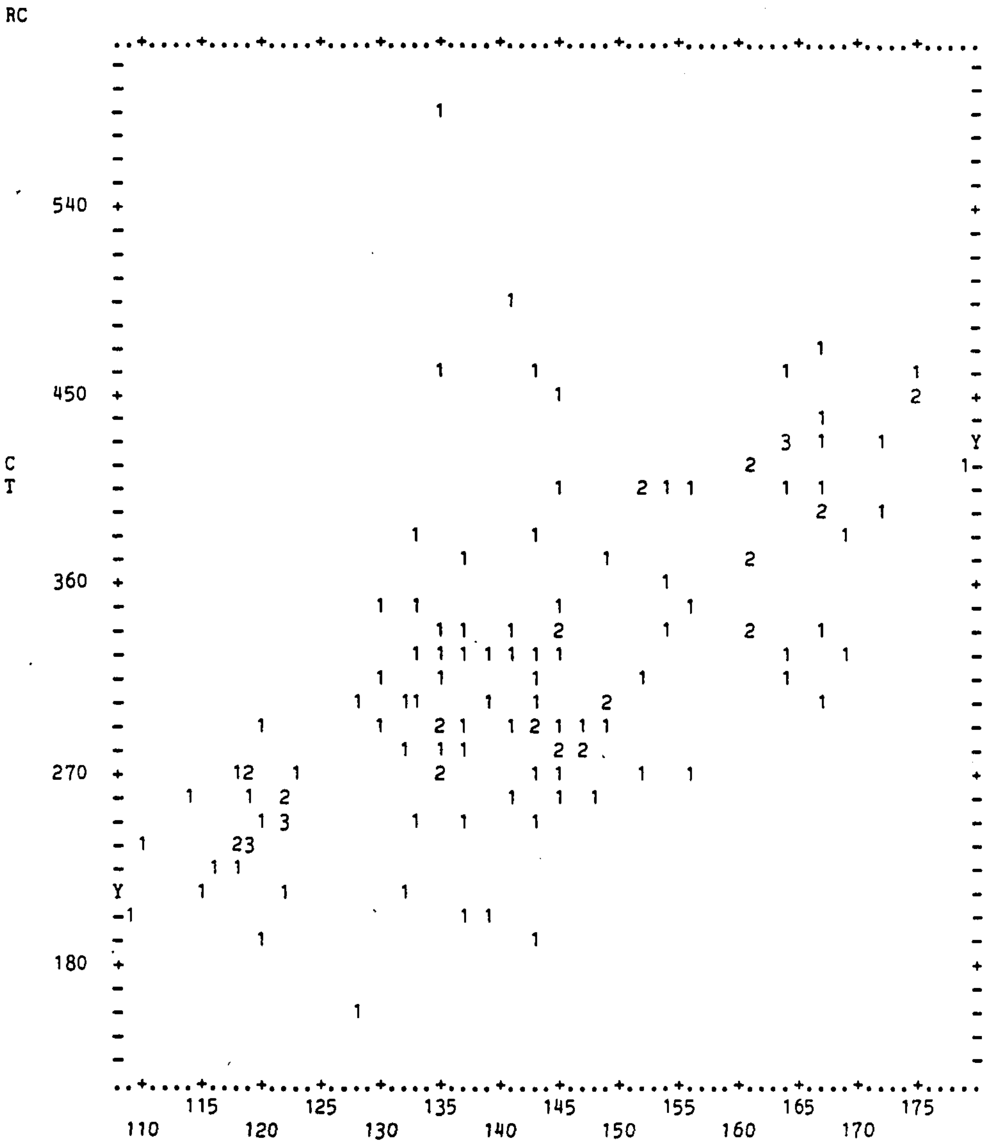
```
N =  106
R = .690
P < .001
                                          MEAN    S.D.
--REGRESSION LINE--   -RES.MS-     X    234.50   48.418
Y= 4.1720 +.02906*X     2.2007     Y    10.987   2.0395
```

Figure 30. Correlation of subglottal pressure and fundamental frequency for EB.

significant. Within individual stress conditions for RC English and EB, however, correlations are more variable. For RC English, only the Double stress tokens show a significant correlation. For EB, correlations are significant for Double and Neutral stress tokens, but not for Early and Late tokens, the latter actually showing a slightly negative correlation and a tendency for Ps to decrease as F0 increases.

_F0 vs. CT_: When peak F0 is plotted against peak CT activity for all tokens, correlation coefficients are high and significant for both subjects (RC: r=.662, p<.001; RC English: r=.747, p<.001; EB: r=.796, p<.001) (Figs. 31-33). Appendix L shows the correlations for individual stress and phonetic conditions. For both RC and EB, F0 and CT show significant correlations for all stress conditions. For RC English, however, there are no significant correlations within any stress conditions. As with the subglottal pressure data, all phonetic conditions for RC show significant correlations for CT and F0.

_CT vs. Ps_: There are two related reasons to expect CT and Ps to be correlated. The first, and the most obvious, is that they both correlate with F0. The second is that they are both mechanisms by which stress is implemented, so that their respective increases should be correlated.

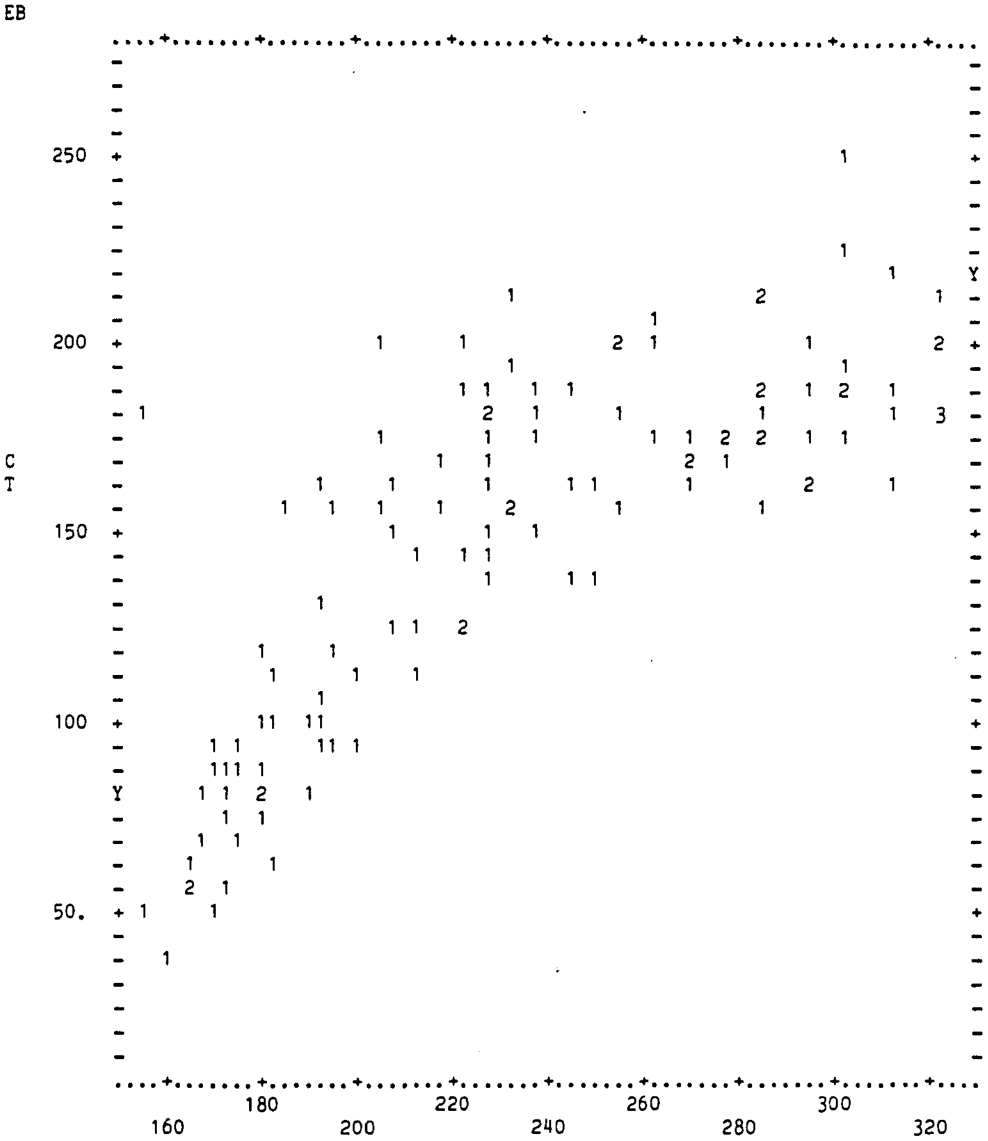


```
N =  131
R = .662
P < .001
                                        MEAN     S.D.
--REGRESSION LINE--  -RES.MS-    X   142.65   16.744
Y=-108.00 +3.0060*X    3267.1    Y   320.81   75.996
```

Figure 31. Correlation of cricothyroid muscle activity and fundamental frequency for RC.

```
RC ENGLISH
         .......+......+......+......+......+......+......+......+......+......+..
         -                                                                       -
         -                                                                       -
    540  +                                                          1            +
         -                                                                       -
         -                                                                       -
         -                                                                       -
         -                                                                 1   1 -
         -                                             1                1  1     Y
    480  +                                                              1  2     +
         -                                                     1                 -
         -                                          1                            -
         -                                                     1                 -
         -                                                                       -
         -     1                                                                 -
    420  +                                                              1        +
         -                                  1       1                            -
         -                                          1                            -
C        -                                                1                      -
T        -              1              1               1                         -
         -                             1                                         -
    360  +              1                 1    1                                 +
         -        1 1                                  2  1                1     -
         -                                                                       -
         -                                1 1                                    -
         -     1                                                                 -
         -                                                1                      -
    300  +     1                                                                 +
         -                                                                       -
         -      1                                                                -
         -                                                                       -
         -                                                                       -
         -                                                                       -
    240  Y                                                                       +
         -   1    1                                                              -
         -                                                                       -
         -                                                                       -
         -                                  1                                    -
         -1     1                                                                -
    180  + 1                                                                     +
         -        1                1                                             -
         -                                                                       -
         -   1                                                                   -
         -                                                                       -
         .......+......+......+......+......+......+......+......+......+......+..
               133           147           161           175           189
                      140           154           168           182           196
N =  46
R = .747                                 FO
P < .001
                                    MEAN    S.D.
--REGRESSION LINE--  -RES.MS-    X  161.22  22.449
Y=-198.60 +3.4664*X    4903.7    Y  360.24  104.16
```

Figure 32. Correlation of cricothyroid muscle activity and fundamental frequency for RC English.

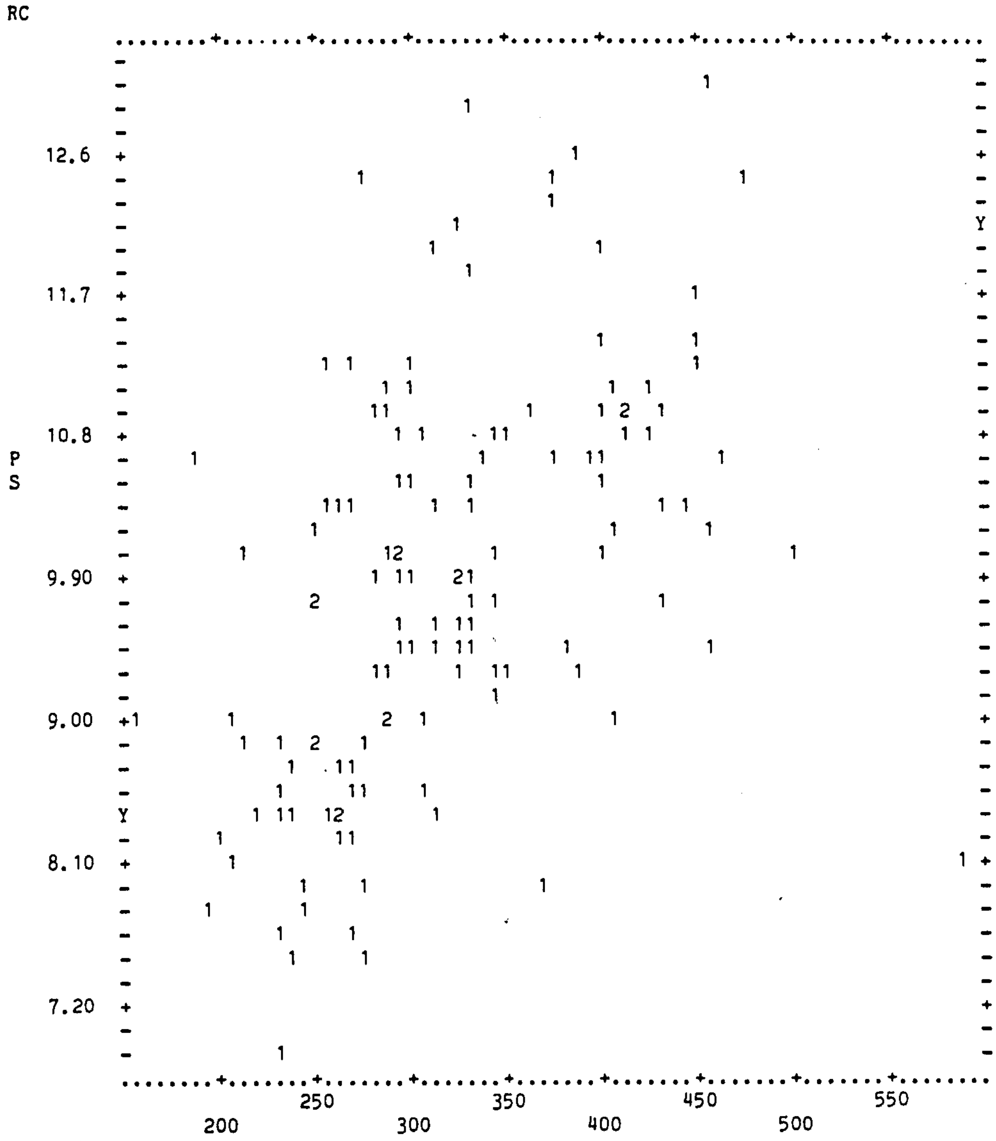


```
N =  121
R = .796
P < .001
                                          MEAN    S.D.
--REGRESSION LINE--   -RES.MS-     X    233.17   49.461
Y=-31.334 +.76712*X    837.41      Y    147.54   47.645
```

Figure 33. Correlation of cricothyroid muscle activity and fundamental frequency for EB.

For both subjects, correlations of CT and Ps are significant (RC: $r=.486$, $p<.001$; RC English: $r=.676$, $p<.001$; EB: $r=.63$, $p<.001$) (Figs. 34-36). For RC and RC English, these correlations are lower than either variable's correlation with F0, while, for EB, the correlation of Ps and CT is higher than that for Ps and F0, but lower than that for CT and F0. Correlation coefficients for individual stress and phonetic conditions are shown in Appendix M. In general, where both Ps and CT show significant correlations with F0, they are significantly correlated with each other as well (see Figs. 28-33 and App. K and L). Only RC Double and EB Double are exceptions here. Conversely, there is no case where there is a significant correlation of CT and Ps when either variable fails to show a significant correlation with F0. It would thus appear that Ps and CT activity are unrelated except to the extent that they each relate to fundamental frequency.

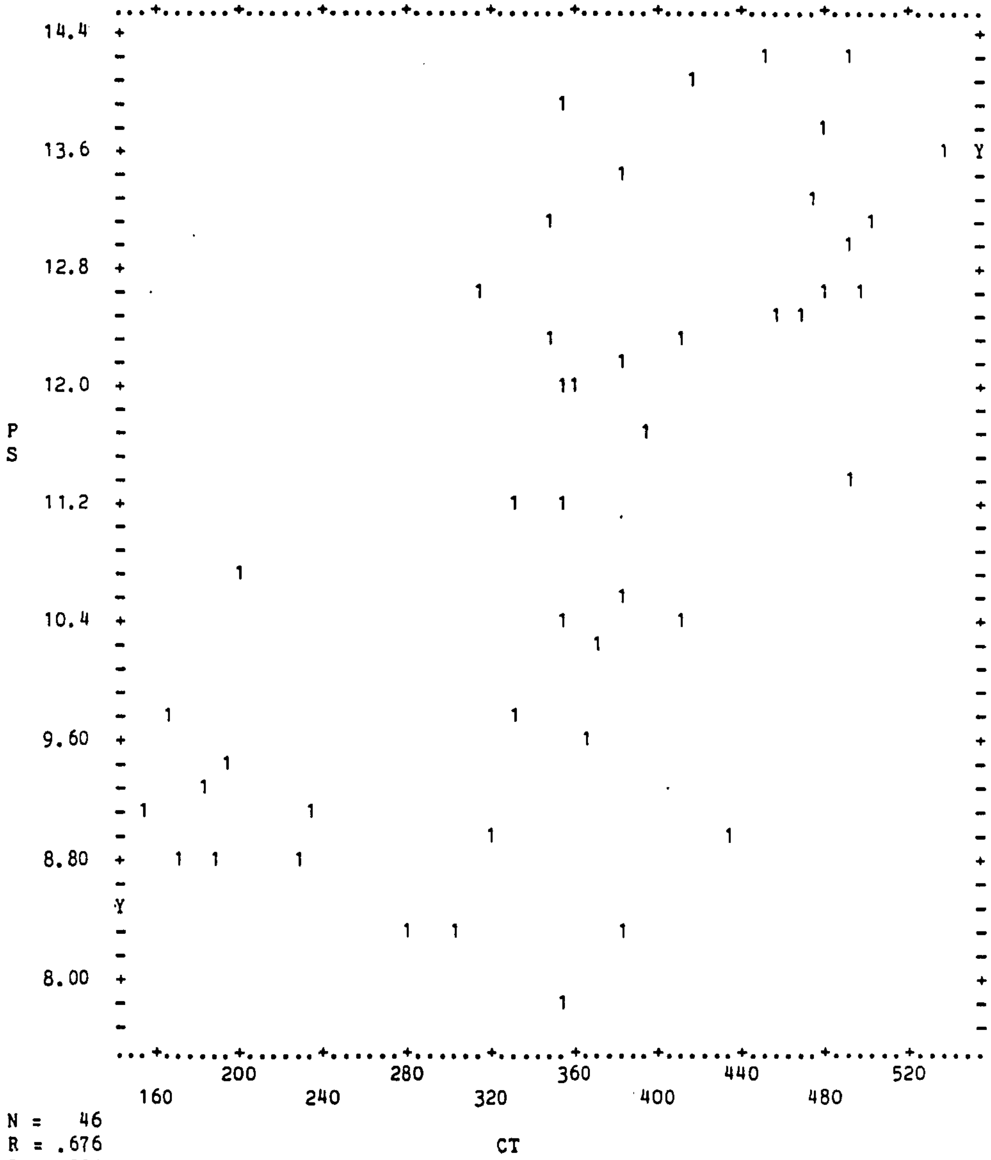
There remains the question of the extent to which either variable's correlation with F0 is a function of the degree to which the other variable is contributing simultaneously. However, this question cannot be answered by the simple correlations performed above. Table 12 shows the resulting correlations (R) of both Ps and CT with F0 when a partial correlation procedure is employed (Williams, 1986). In all cases, there is a decrease in the original correlation of either variable with F0 when the influence of the other variable is removed. However, the amount of decrease is variable. For example, for RC, both the original Ps and CT correlations with F0 drop by relatively small



```
N =  131
R = .486                                 CT
P < .001
                                    MEAN     S.D.
--REGRESSION LINE--  -RES.MS-   X   320.81   75.996
Y= 7.2179 +.00828*X    1.2900   Y   9.8755   1.2948
```

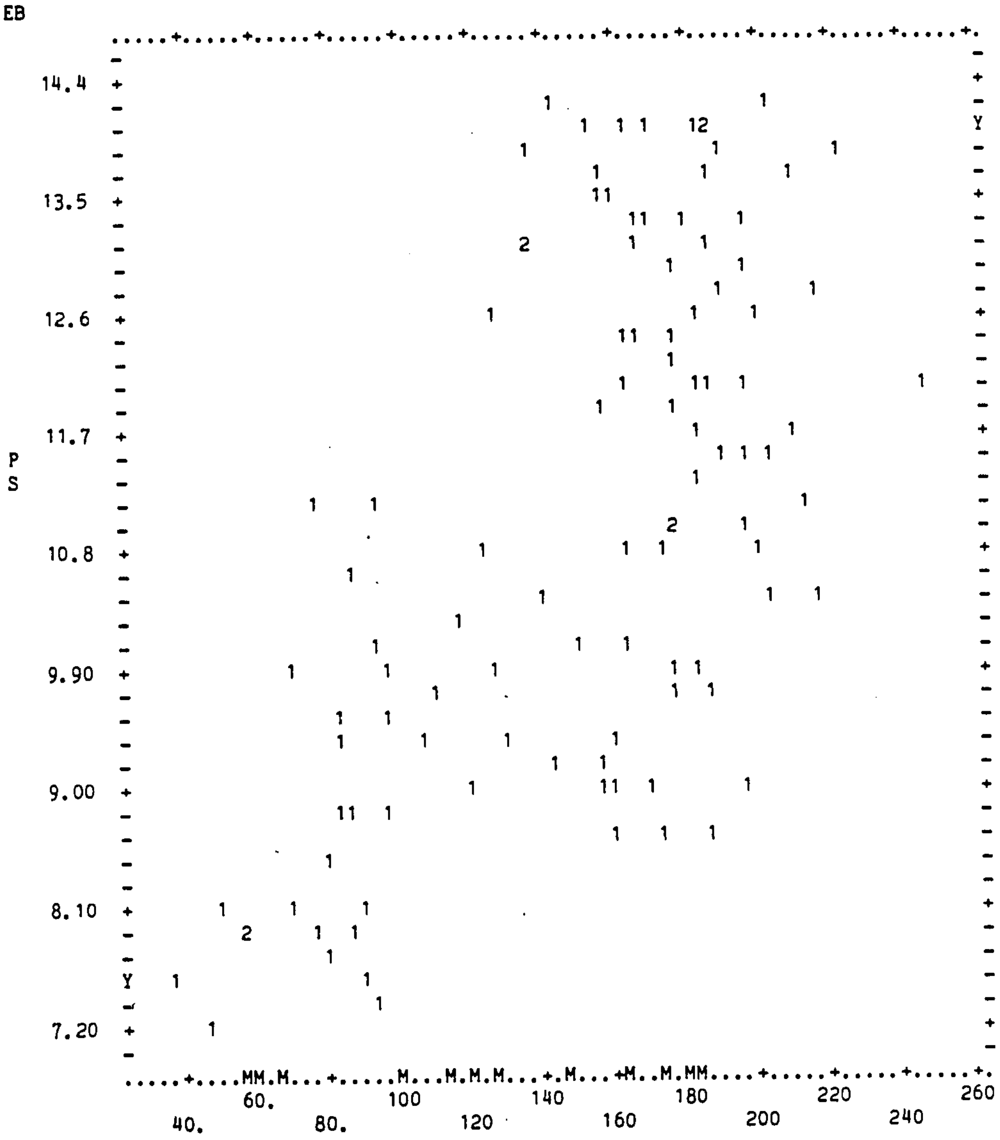
Figure 3c. Correlation of subglottal pressure and cricothyorid muscle activity for RC.

RC ENGLISH



```
N =   46
R = .676
P < .001

                                  MEAN    S.D.
--REGRESSION LINE--  -RES.MS-   X  360.24  104.16
Y= 6.6478 +.01256*X    2.0861   Y  11.174  1.9371
```

Figure 35. Correlation of subglottal pressure and cricothyorid muscle activity for RC English.



```
N =  106
R = .630
P < .001
                                        MEAN    S.D.
--REGRESSION LINE--   -RES.MS-     X   150.32   47.560
Y= 6.9266 +.02701*X     2.5335     Y   10.987   2.0395
```

Figure 36. Correlation of subglottal pressure and cricothyroid muscle activity for EB.

Table 12

Partial Correlations of Initial Peak Subglottal Pressure and Cricothyroid Muscle Activity with Fundamental Frequency

	r	*R*	*R*	*r-R*	
Ps	.786	-	.709	.077	*RC*
CT	.664	.522	-	.142	
Ps	.937	-	.883	.054	*RC English*
CT	.747	.443	-	.304	
Ps	.568	-	.294	.274	*EB*
CT	.750	.640	-	.110	

amounts, indicating that most of the relationship of either variable with FO is independent of the other variable's relationship with FO. The same can also be said for Ps for RC English and CT for EB. However, for CT for RC English and Ps for EB, the original correlations with FO decrease by almost half their original r values, suggesting that their respective correlations with FO are largely dependent on the relationship of the other variable with FO.

Still, correlations are illuminating only to the extent that they show the degree to which any two variables are related statistically. They fall short, however, in being able to demonstrate a causal relationship between either Ps or CT and FO. In other words, they cannot answer the question of whether either variable alone can account for the magnitude of the observed increase in peak FO.

Previous research in this area has suggested that the cricothyroid muscle is the primary mechanism by which the increases in FO for accented syllables are implemented, with concommitant increases in subglottal pressure accounting for only a fraction of these FO increases (eg., Fromkin & Ohala, 1968). Table 13 shows the difference in FO and Ps between the Early and -Early stress conditions (derived from Tables 7 and 8) and the ratios of these differences (Hz/cm-H2O) for each length condition for RC, RC English and EB. It can be seen that these ratios fall well outside the accepted range of 2-7 Hz/Cm-H2O that would be predicted if subglottal pressure alone were the mechanism responsible for the difference in FO, thus suggesting that it is the cricothyroid muscle that accounts for the rapid and significant

Table 13

Differences Between +Early and -Early Stress Conditions for Fundamental Frequency and Subglottal Pressure and Ratios of Hz/Cm-H2O

	Stress-No Stress		Ratio	
	FO (Hz)	Ps (Cm-H2O)	FO/Ps (Hz/Cm-H2O)	
Length 1	20	1.17	17.09	
Length 2	26	2.28	11.40	RC
Length 3	27	1.54	17.53	
Mean	24	1.66	15.34	
Length 1	28	2.77	10.11	
Length 2	42	3.66	11.48	RC English
Length 3	54	3.66	14.75	
Mean	41	3.36	12.11	
Length 1	89	3.10	28.71	
Length 2	51	2.81	18.15	EB
Length 3	83	3.16	26.27	
Mean	74	3.02	24.38	

increases in F0 that occur at syllable peaks.

Summary

The discussion of this part of Experiment 1 has reported the effects of utterance length and initial stress configuration on initial peak fundamental frequency, subglottal pressure, cricothyroid muscle activity and, for one subject, depth of inspiration. It has also attempted to relate the behavior of peak F0 to that of the concommitant levels in Ps and CT activity in order to determine the primary mechanism by which local, rapid increases in F0 are implemented.

The effects of utterance length on initial conditions were summarized above. These results showed an inconsistent length effect across subjects, ranging from significant effects for RC to no effect for EB, with one subject (RC) showing a different pattern of initial activity when speaking native and non-native languages.

The degree of stress placed on a syllable (i.e., lexical or emphatic) was reflected in the peak F0 values for that syllable and in the accompanying peak Ps and CT activity. However, the difference in F0 between initial syllables receiving lexical stress (i.e., -Early) and those receiving emphatic stress (i.e., Early) could not be accounted for by the difference in subglottal pressure for the same syllables. Thus, it must be the cricothyroid muscle that is responsible for the implementation of the rapid, local increases in F0 that occur for stressed syllables.

Levels of peak Ps and CT activity appear to be related only to the extent that they share a common correlation with F0, although, in two cases (i.e., CT for RC English and Ps for EB), the correlation of each of these variables with F0 appears to be largely a function of the simultaneous influence of the other physiological variable on F0, making the significance of their respective correlations with F0 more apparent than real. For the most part, however, when there is an increase in F0, there will be an increase in the level of both Ps and CT activity as well.

EXPERIMENT 1: DECLINATION

Having established the influence of utterance length and stress configuration on peak F0, Ps and CT values, as well as their relationship to each other, the remainder of the analyses will focus on the more global aspect of the behavior of these variables as a function of the same conditions for both RC and EB. In addition, an attempt will be made to relate fundamental frequency declination to the underlying physiological variables.

1. The Effect of Utterance Length

Fundamental Frequency: Figures 37-39 show the averaged tokens for the three length conditions within each stress condition for RC's Dutch, /ma/ and /fa/ utterances, respectively. For purposes of visual clarity, the contours are superimposed for the continuously voiced /ma/

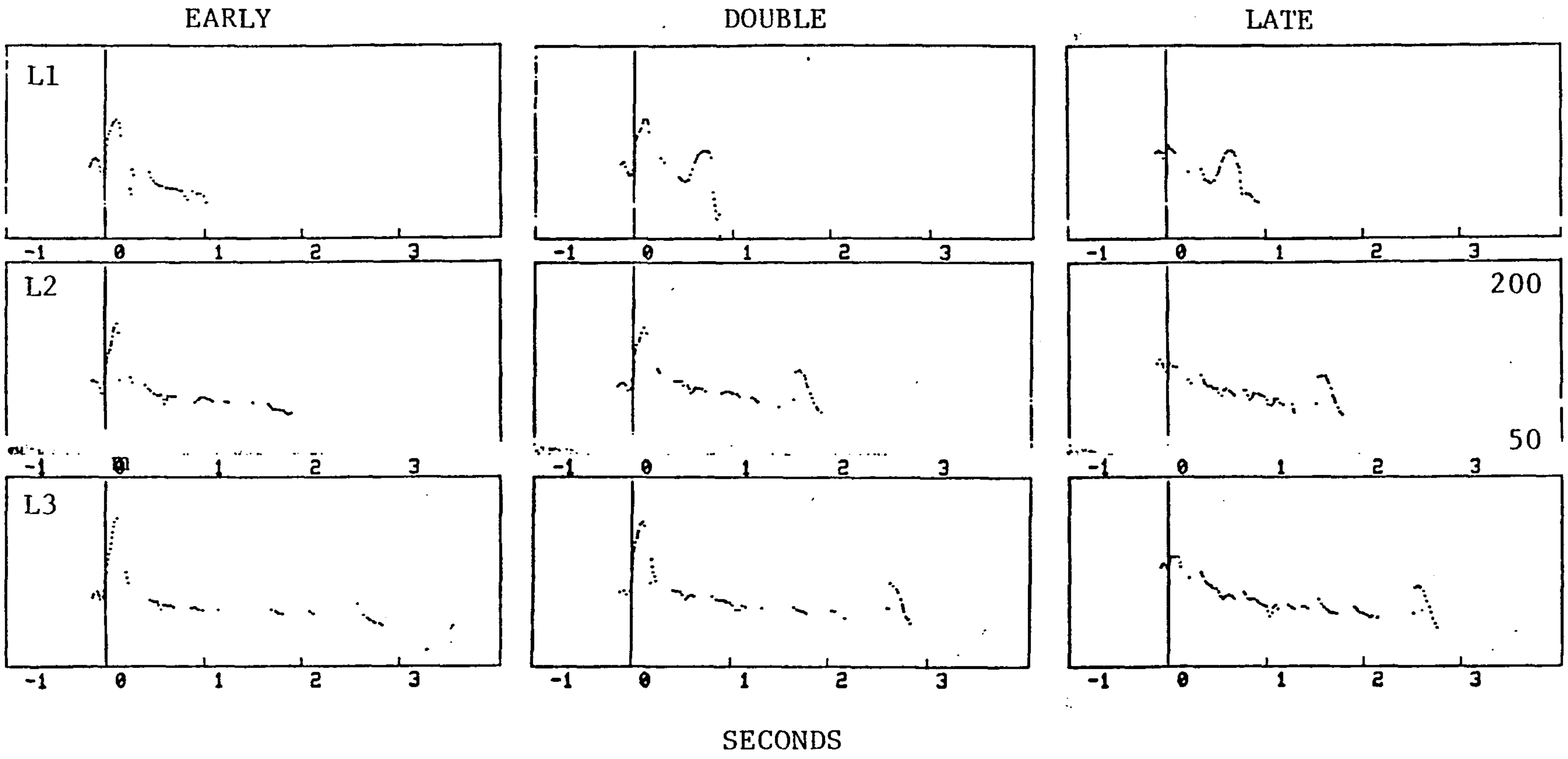


Figure 37. Comparison of F0 contours across utterance lengths for RC Dutch.

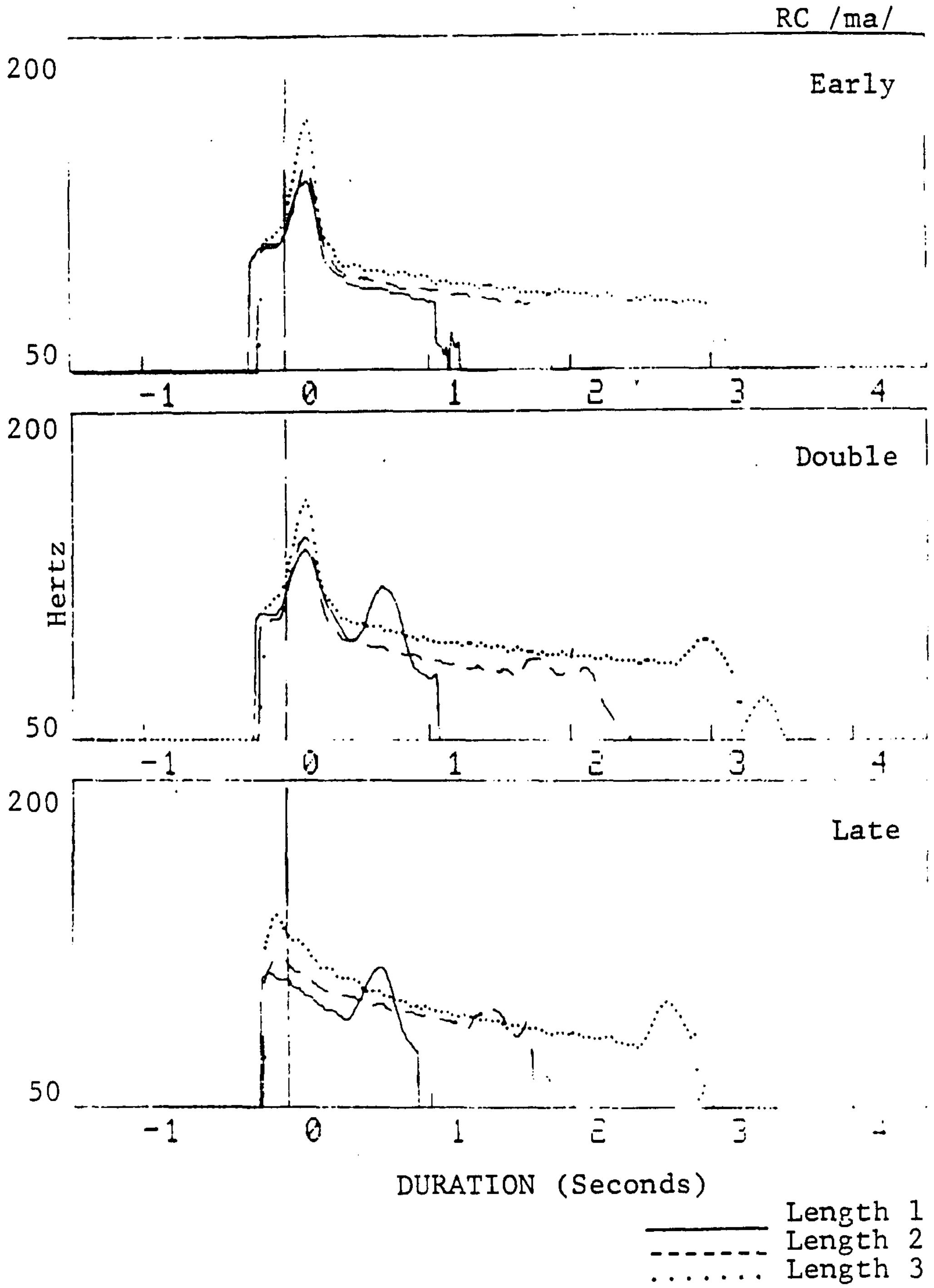


Figure 38. Comparison of F0 contours across utterance lengths for RC /ma/.

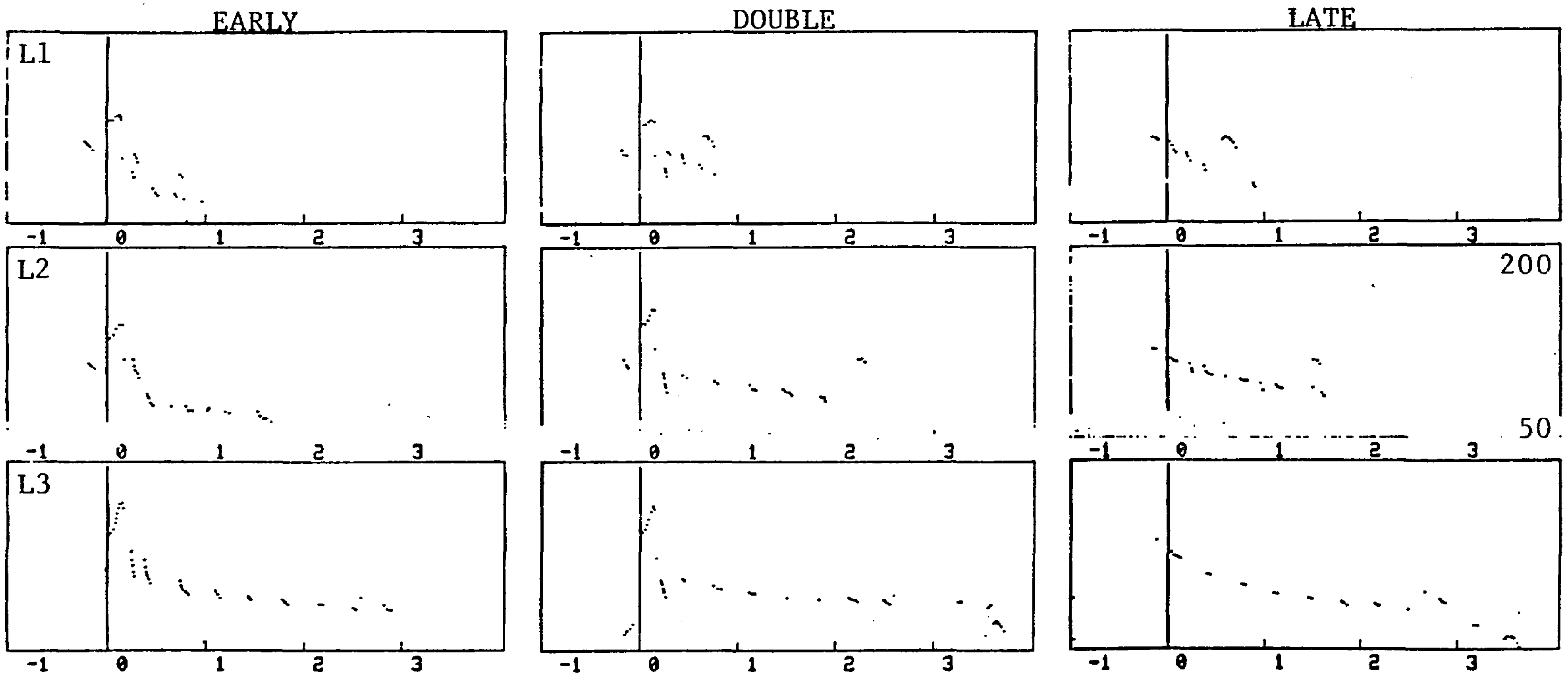


Figure 39. Comparison of F0 contours across utterance lengths for RC /fa/.

utterances, but are shown individually for the intermittently voiced Dutch and /fa/ utterances.

For the Early and Double stress utterances, there is an obvious peak associated with the initial emphasized syllables. However, F0 does not decline steadily from these peaks. Rather, there is a rapid drop in frequency to a point from which F0 then begins a steady decline. And, while the duration of this initial frequency drop is constant across lengths, despite differences in peak height, the frequencies from which the slow decline begins for each length are not, bearing instead the same relationship as the initial peaks. This relationship is maintained throughout the course of at least the longer utterances, with the F0 contours declining in parallel.

In the absence of early emphasis in the Late stress utterances, the F0 peaks occur upon initiation of the utterances and are thus displaced in time relative to the early stress peaks in the Early and Double stress conditions. Furthermore, the decline of F0 from these peaks is more gradual.

The contours for the /ma/ utterances in Figure 38 are of special interest. These probably best represent what has been termed baseline declination in as pure a form as possible in that, due to a composition of only voiced continuants, as well as to a relatively straightforward stress configuration, there are few perturbations superimposed on their contours.

Figures 40 and 41 depict the same length comparisons for each stress condition for RC English and for EB, respectively. It is interesting that the displacement of initial peaks noted above for RC's Late stress utterances does not occur for EB's Neutral or Late stress data or even for RC's English Late utterances. The relationships among the individual curves are difficult to interpret because of the numerous discontinuities secondary to segmental effects. It can, however, be seen that, beyond the initial peaks, the trajectories of the F0 'curves' are very similar, despite differences in utterance length.

<u>Subglottal</u> <u>Pressure</u>: Subglottal pressure contours are displayed in the same manner for both subjects. RC's data are displayed in Figures 42-44 his English data in Figure 45, and EB's data in Figure 46.

The Ps contours show the same general pattern as the F0 contours. That is, there is a relatively rapid initial pressure drop from the initial peaks into a generally declining function for the longer utterances. The temporal displacement of initial peaks for RC's Late stress utterances is again apparent in his pressure curves for the Dutch, /ma/ and /fa/ utterances, but not for the English.

Segmental effects on the subglottal pressure, particularly for the Dutch and English, are quite obvious. However, unlike F0, Ps remains continuous, so that the relationships among the various lengths are

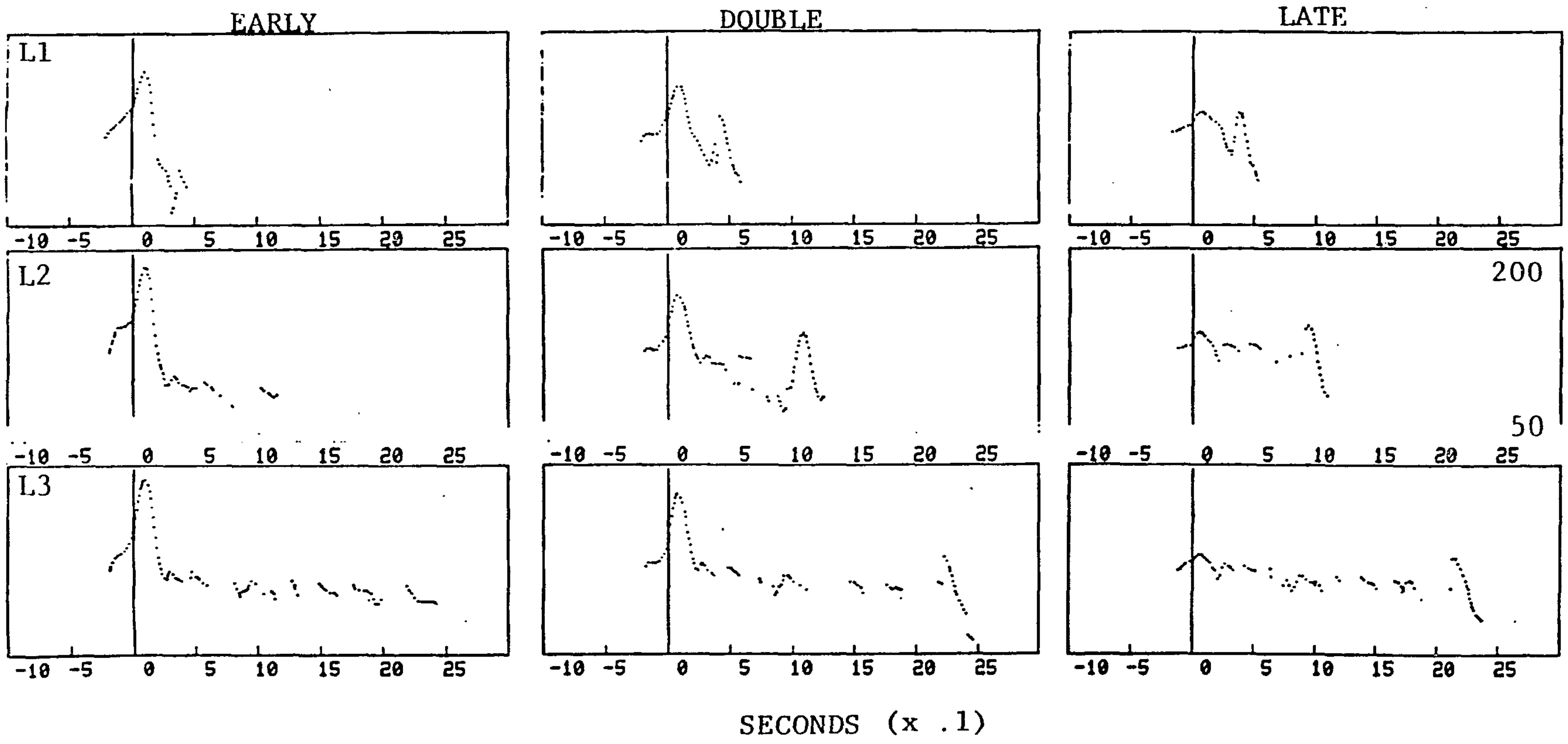


Figure 40. Comparison of F0 contours across utterance lengths for RC English.

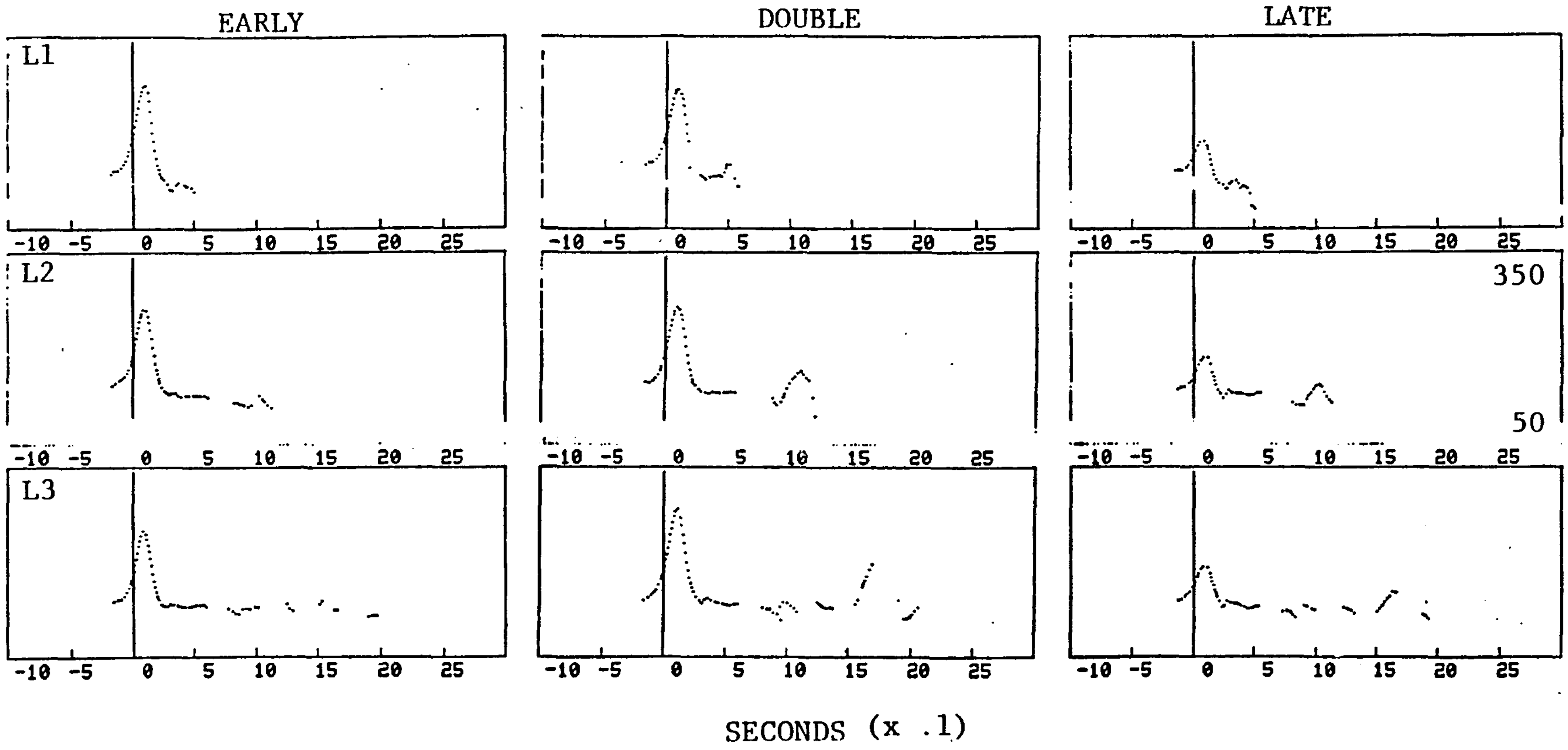


Figure 41. Comparison of F0 contours across utterance lengths for EB.

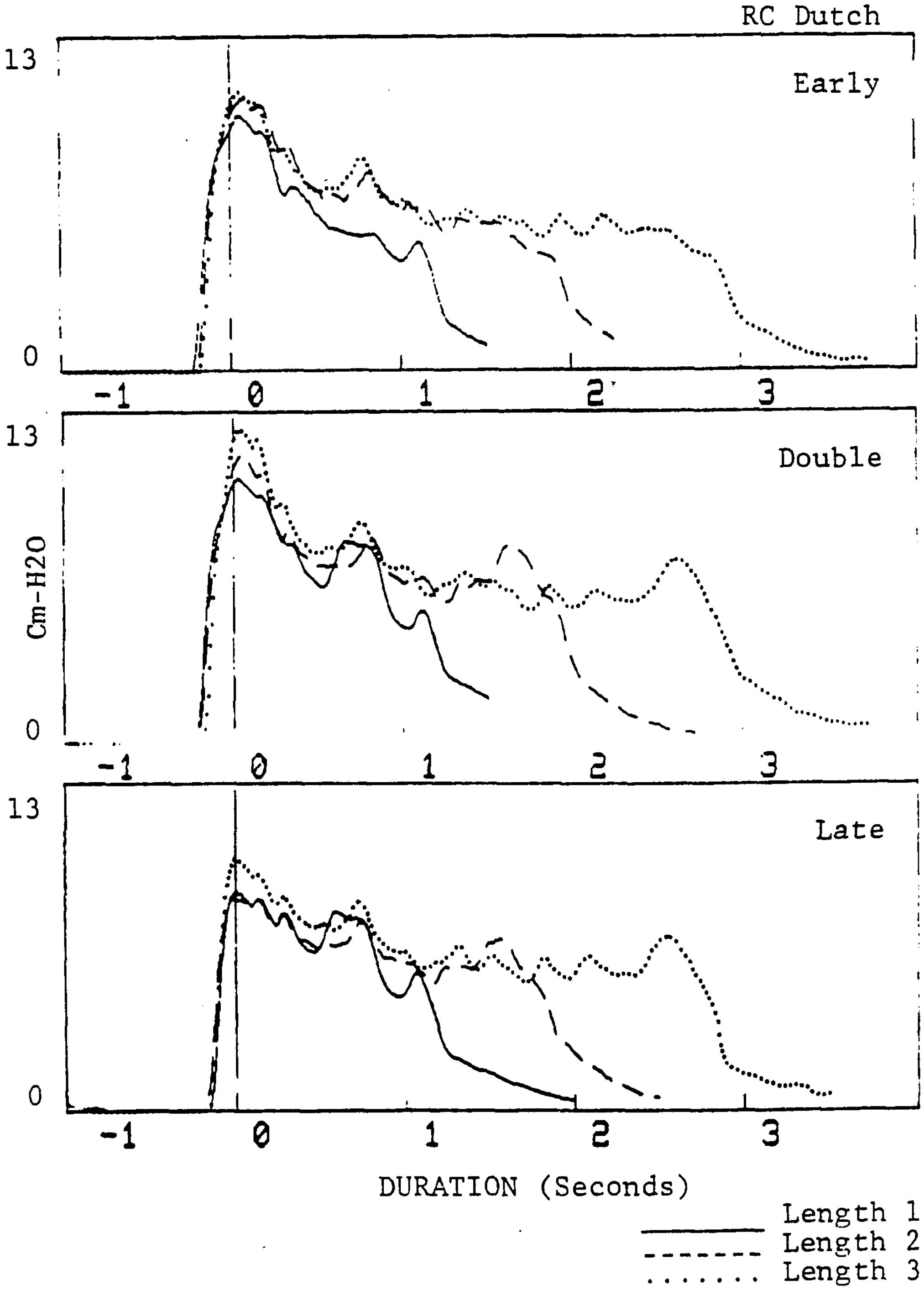


Figure 42. Comparison of Ps contours across utterance lengths for RC Dutch.

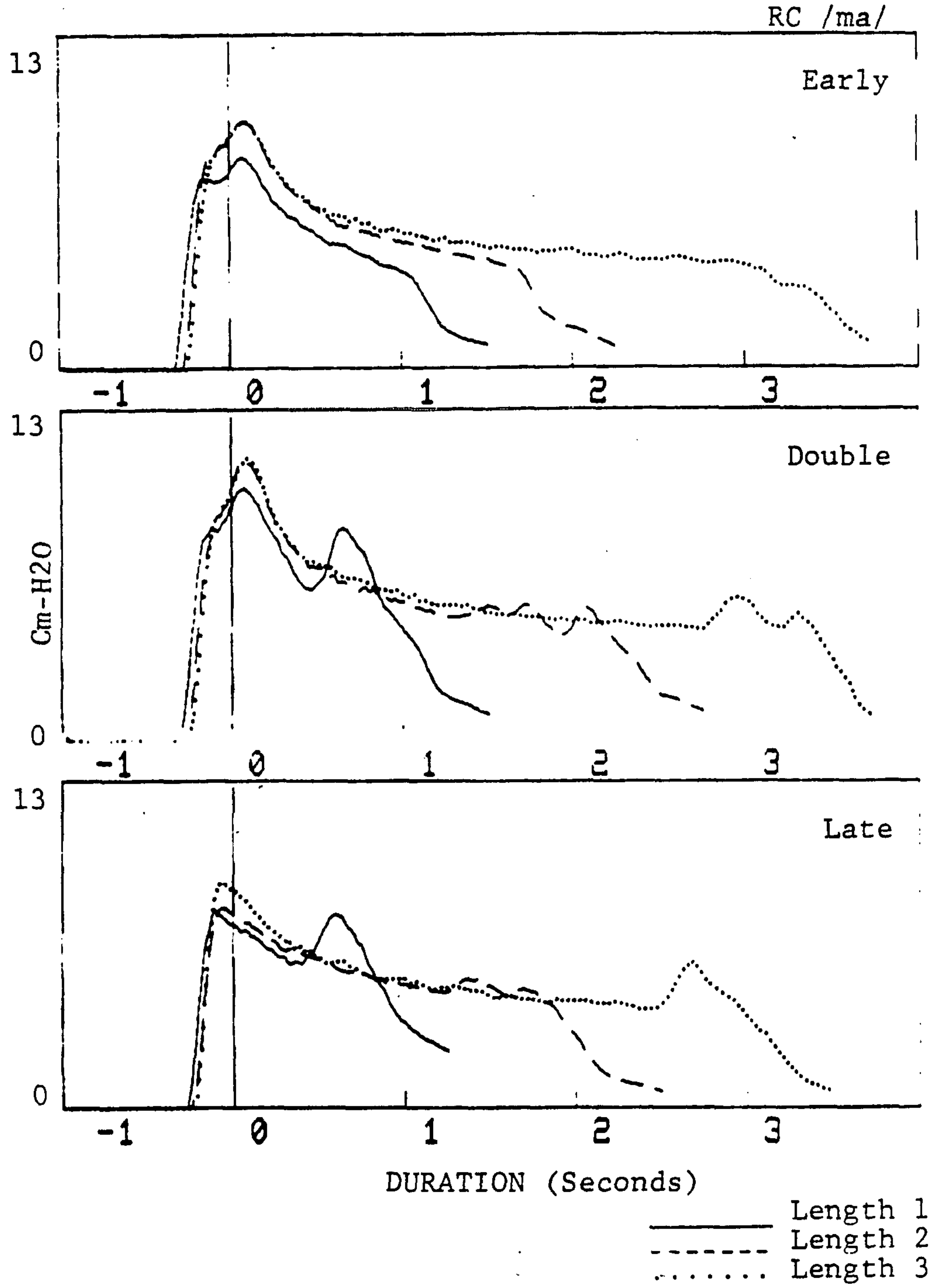


Figure 43. Comparison of Ps contours across utterance lengths for RC /ma/.

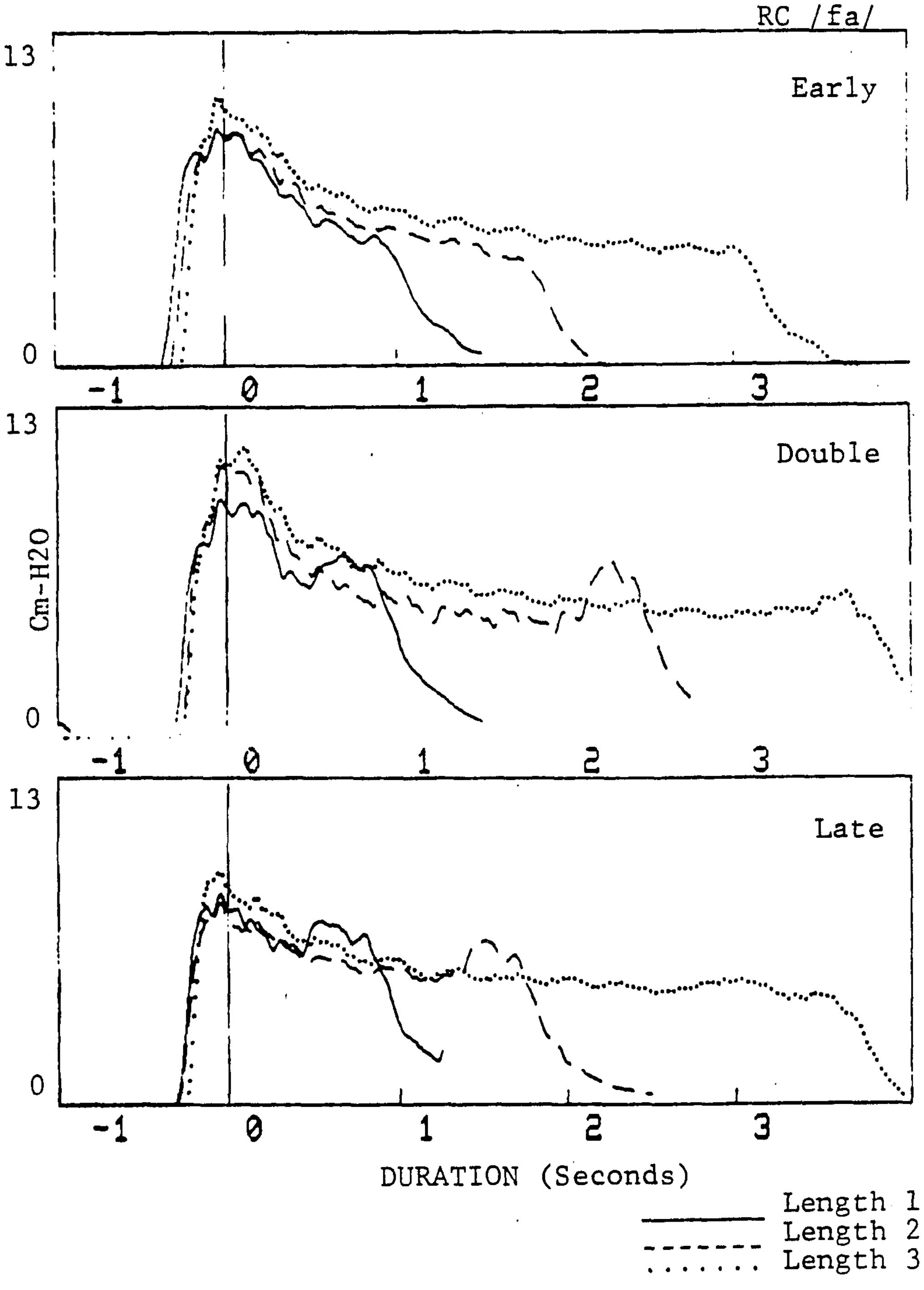


Figure 44. Comparison of Ps contours across utterance lengths for RC /fa/.

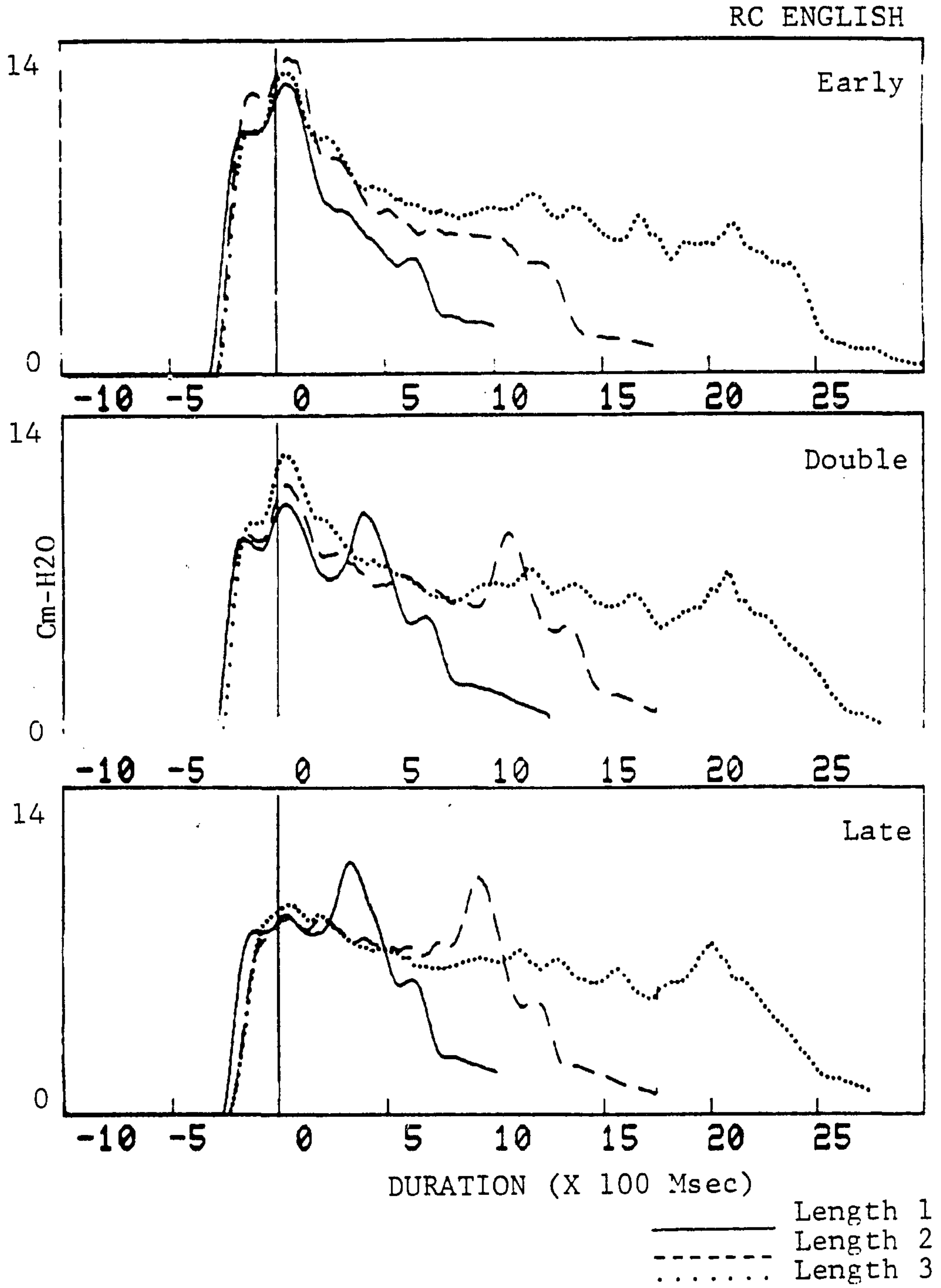


Figure 45. Comparison of Ps contours across utterance lengths for RC English.

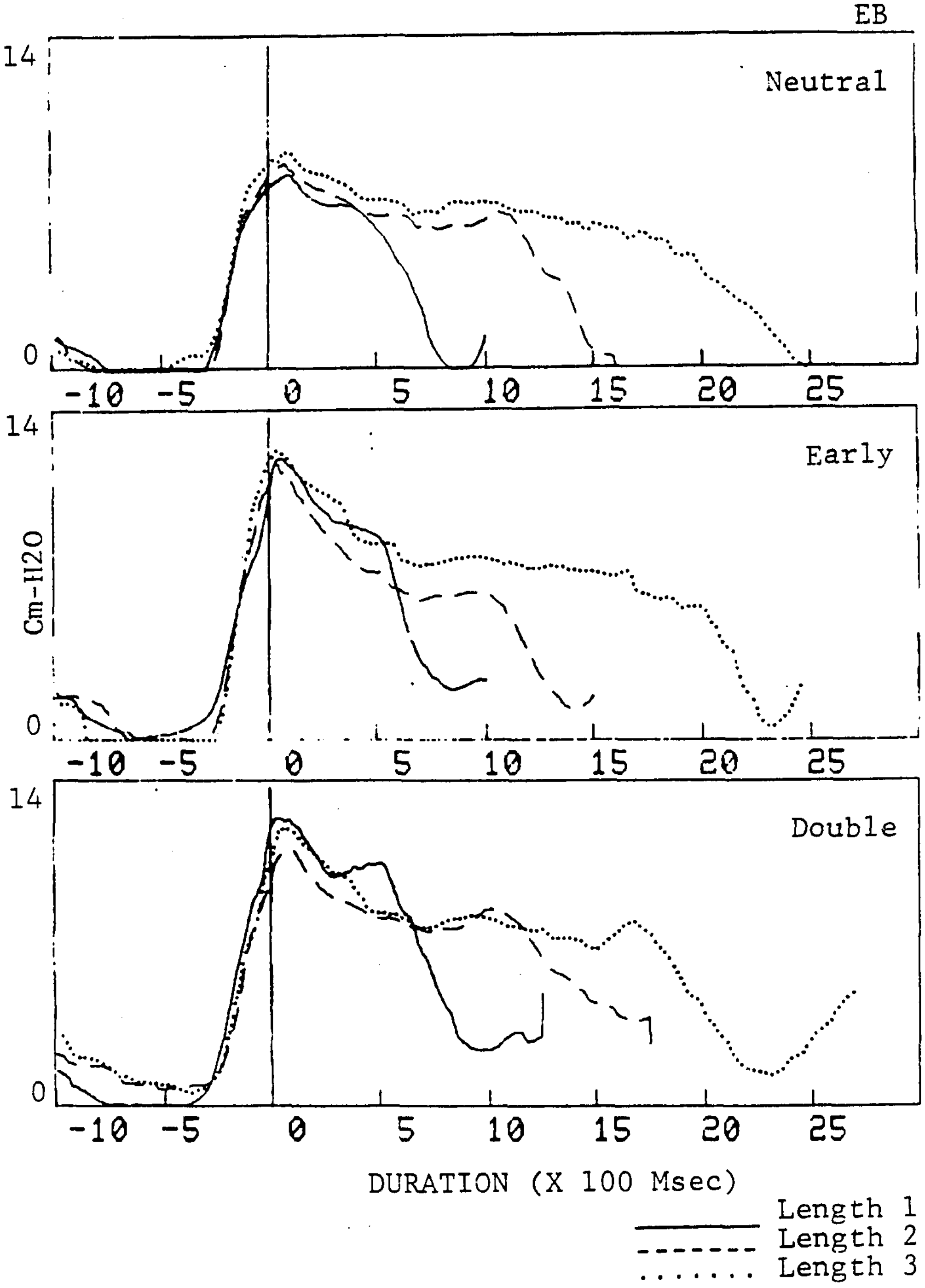


Figure 46. Comparison of Ps contours across utterance lengths for EB.

somewhat easier to discern. As with the F0 data, the relationship between the Ps curves appears to be a function of the relationship between their respective peaks. For almost every length comparison for both subjects, after the initial pressure drop the Ps curves are either superimposed or they decline more-or-less in parallel. However, for the Length 1 utterances, particularly those for which there is no final stressed syllable (i.e., Early or Neutral), there is a much more rapid fall-off in pressure, so that the parallelism observed for the longer utterances is less apparent for the shortest utterances.

Cricothyroid Muscle Activity: Corresponding cricothyroid muscle data are depicted in Figures 47-49 for RC and in Figures 50 and 51 for RC English and for EB, respectively. The superimposition of these data makes it rather difficult to discern the individual tracings, so that the peaks associated with final stressed syllables are numbered according to utterance length in the Double and Late stress utterances in order to separate these peaks from activity associated with segmental adjustments. It should be noted, however, that the double peaks that appear to be associated with the final stressed syllables of Lengths 2 and 3 for the /ma/ utterances in particular are the result of averaging events that are distant from the line-up point in tokens of unequal lengths, and are not characteristic of CT activity for final stressed peaks.[3]

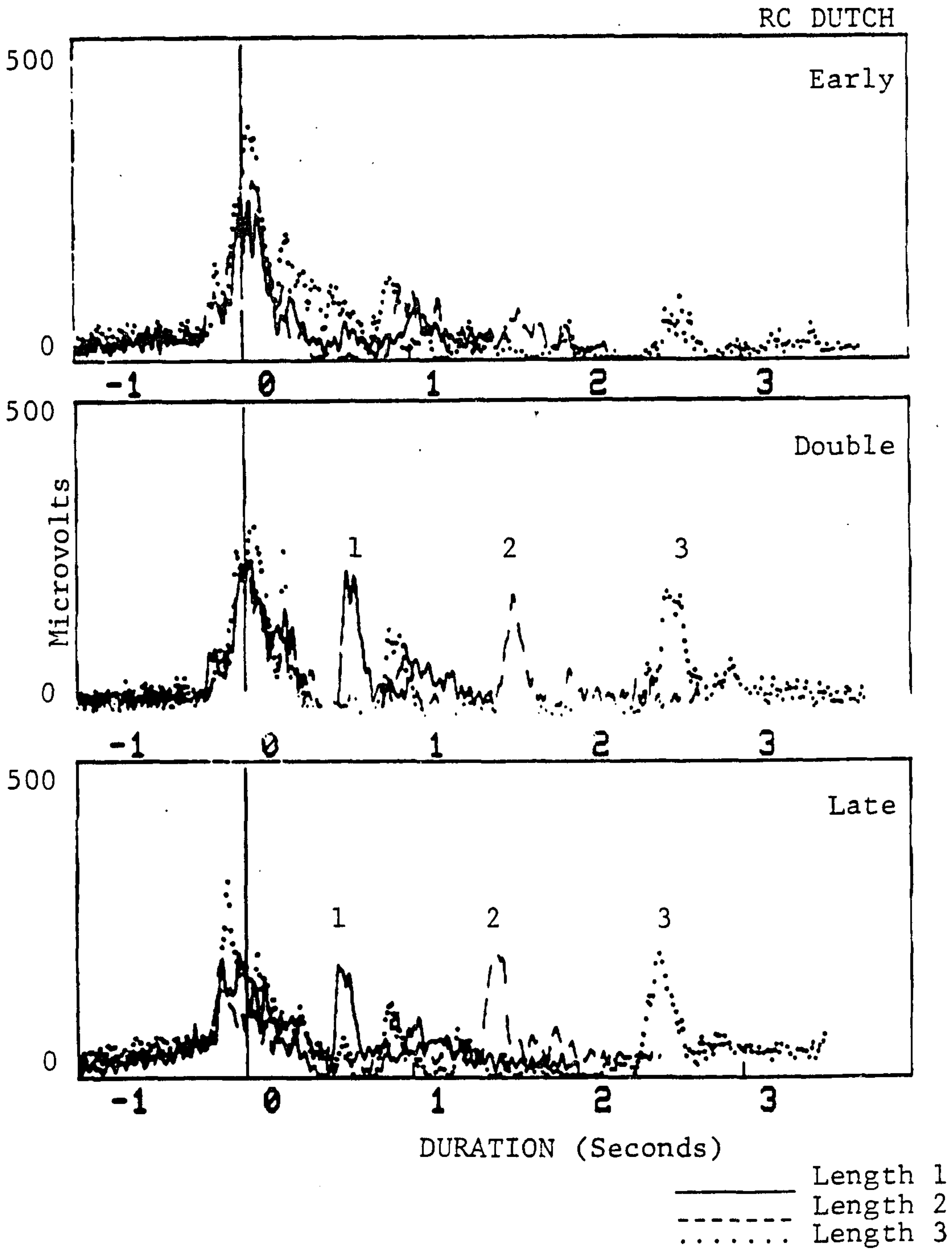


Figure 47. Comparison of CT muscle activity across utterance lengths for RC Dutch.

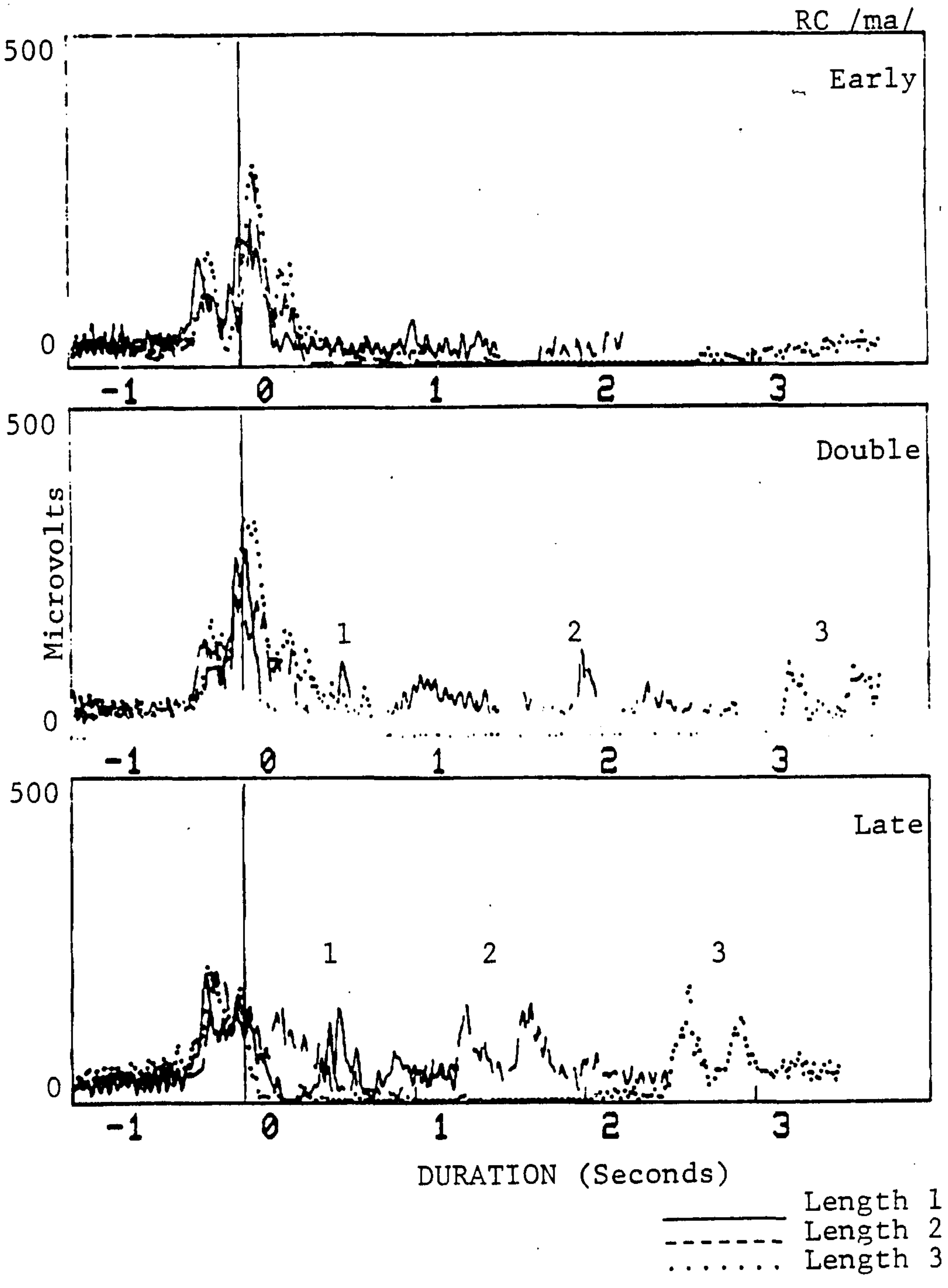


Figure 48. Comparison of CT muscle activity across utterance lengths for RC /ma/.

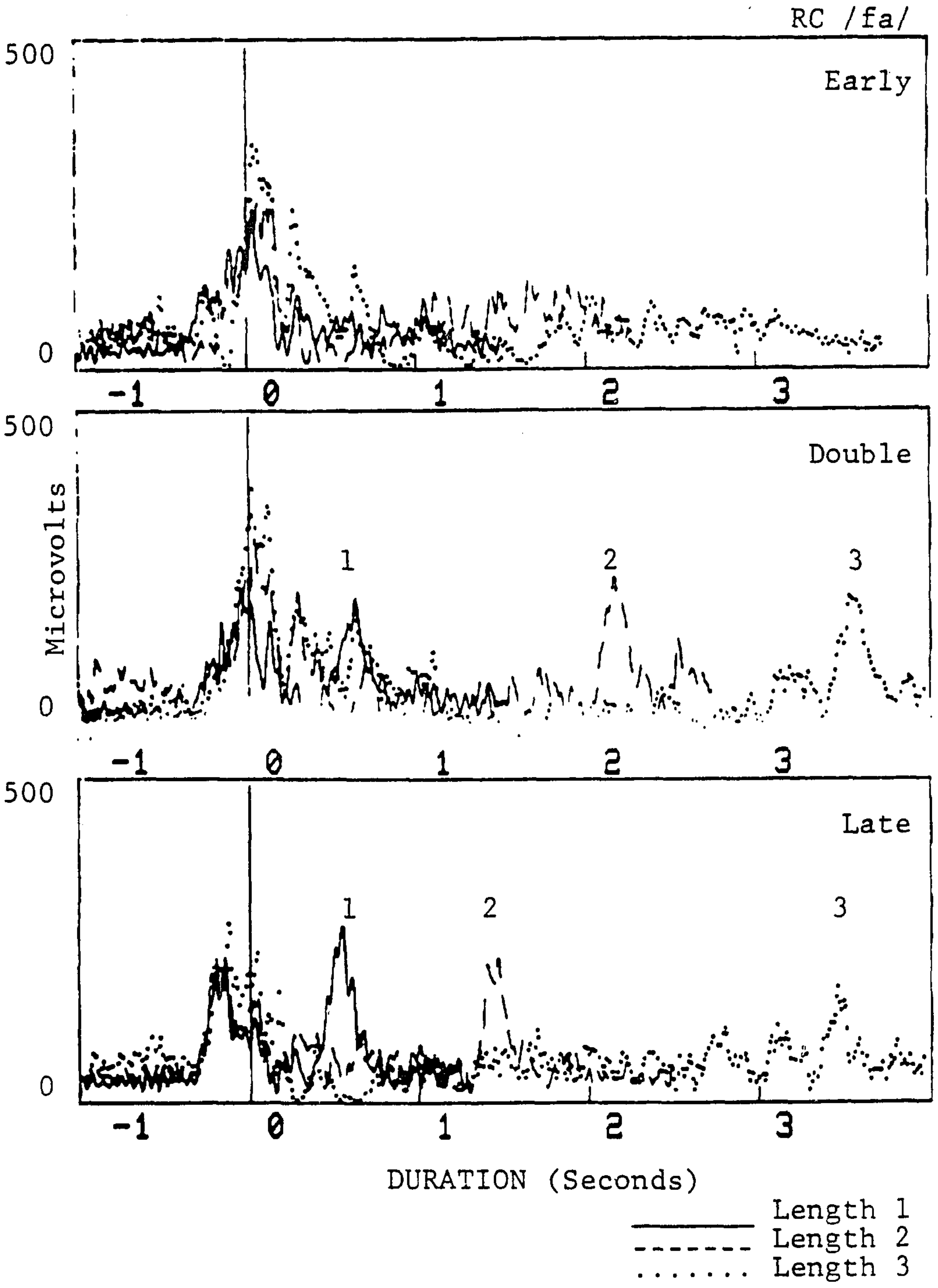


Figure 49. Comparison of CT muscle activity across utterance lengths for RC /fa/.

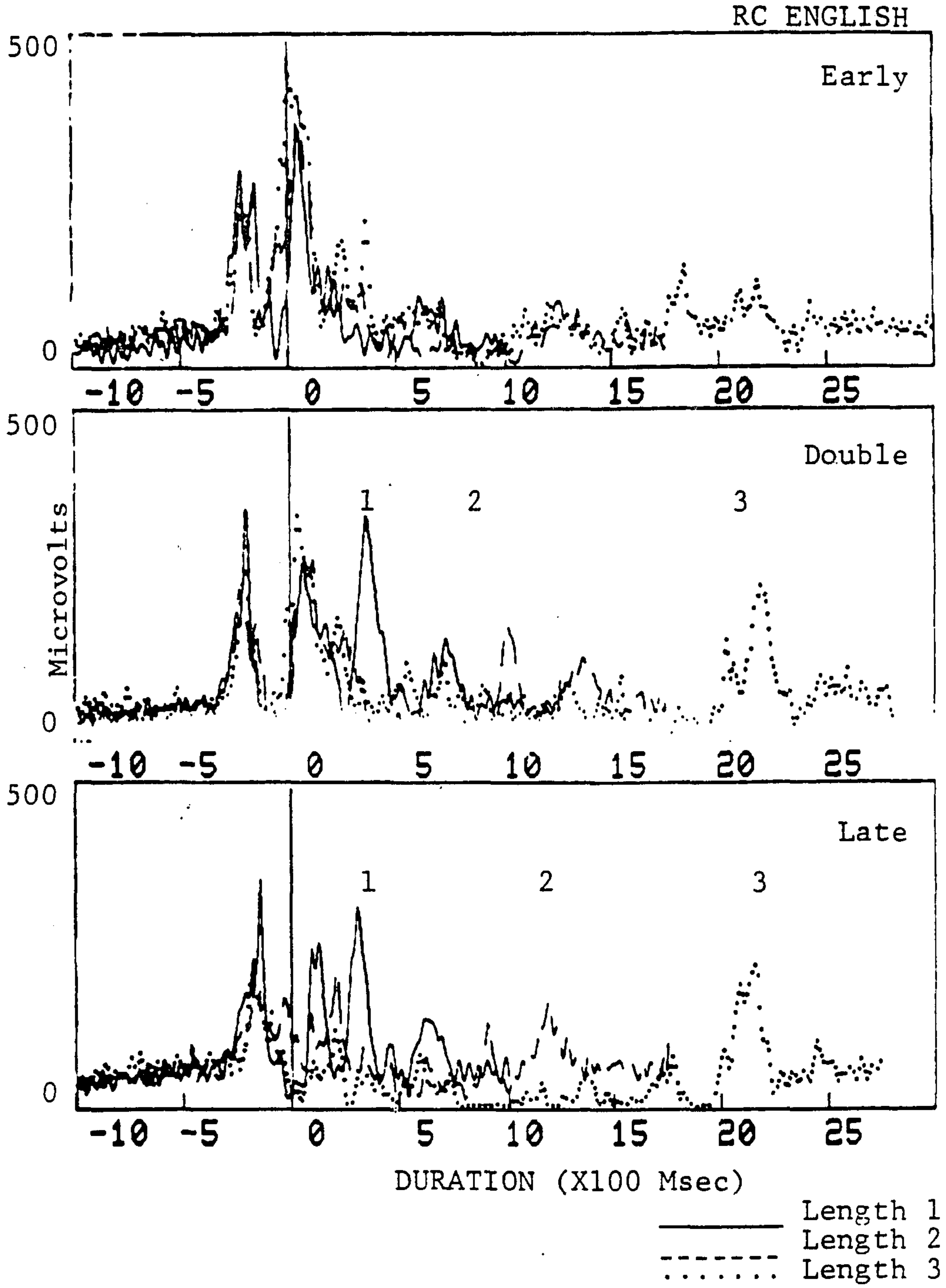


Figure 50. Comparison of CT muscle activity across utterance lengths for RC English.

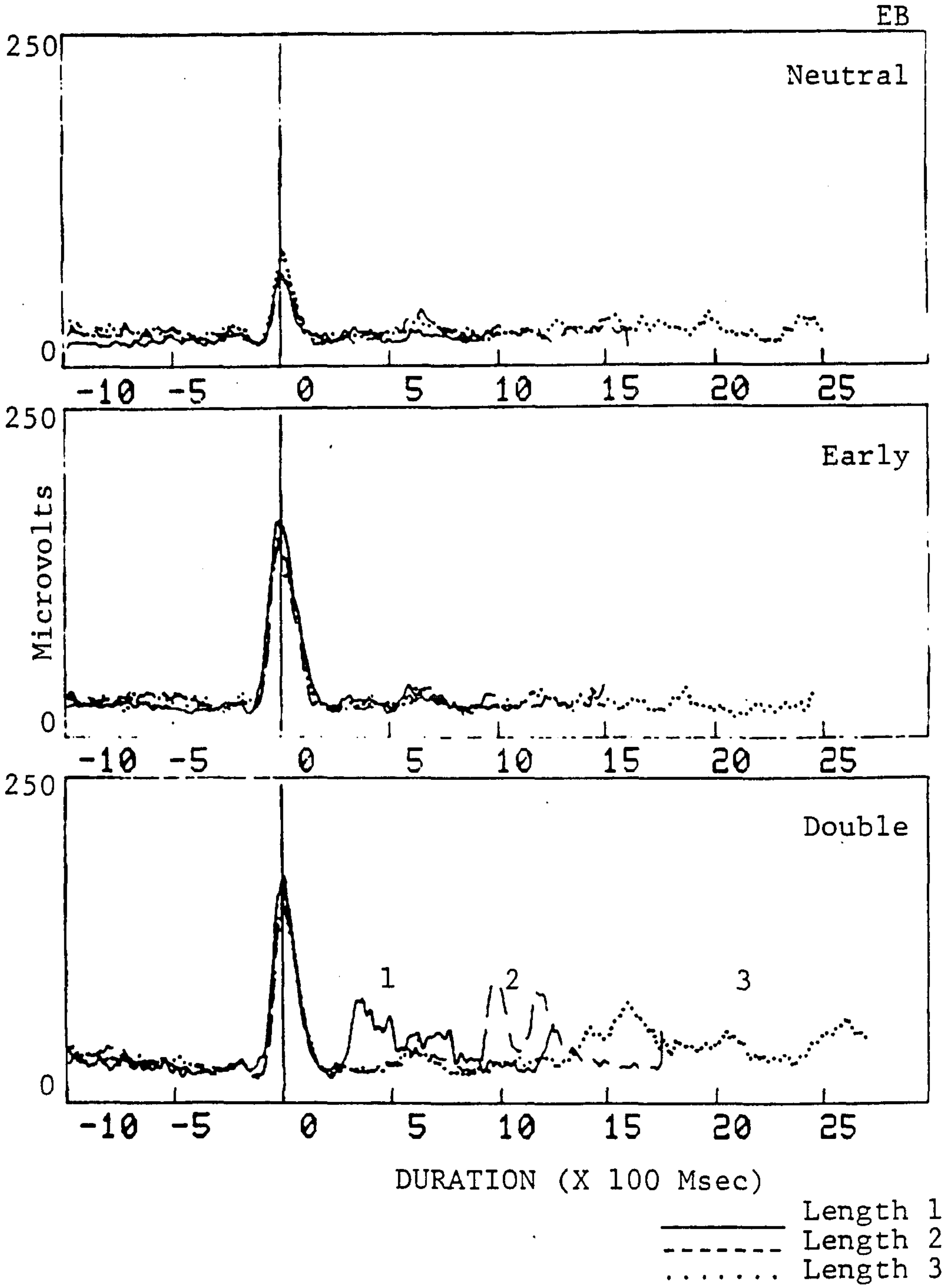


Figure 51. Comparison of CT muscle activity across utterance lengths for EB.

The most striking feature of the CT data is the on-off (i.e., binary) nature of CT activity and the absence of a gradual decline in the level of that activity. That is, there is significant activity associated with emphatically stressed syllables, with CT activity decreasing to baseline level between peaks.

Expiration: Corresponding Respitrace data are displayed in Figures 52-55 for RC and RC English. These traces depict changes in lung volume over time. It can be seen that, while peak inspiration may vary across lengths, within the same stress and phonetic condition, the rate at which lung volume declines over utterances of various lengths is quite similar.

2. Effect of Stress Configuration

Variations in F0, Ps, CT and lung volume are examined in this section as a function of stress configuration in order to determine the influence of this variable on the organization of the declination contour or any of the variables underlying it.

Fundamental Frequency: Figures 56-58 show the effects of stress configuration on the F0 contours for each phonetic and length condition for RC. There is an obvious effect of the degree of initial emphasis on the early F0 peaks, with Double and Early stress utterances showing

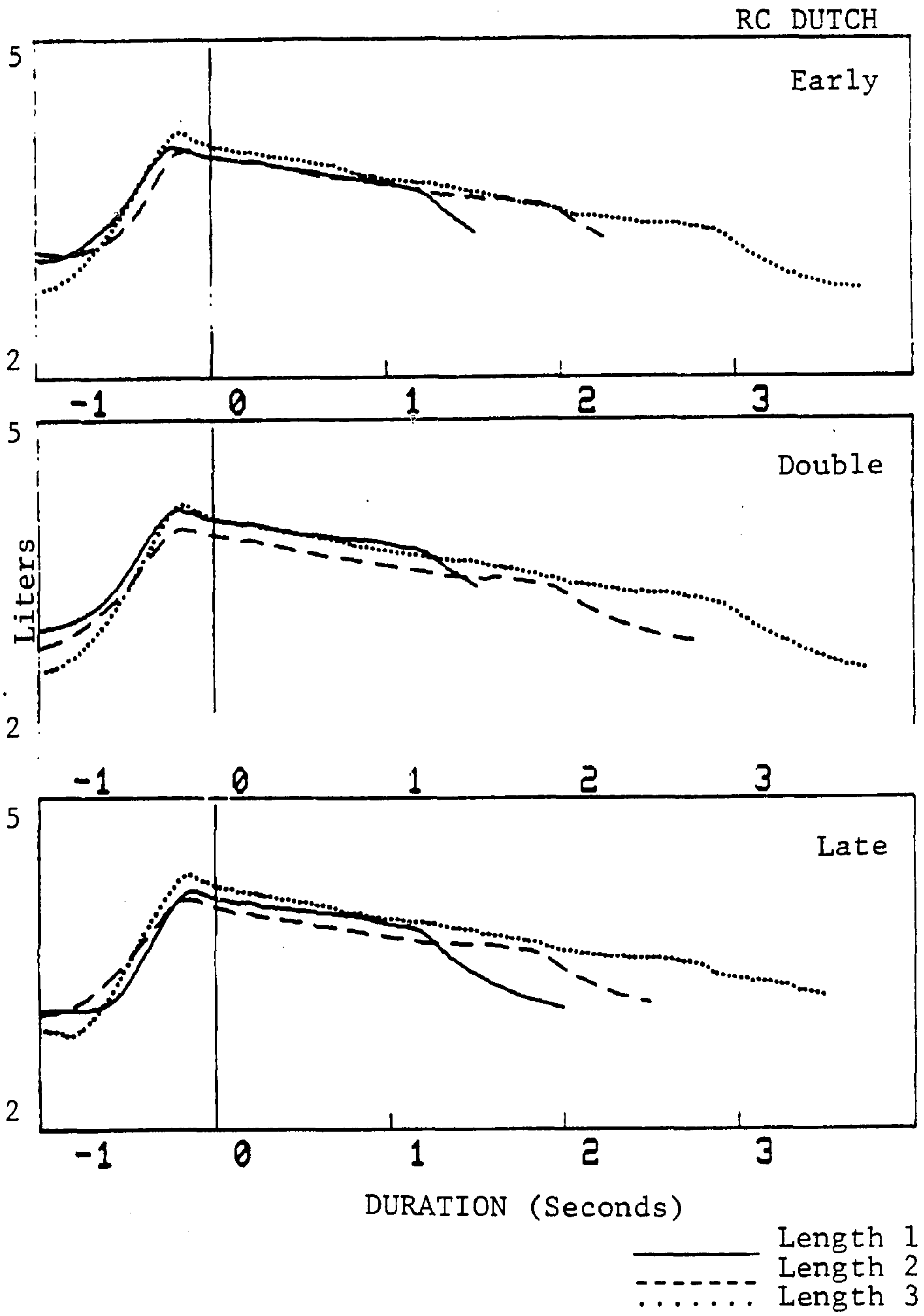


Figure 52. Comparison of lung volume changes across utterance lengths for RC Dutch.

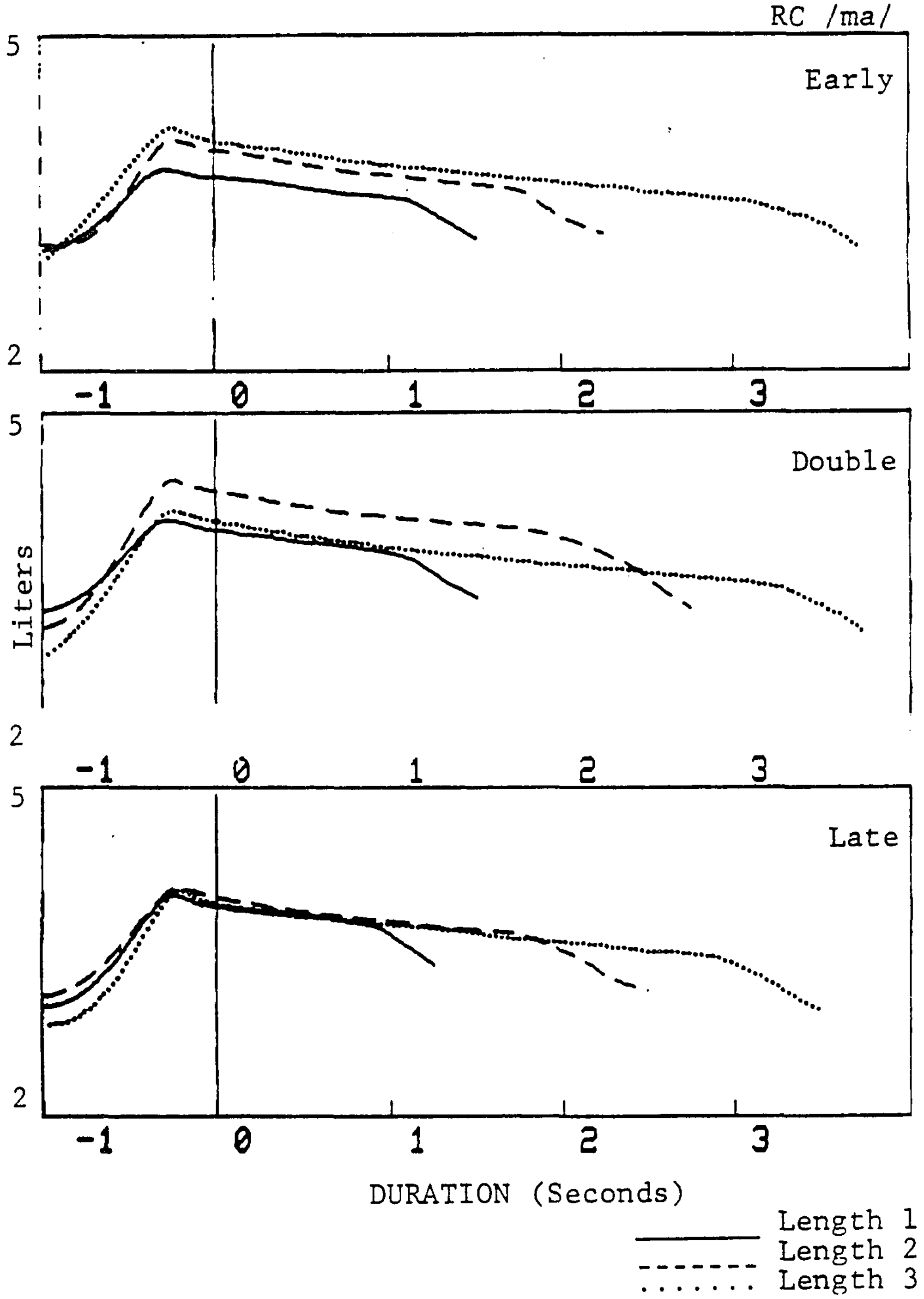


Figure 53. Comparison of lung volume changes across utterance lengths for RC /ma/.

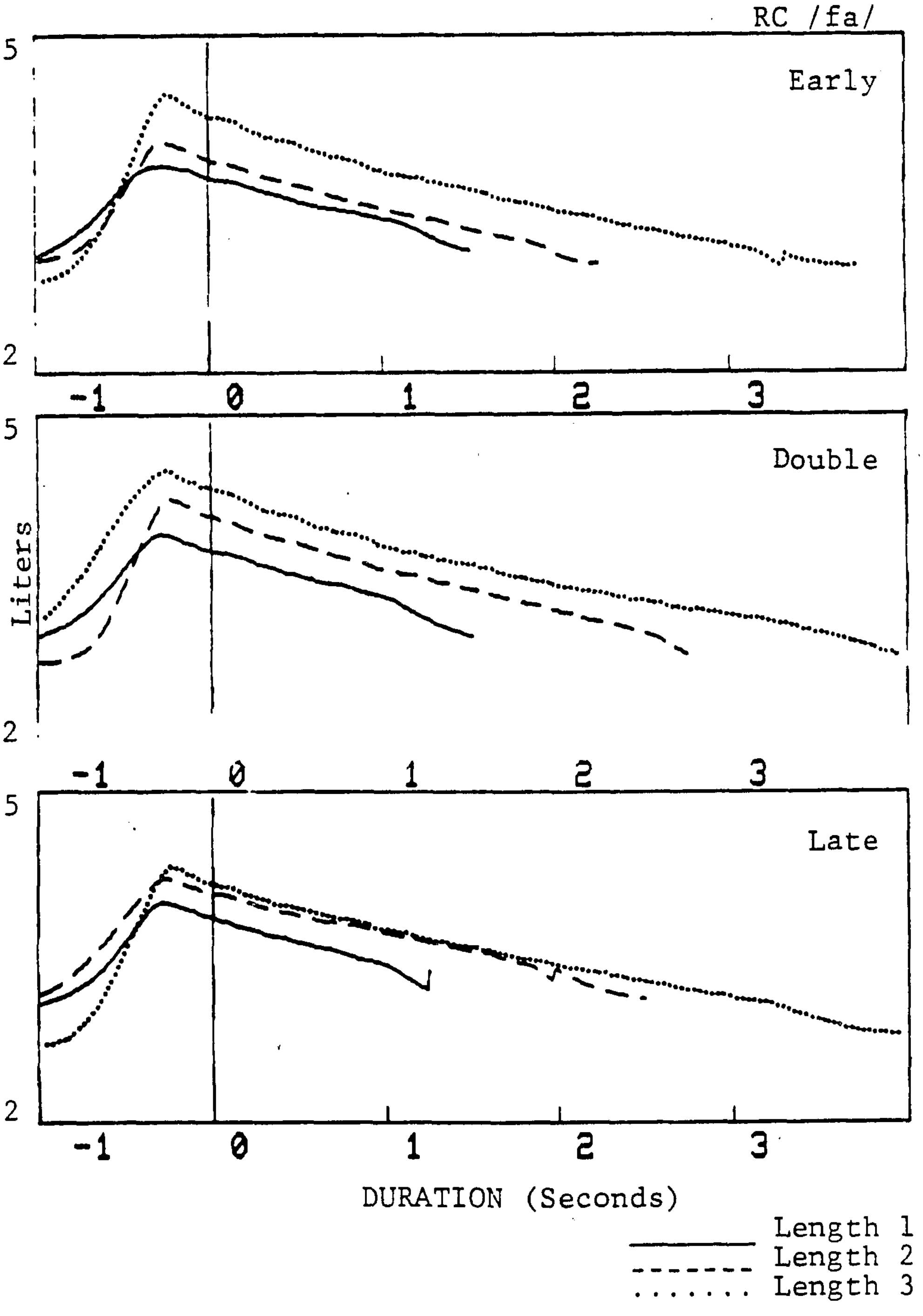


Figure 54. Comparison of lung volume changes across utterance lengths for RC /fa/.

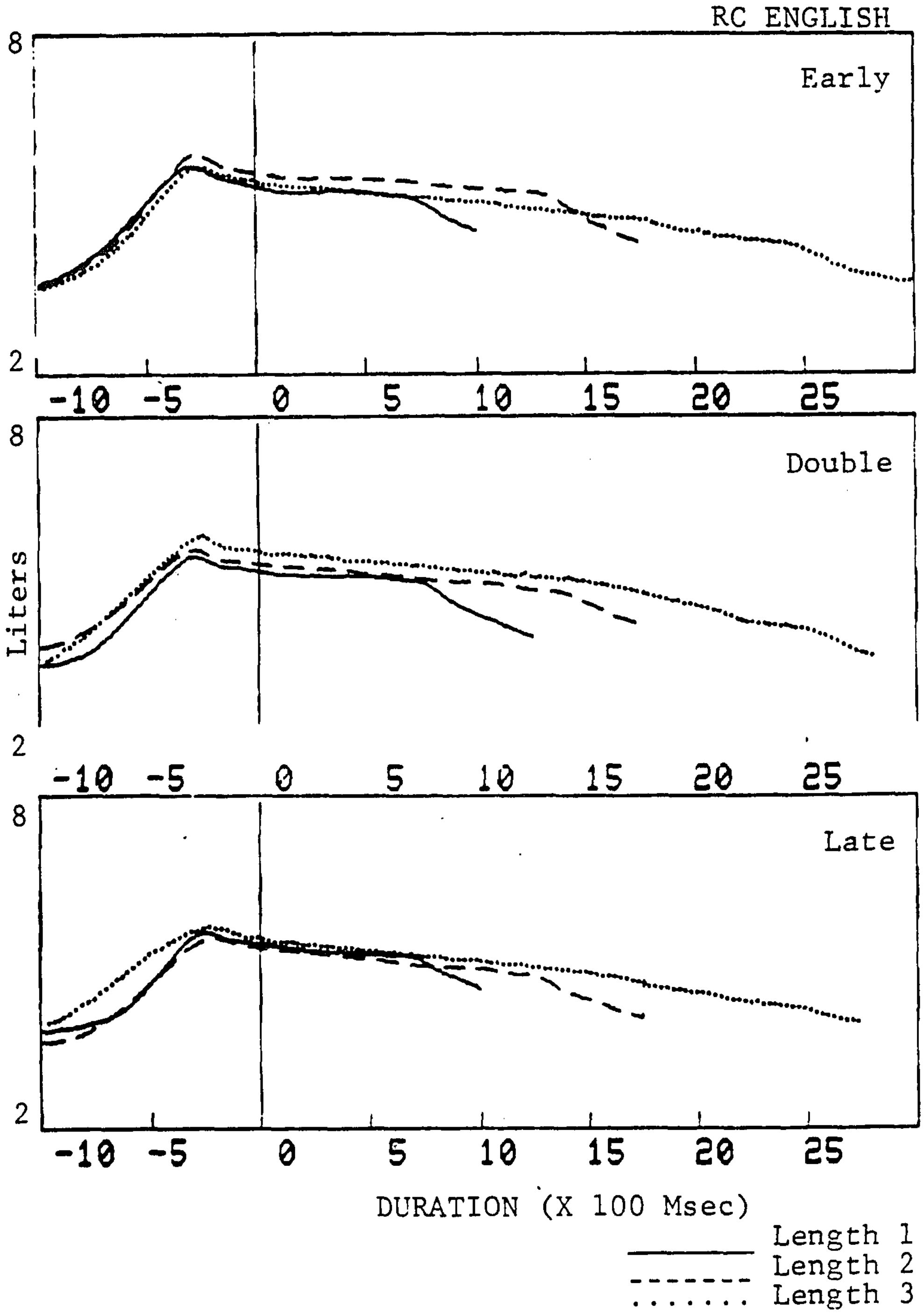


Figure 55. Comparison of lung volume changes across utterance lengths for RC English.

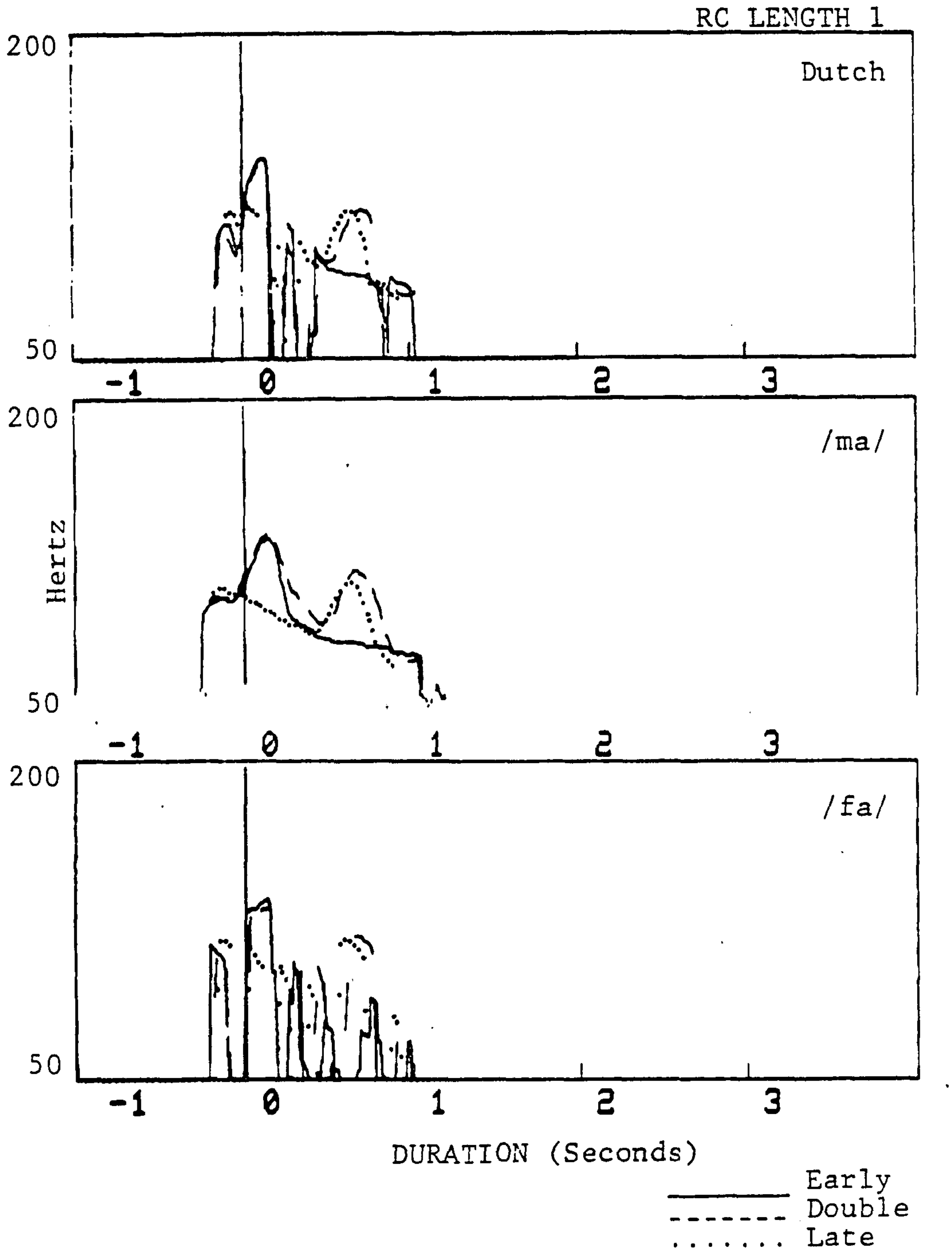


Figure 56. Comparison of F0 contours across stress configurations for RC Length 1 utterances.

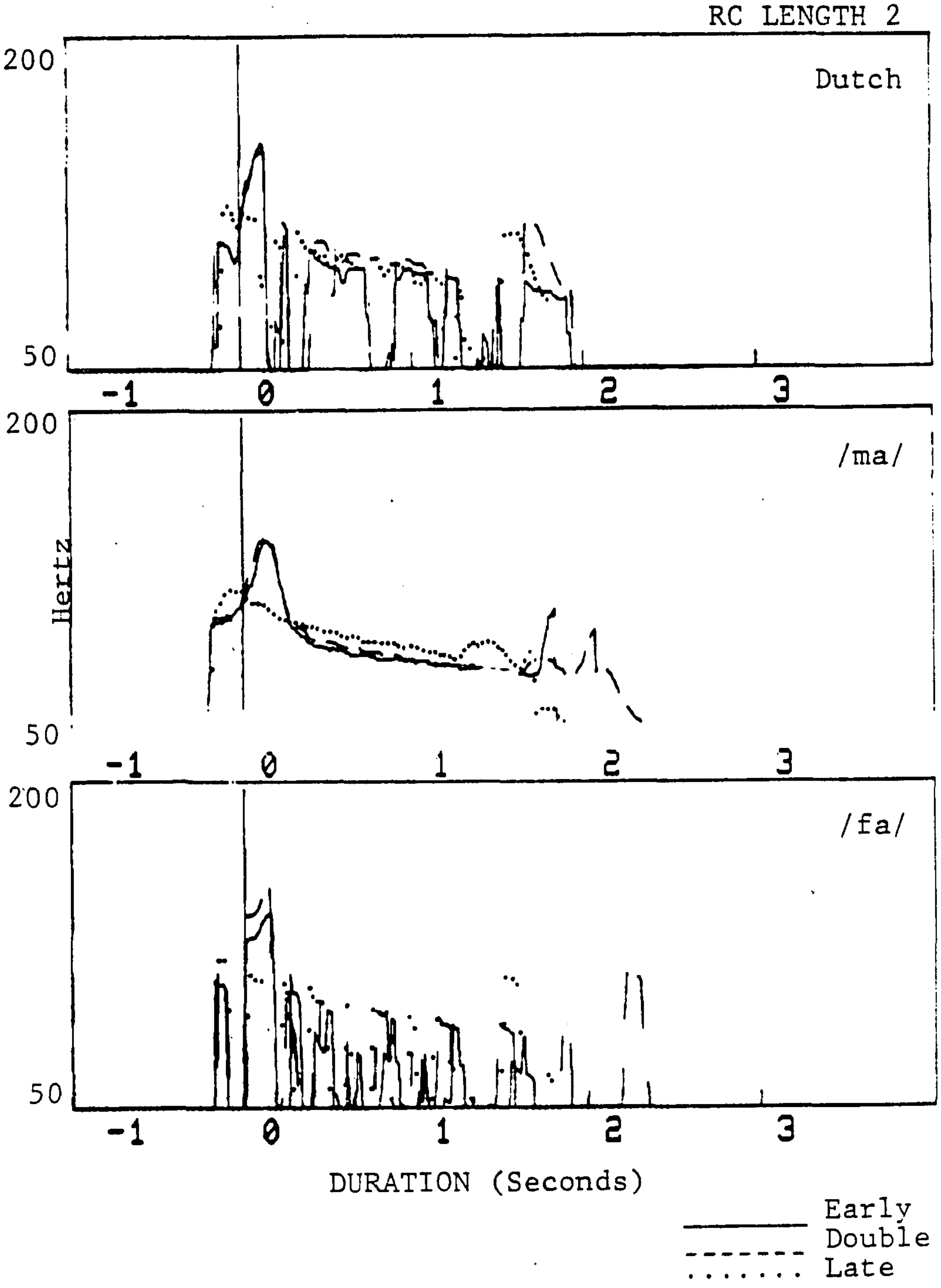


Figure 57. Comparison of F0 contours across stress configurations for RC Length 2 utterances.

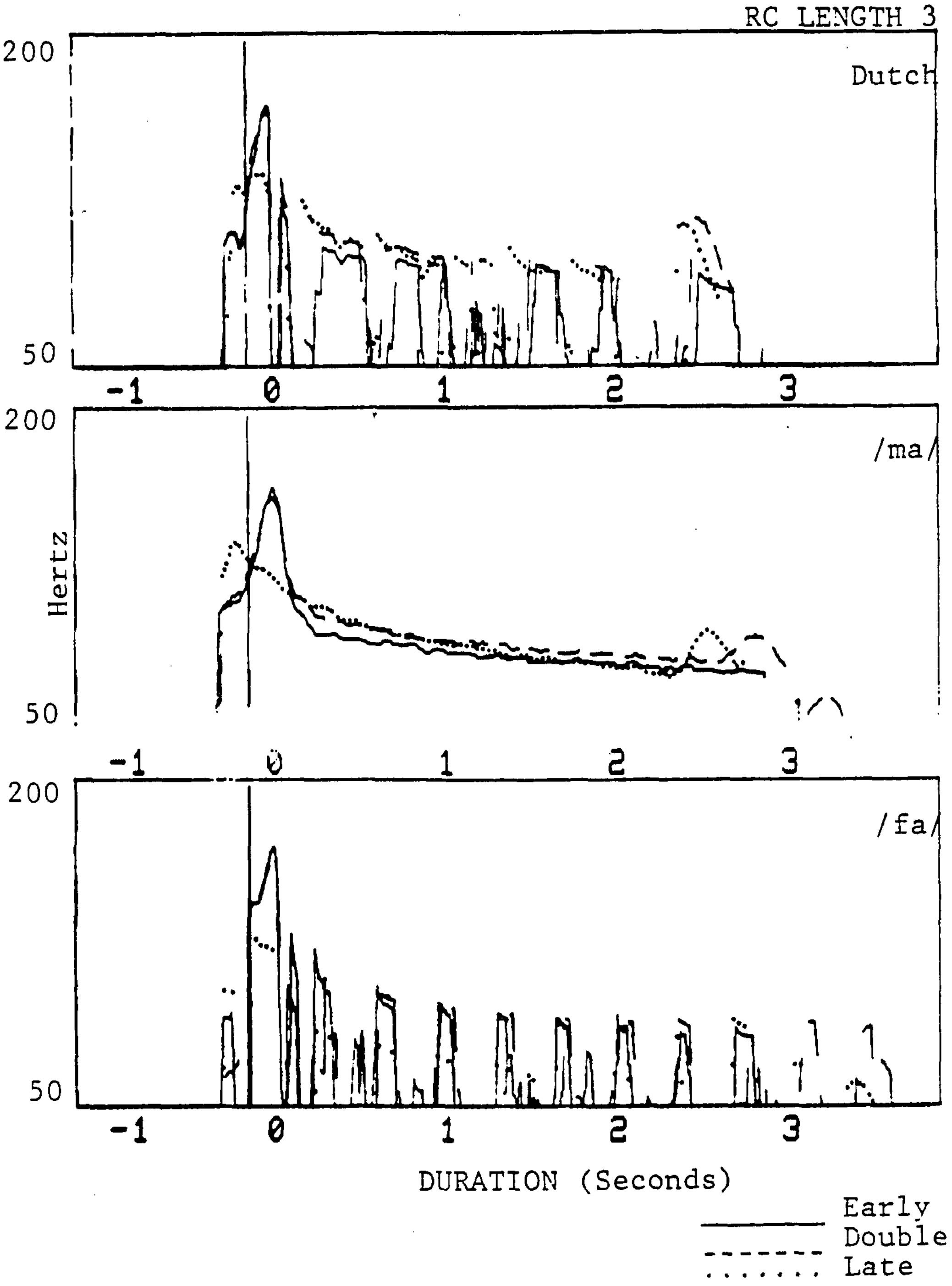


Figure 58. Comparison of F0 contours across stress configurations for RC Length 3 utterances.

higher values than Late stress utterances. However, the influence of the height of the Early and Double stress utterance peaks appears to be limited to the portions of the utterances in which they occur. That is, there is no apparent reorganization of the F0 contour on the basis of those stress configurations. The Late stress utterances, on the other hand, differ from the other two in that their peaks occur earlier and the subsequent decline in F0 is more gradual. However, by the time the initial rapid fall in F0 for the Early and Double stress utterances occurs, the decline in the F0 contours for the three stress configurations is quite similar.

This pattern is also apparent in the F0 curves for the various stress configurations for RC English and EB (Figs. 59 & 60), with the exception, perhaps, of the Length 2 utterances for RC English. That is, apart from the differences in peak frequency between those utterances that receive early stress and those that do not, the F0 curves are quite similar beyond the initial peaks.

Subglottal Pressure: The corresponding pressure curves are shown in Figures 61-63 for RC, in Figure 64 for RC English and in Figure 65 for EB. For both subjects, the contours merge following the initial peaks until they diverge again, depending on the final stress configuration. In some cases, the medial portions of these curves become indistinguishable from one another, while in other cases, they decline in parallel, bearing the same relationship as the initial peaks. Thus, there is no evidence for a reorganization of subglottal

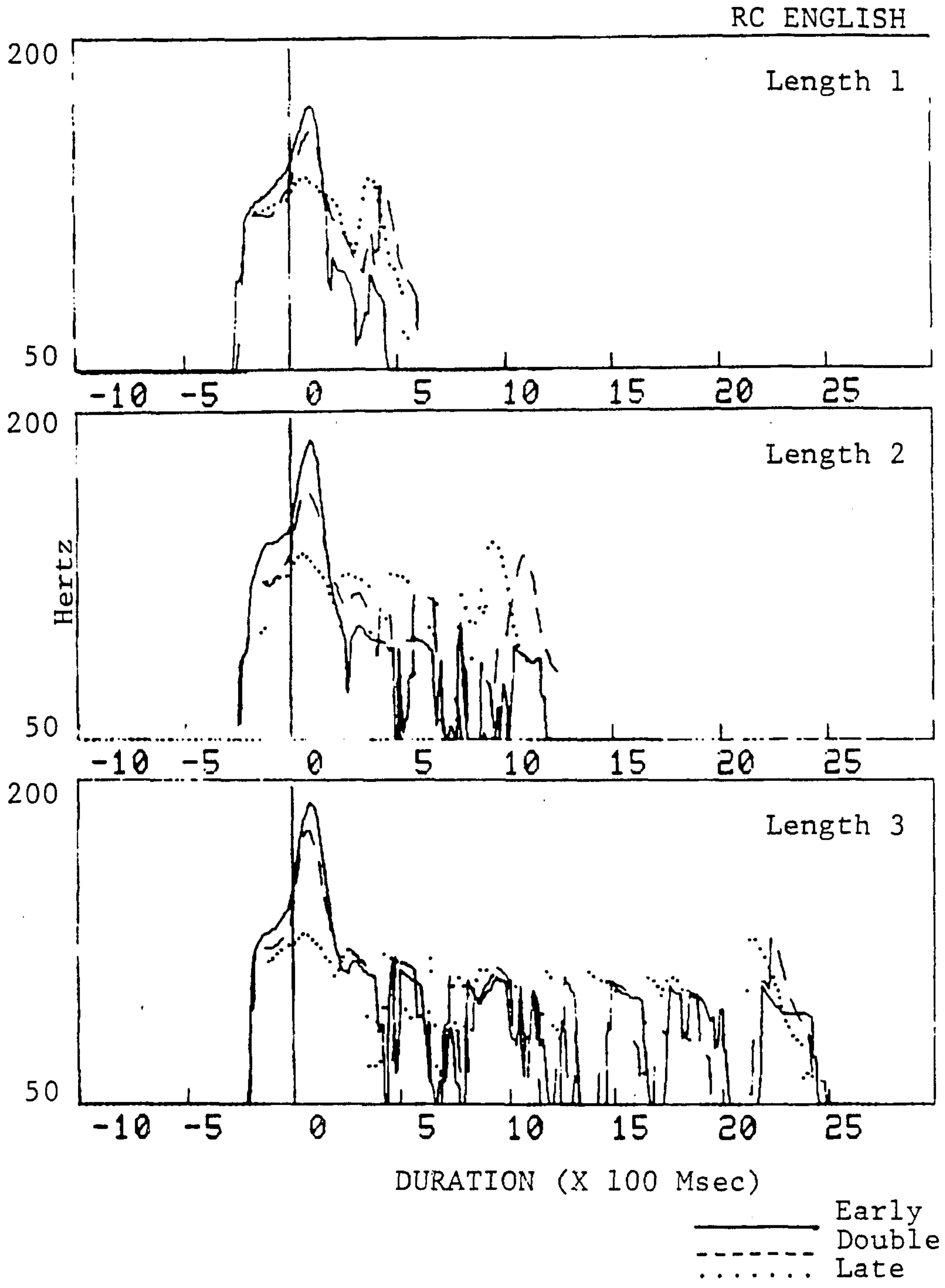


Figure 59. Comparison of F0 contours across stress configurations for RC English.

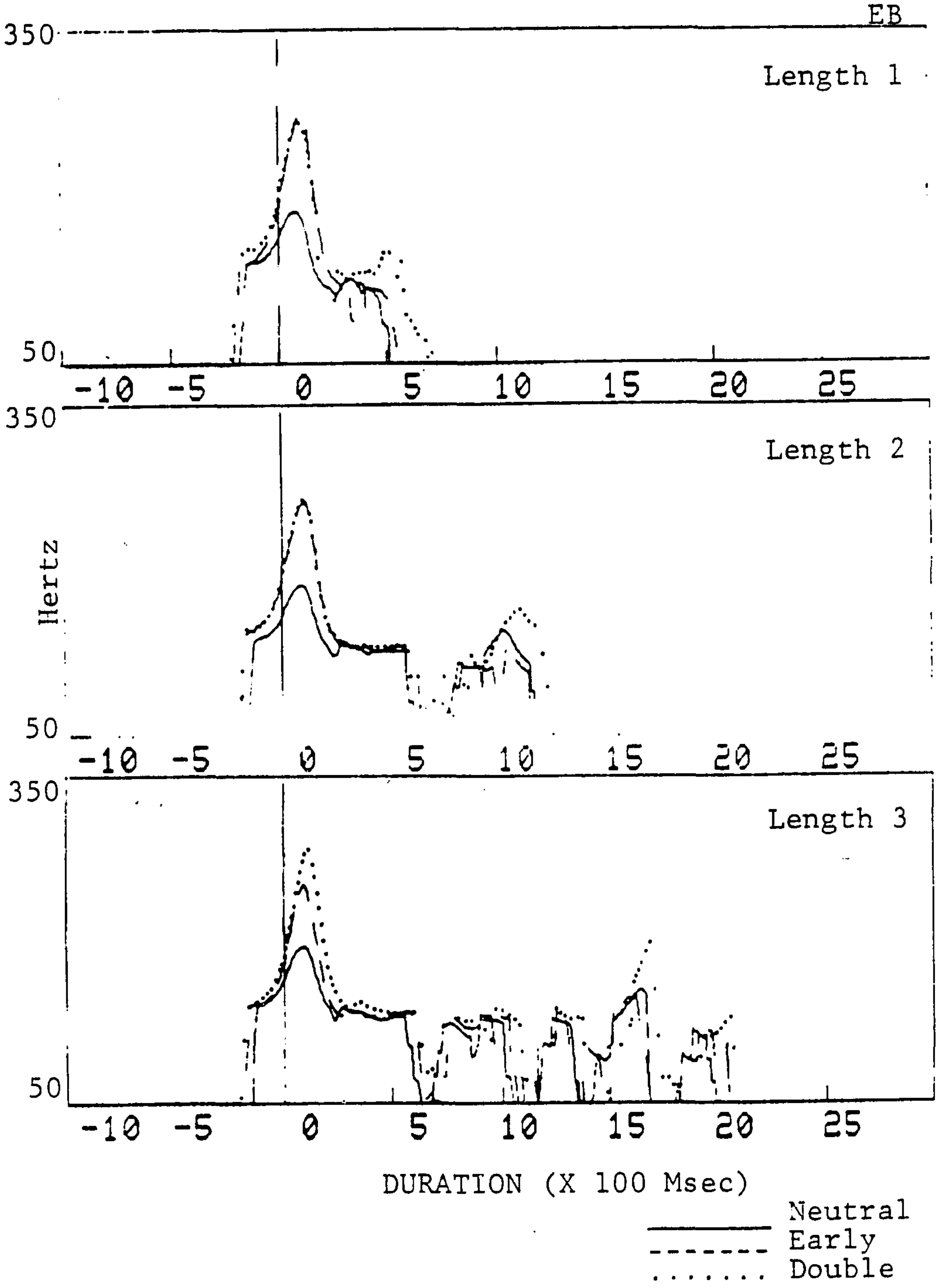


Figure 60. Comparison of F0 contours across stress configurations for EB.

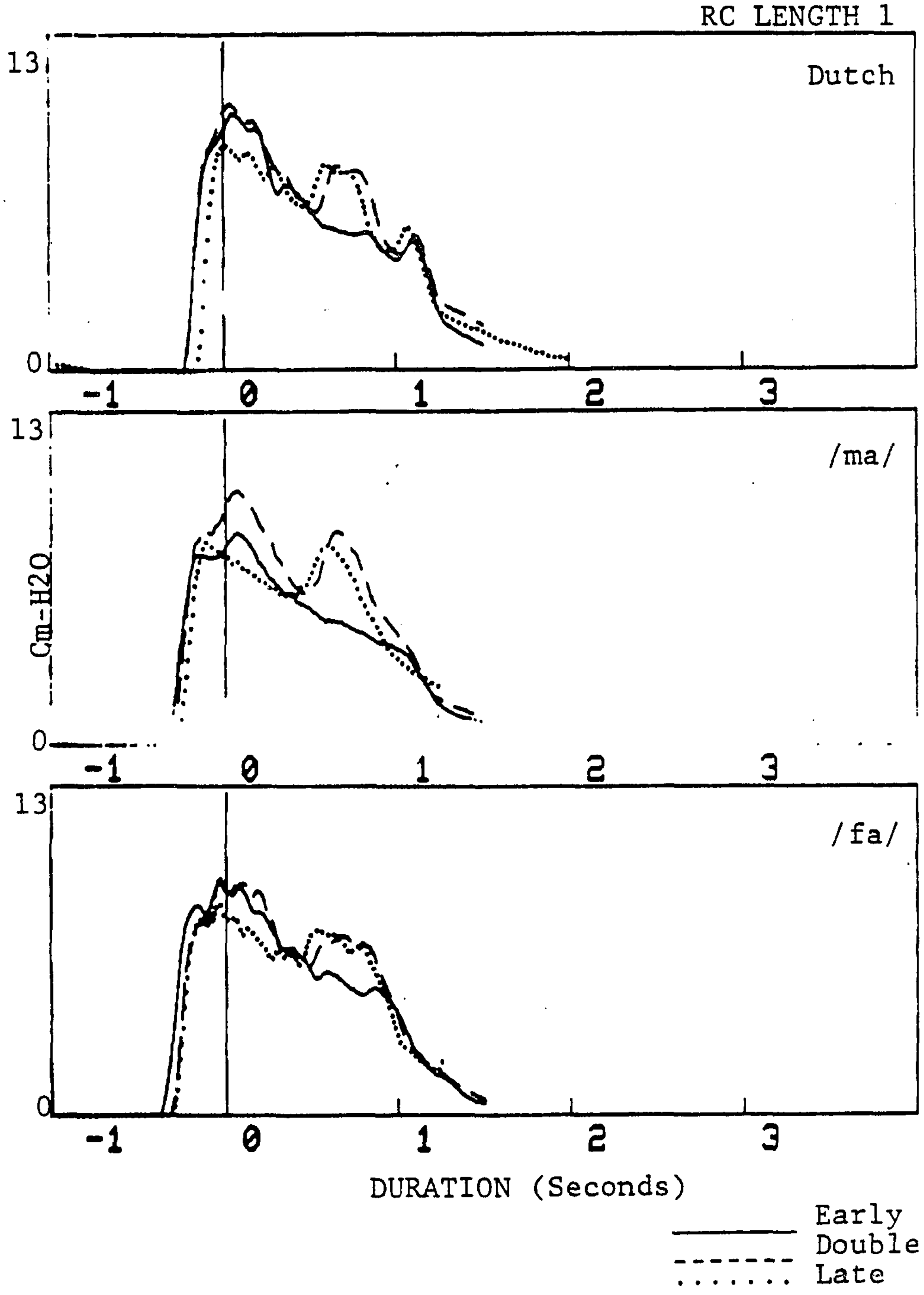


Figure 61. Comparison of Ps contours across stress configurations for RC Length 1 utterances.

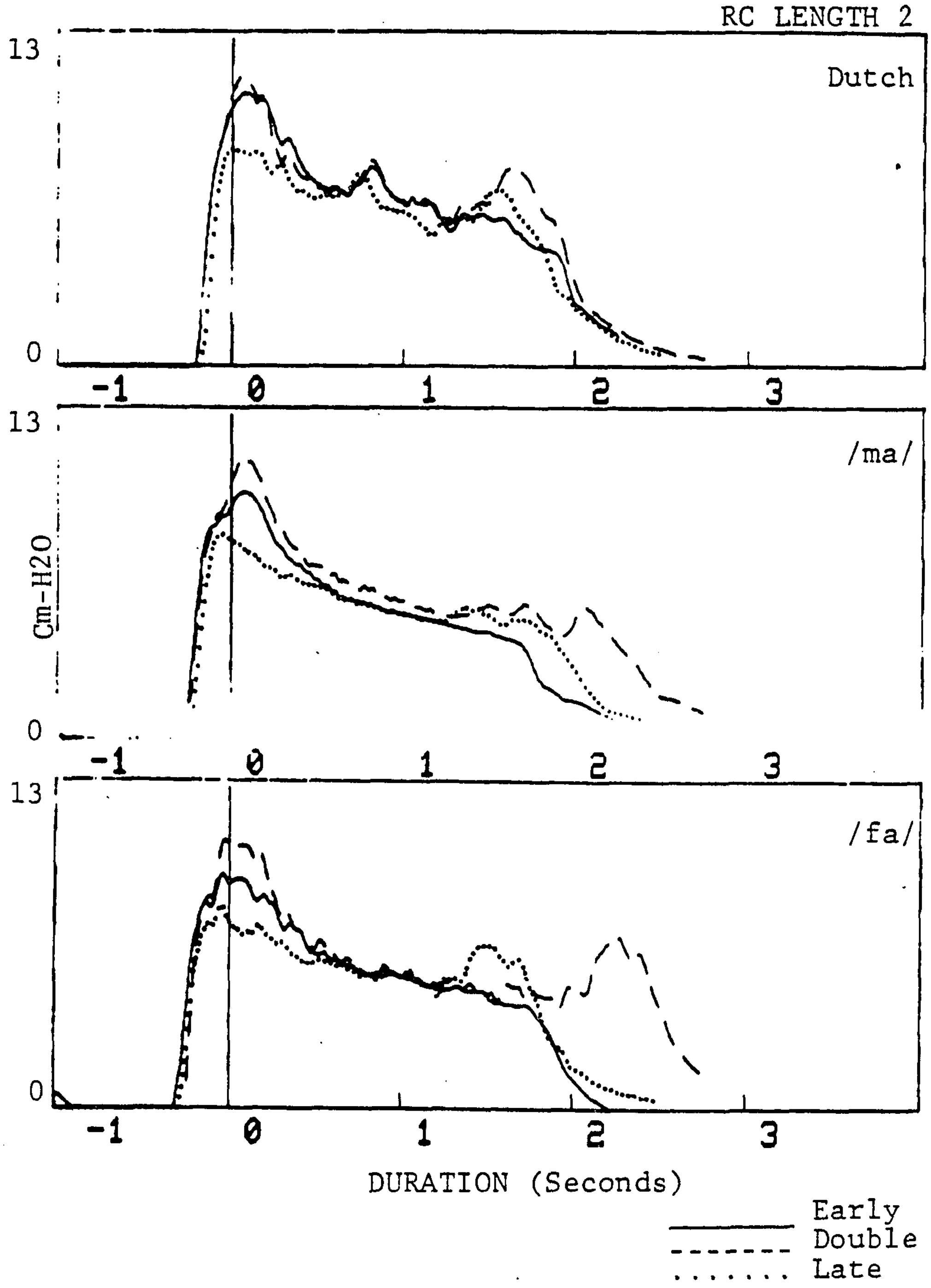


Figure 62. Comparison of Ps contours across stress configurations for RC Length 2 utterances.

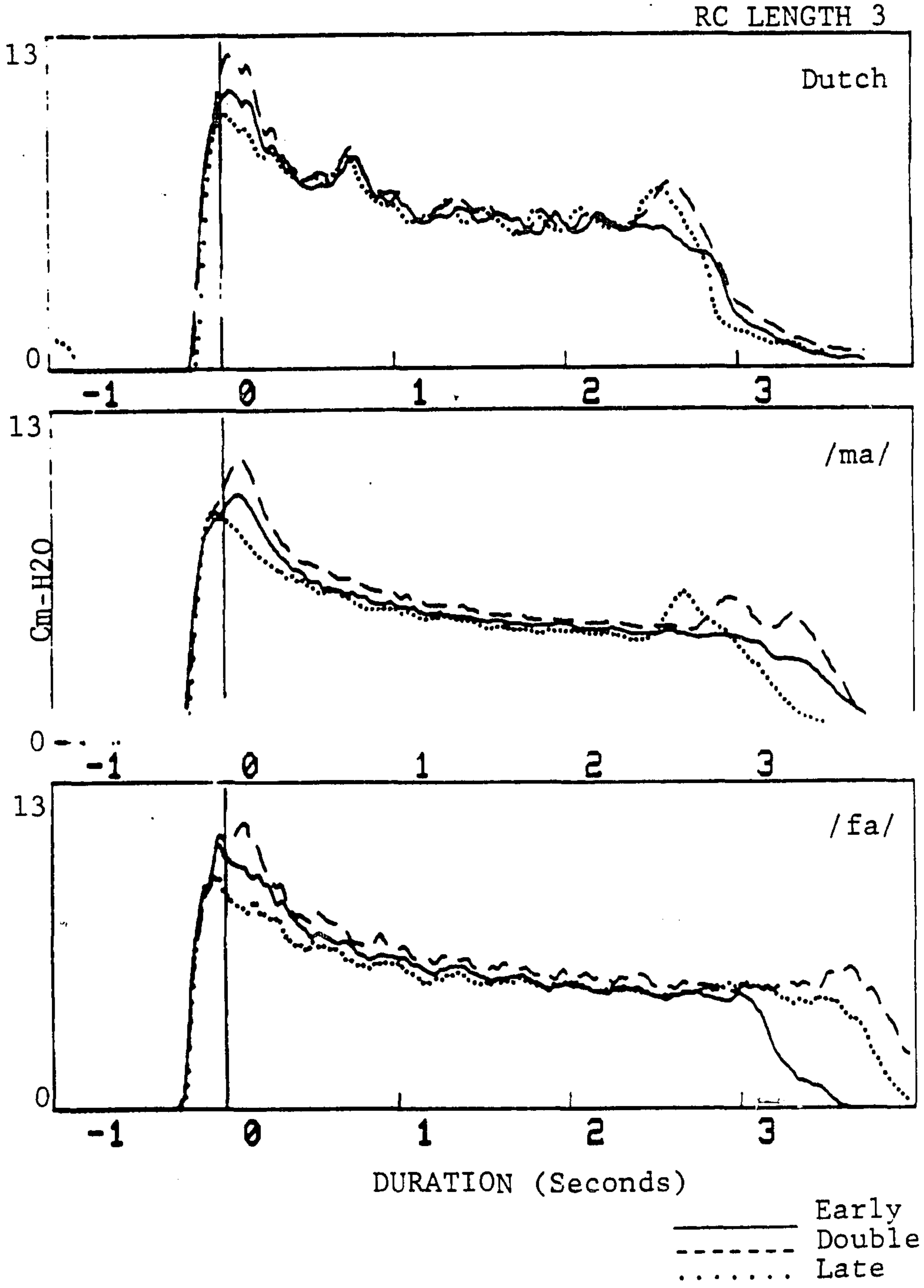


Figure 63. Comparison of Ps contours across stress configurations for RC Length 3 utterances.

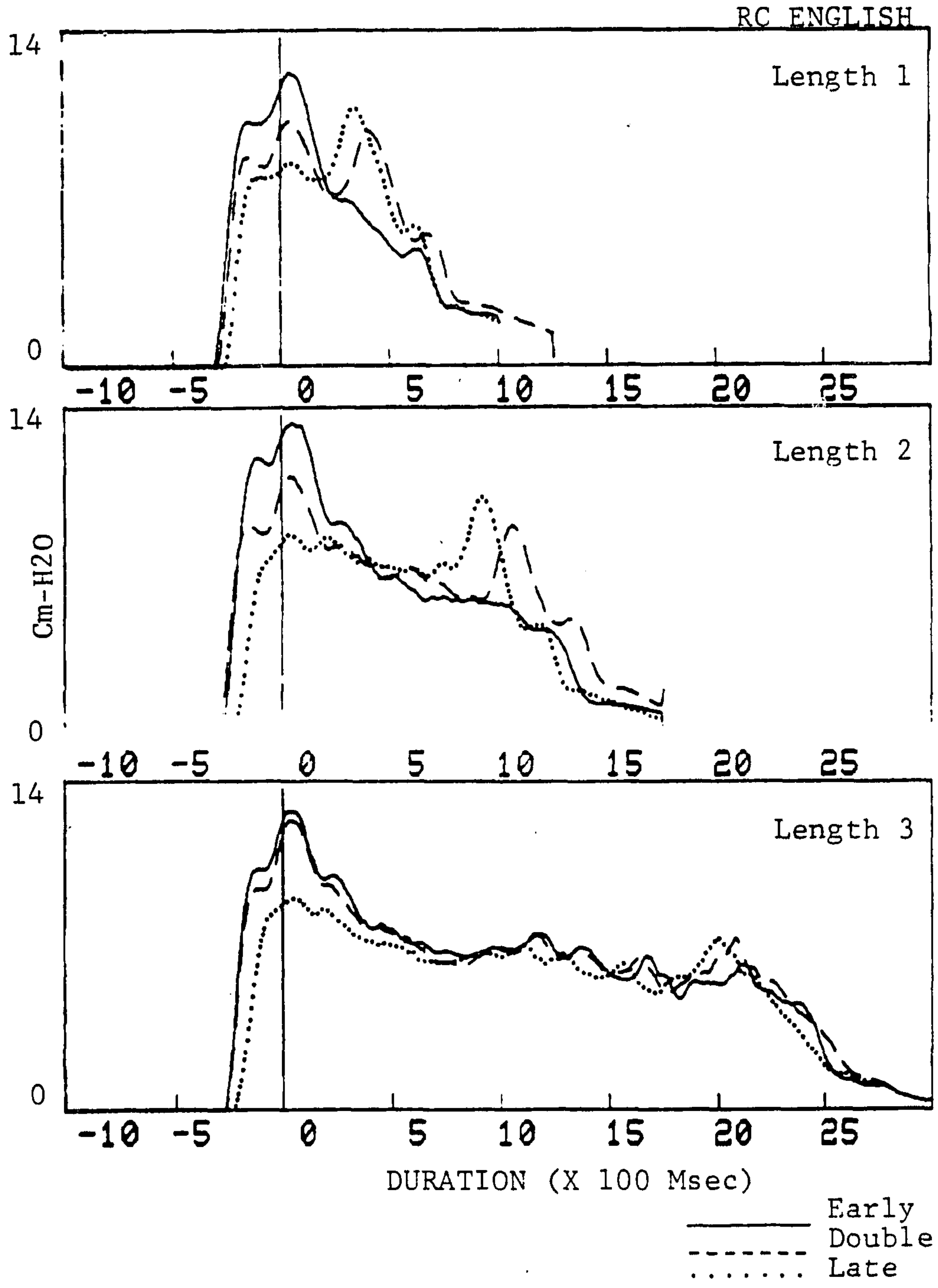


Figure 64. Comparison of Ps contours across stress configurations for RC English.

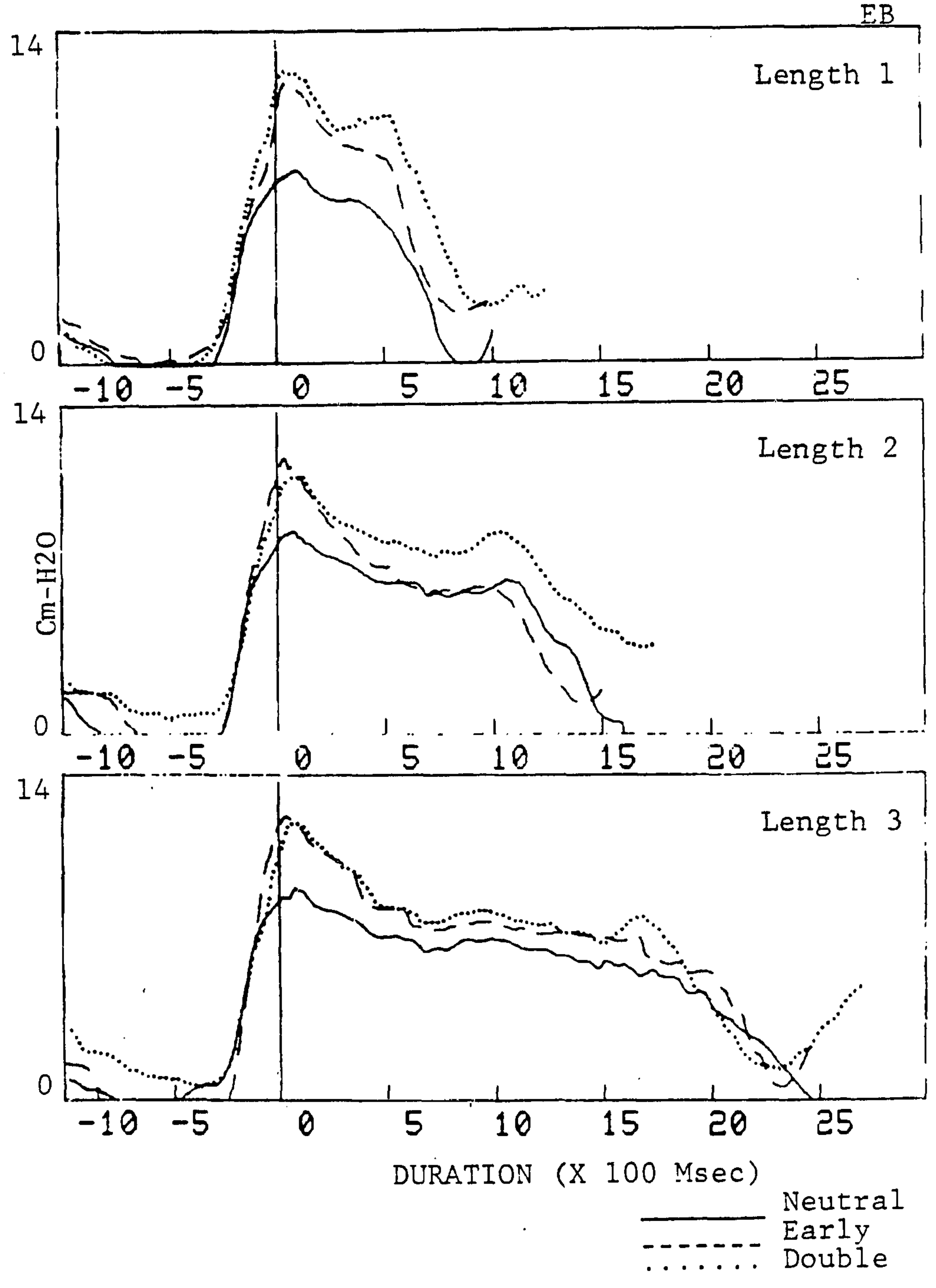


Figure 65. Comparison of Ps contours across stress configurations for EB.

pressure as a function of stress configuration.

Cricothyroid Muscle Activity: It is again more difficult to separate the individual utterance types from the superimposed cricothyroid data, shown in Figures 66-70 for both subjects. However, it is possible to identify the peaks associated with utterance initiation and the individual stress peaks. Beyond that, there are no real differences between the CT traces for the various stress conditions.

Expiration: Lung volume changes over time (Figs. 71-74) show no difference in the rate of decline as a function of stress configuration for RC or RC English.

3. The Effect of Phonetic Composition

The analyses in this section are limited to RC's Dutch, /ma/ and /fa/ data and compare the effects of phonetic composition on F0, Ps, CT and rate of expiration. The order in which these variables are discussed is rearranged, for reasons which will become apparent.

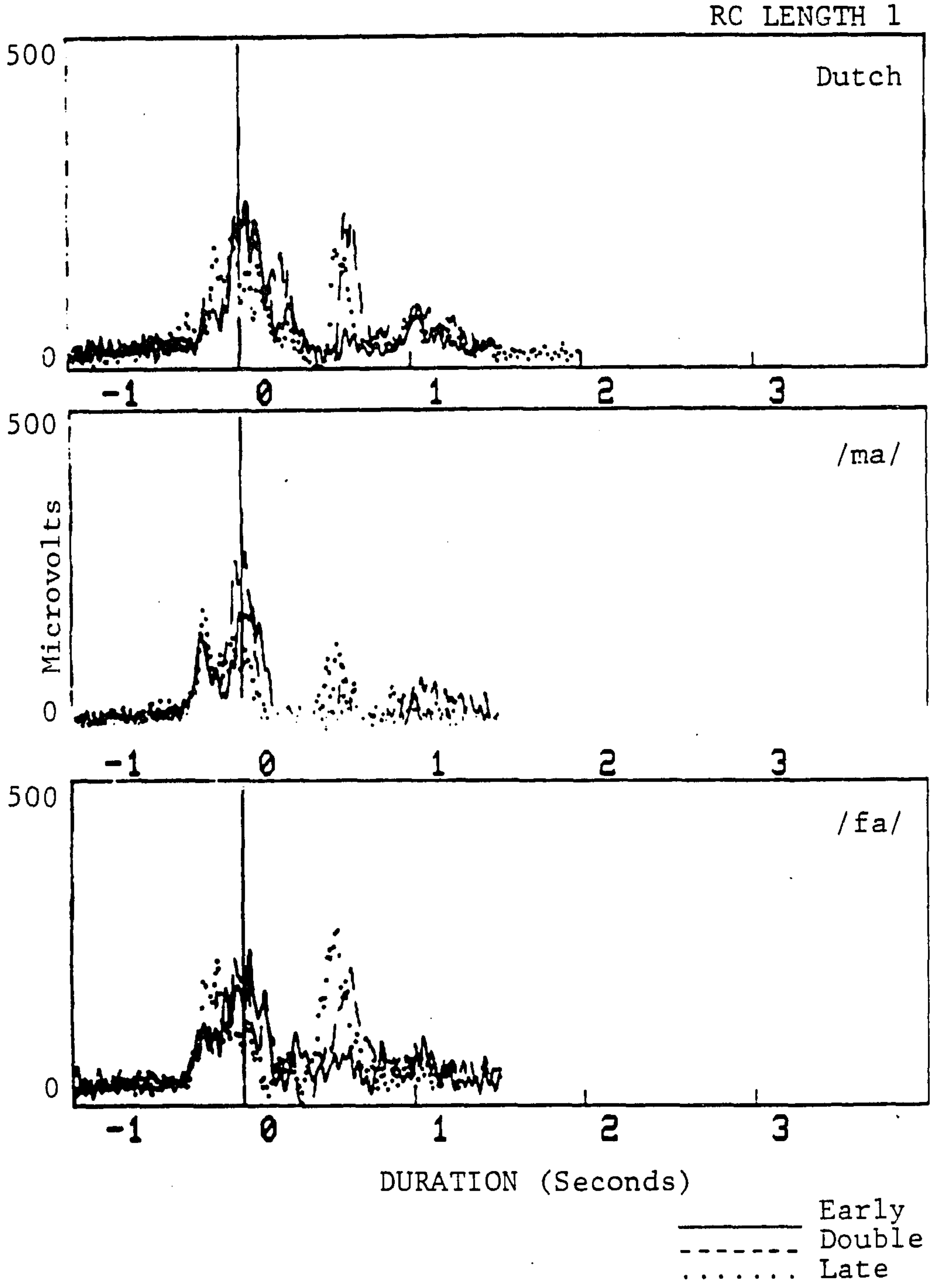


Figure 66. Comparison of CT activity across stress configurations for RC Length 1 utterances.

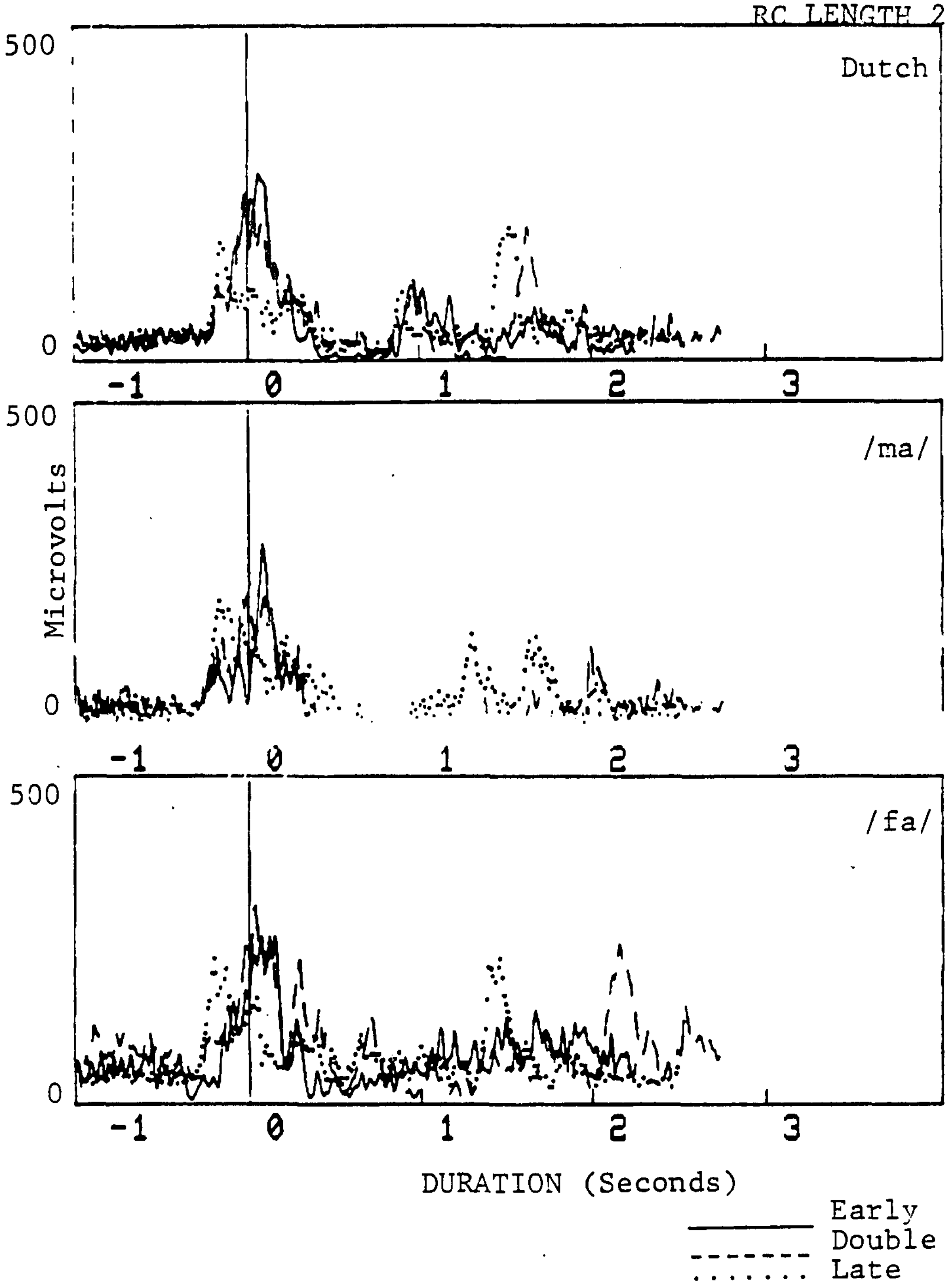


Figure 67. Comparison of CT activity across stress configurations for RC Length 2 utterances.

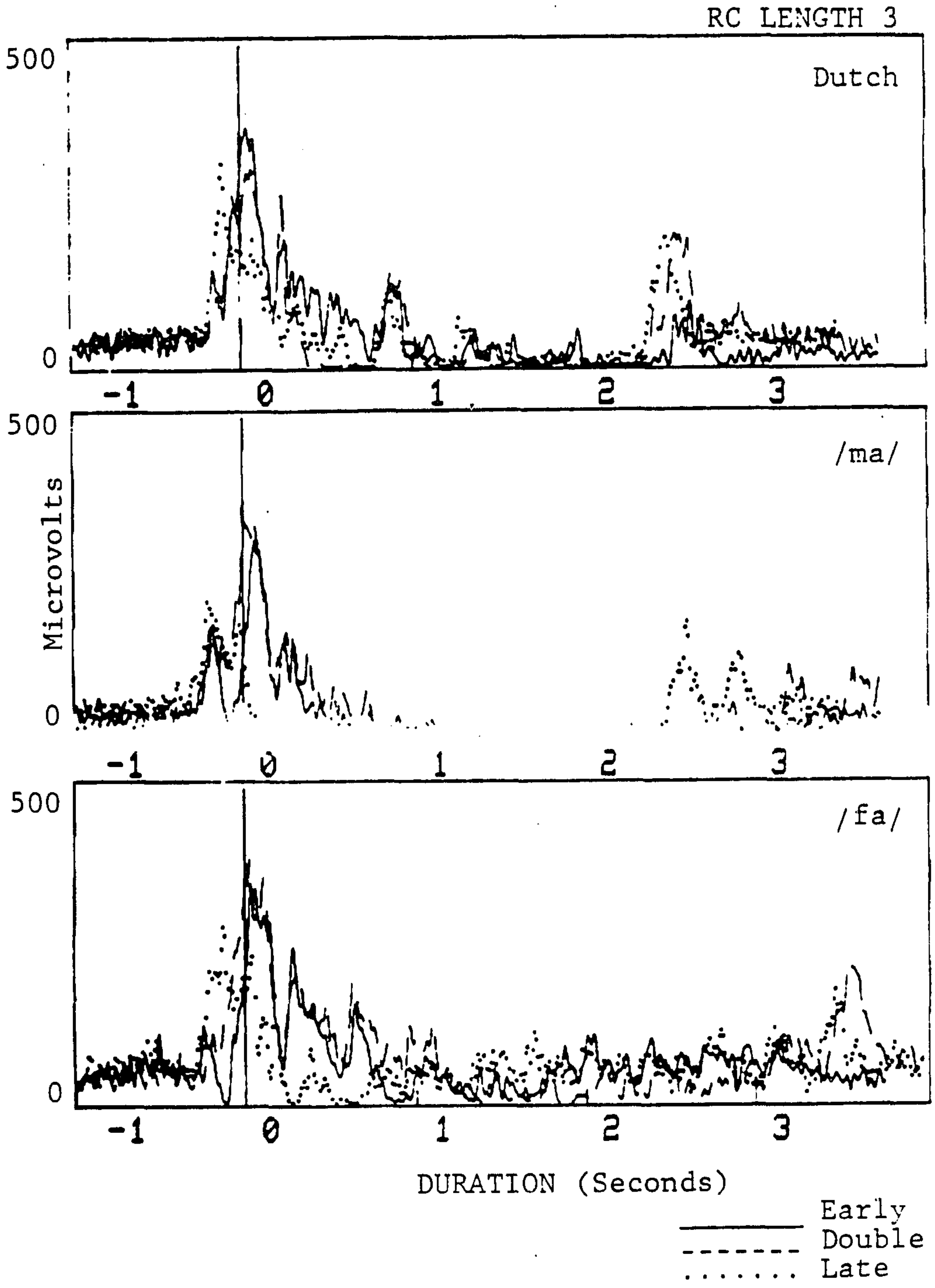


Figure 68. Comparison of CT activity across stress configurations for RC Length 3 utterances.

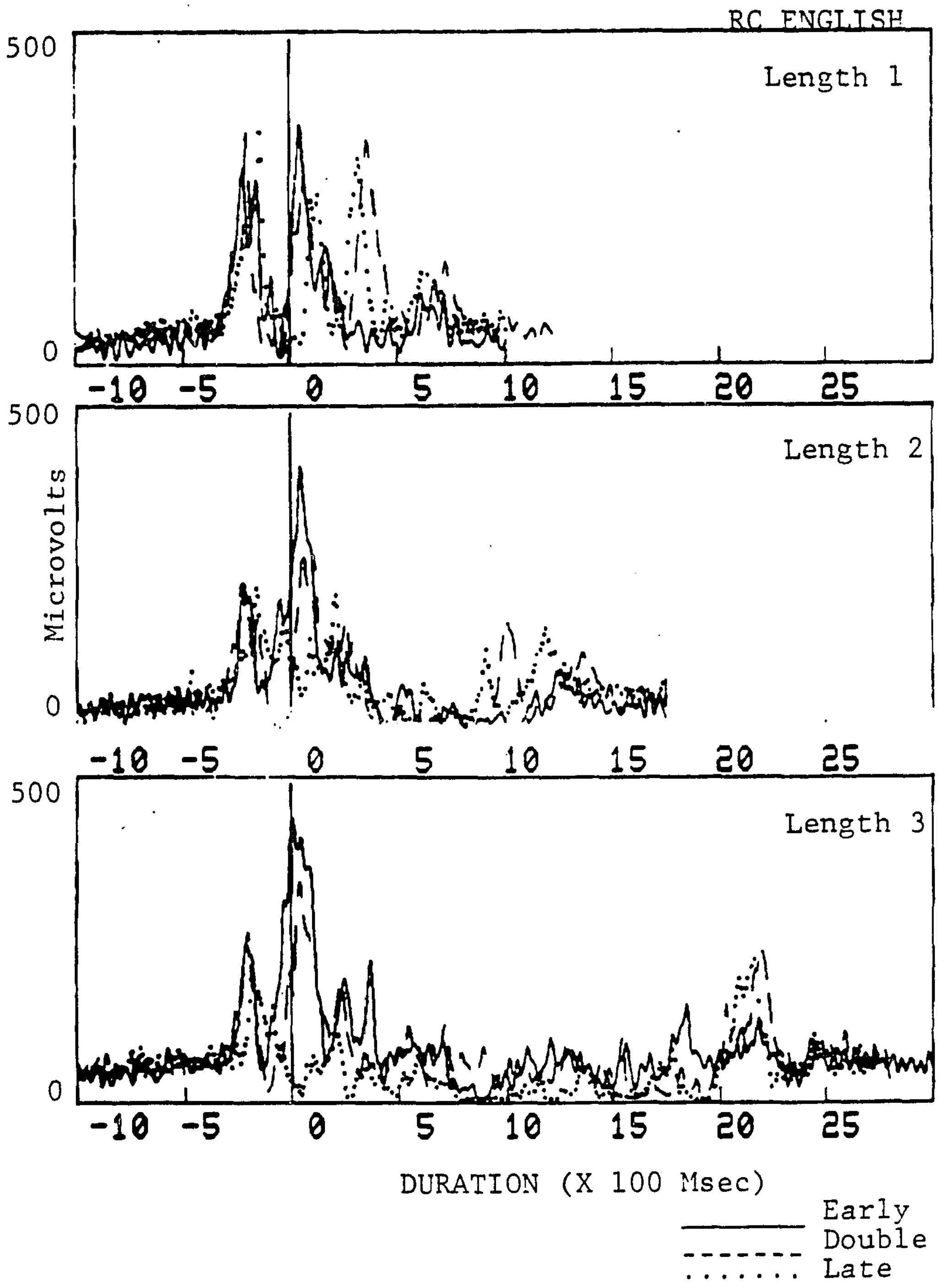


Figure 69. Comparison of CT activity across stress configurations for RC English.

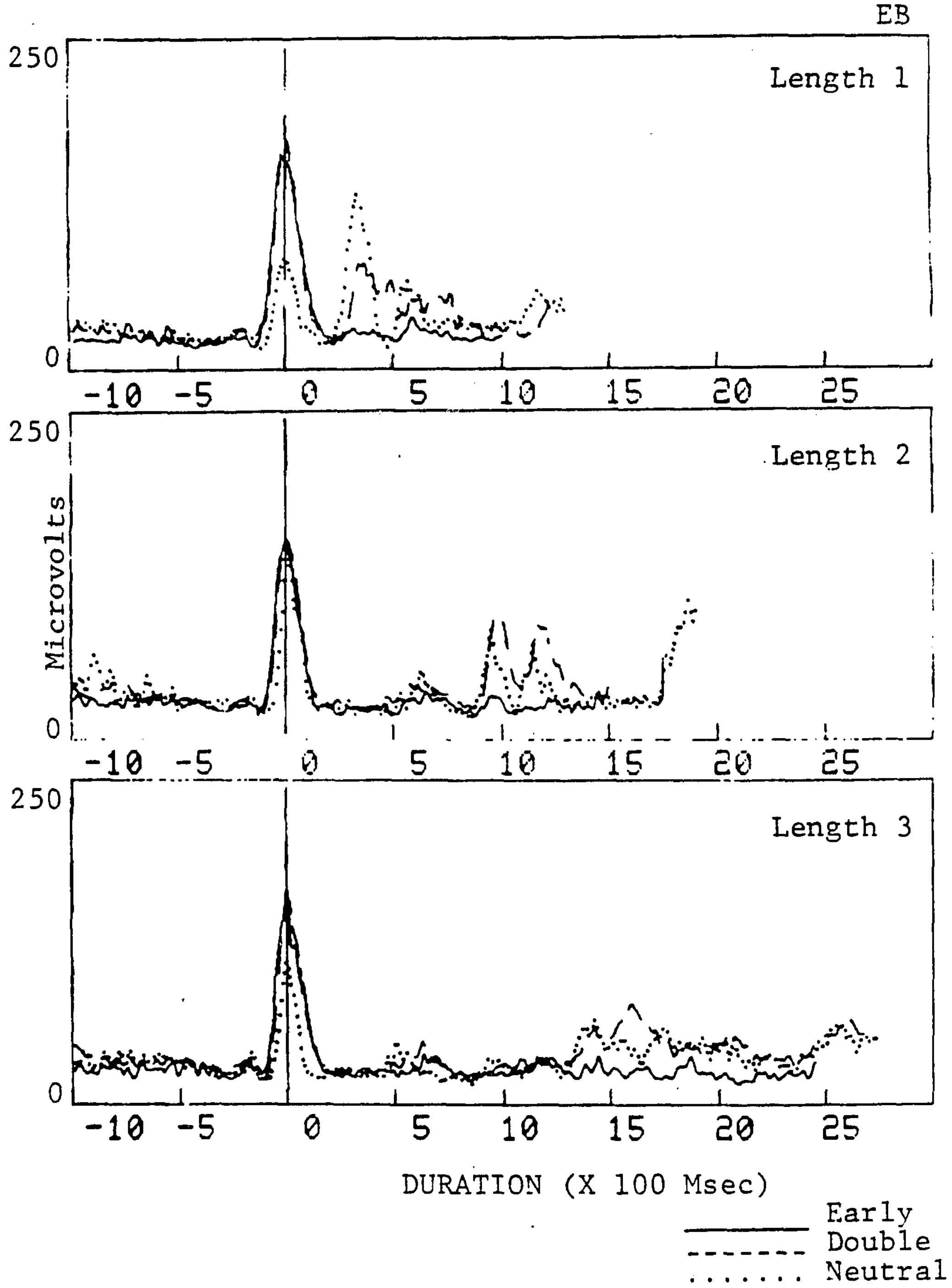


Figure 70. Comparison of CT activity across stress configurations for EB.

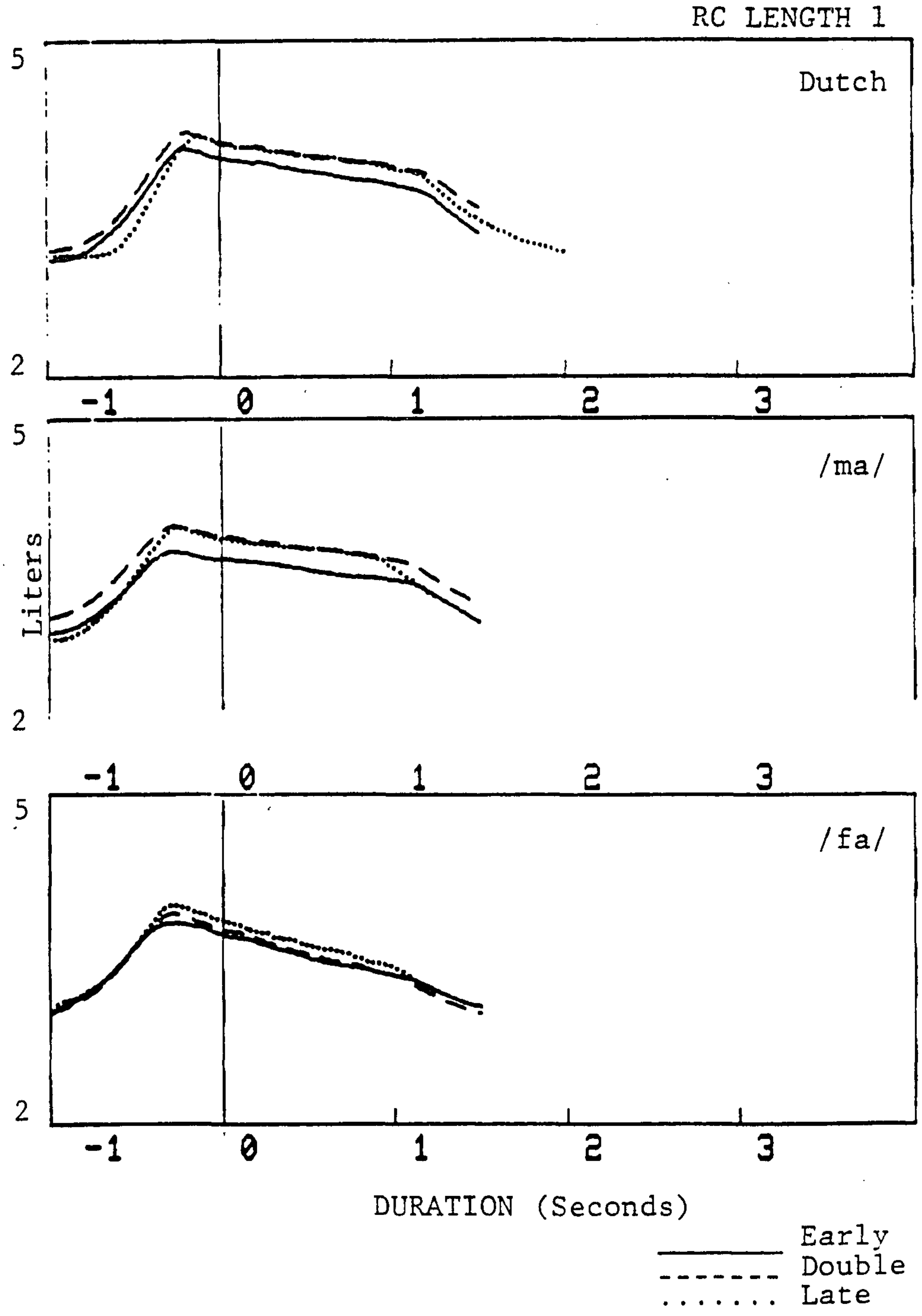


Figure 71. Comparison of lung volume changes across stress configurations for RC Length 1 utterances.

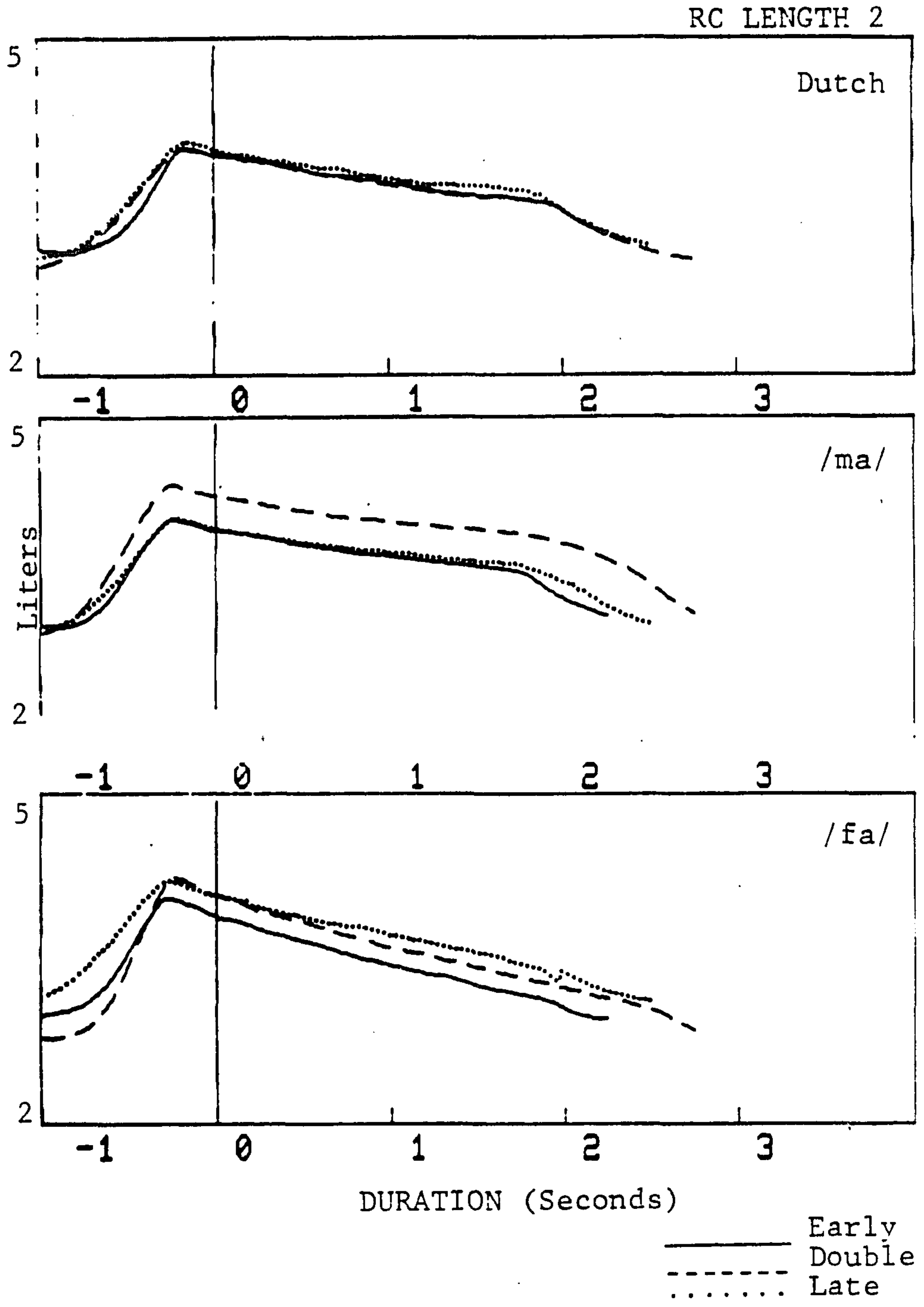


Figure 72. Comparison of lung volume changes across stress configurations for RC Length 2 utterances.

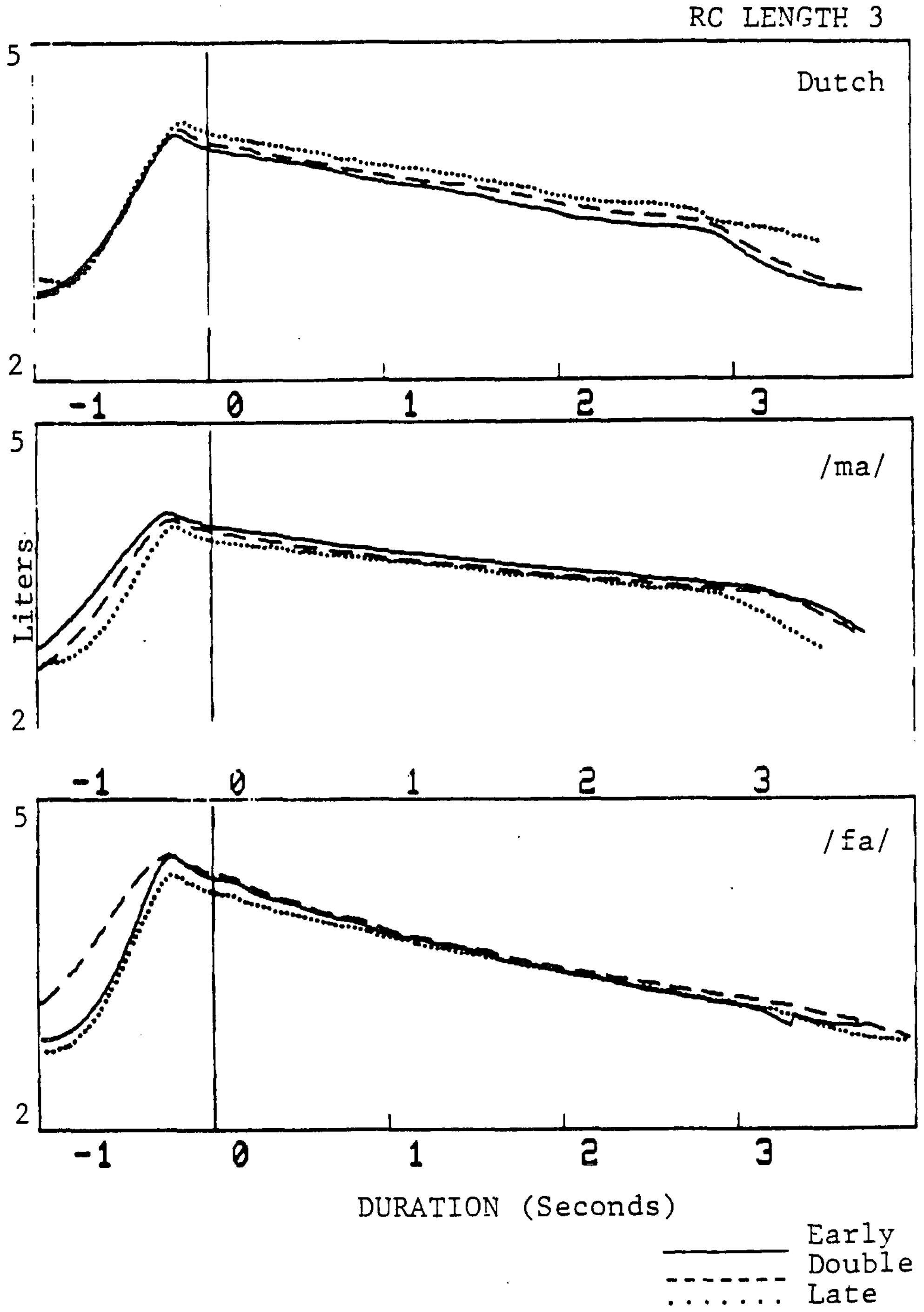


Figure 73. Comparison of lung volume changes across stress configurations for RC Length 3 utterances.

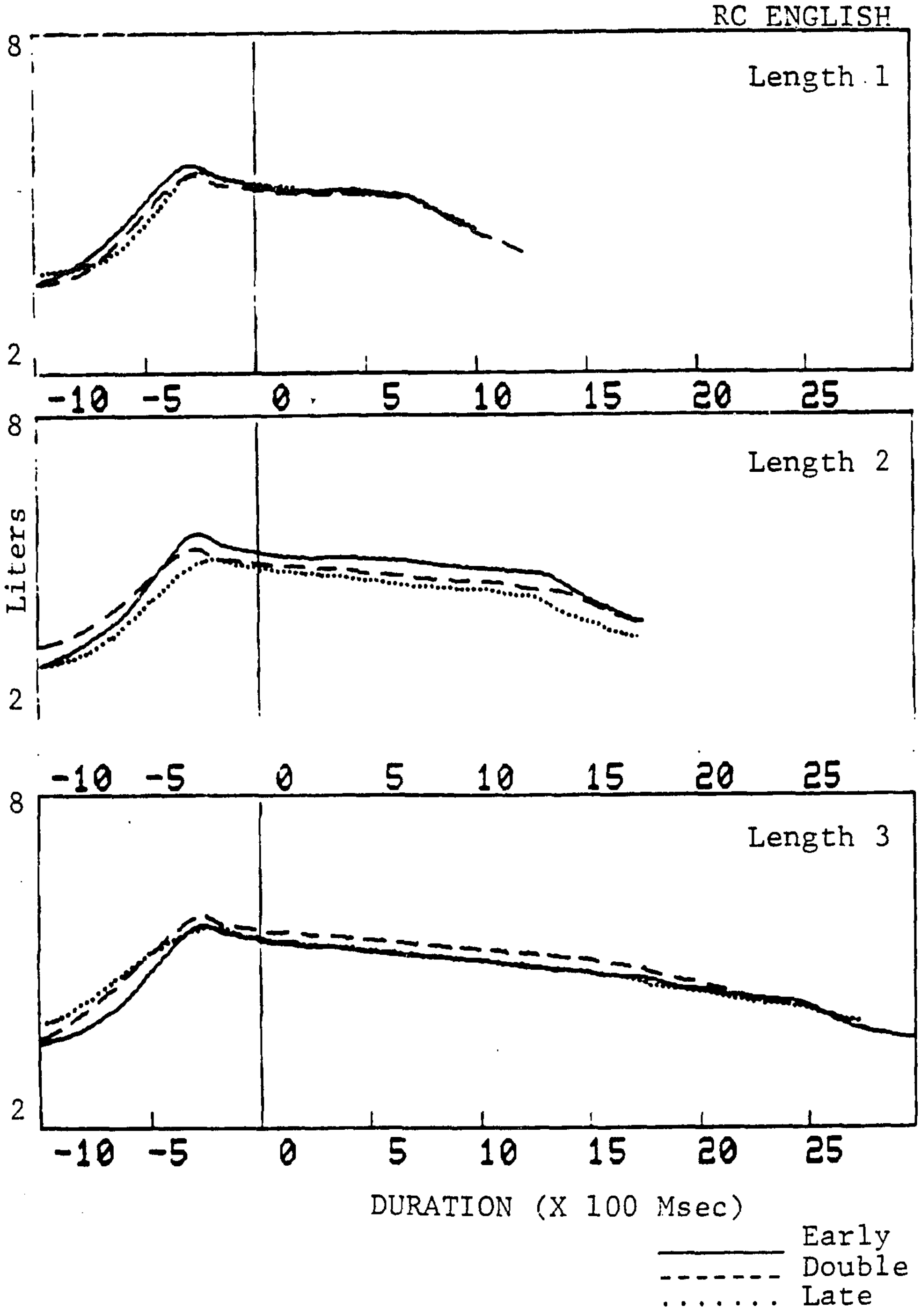
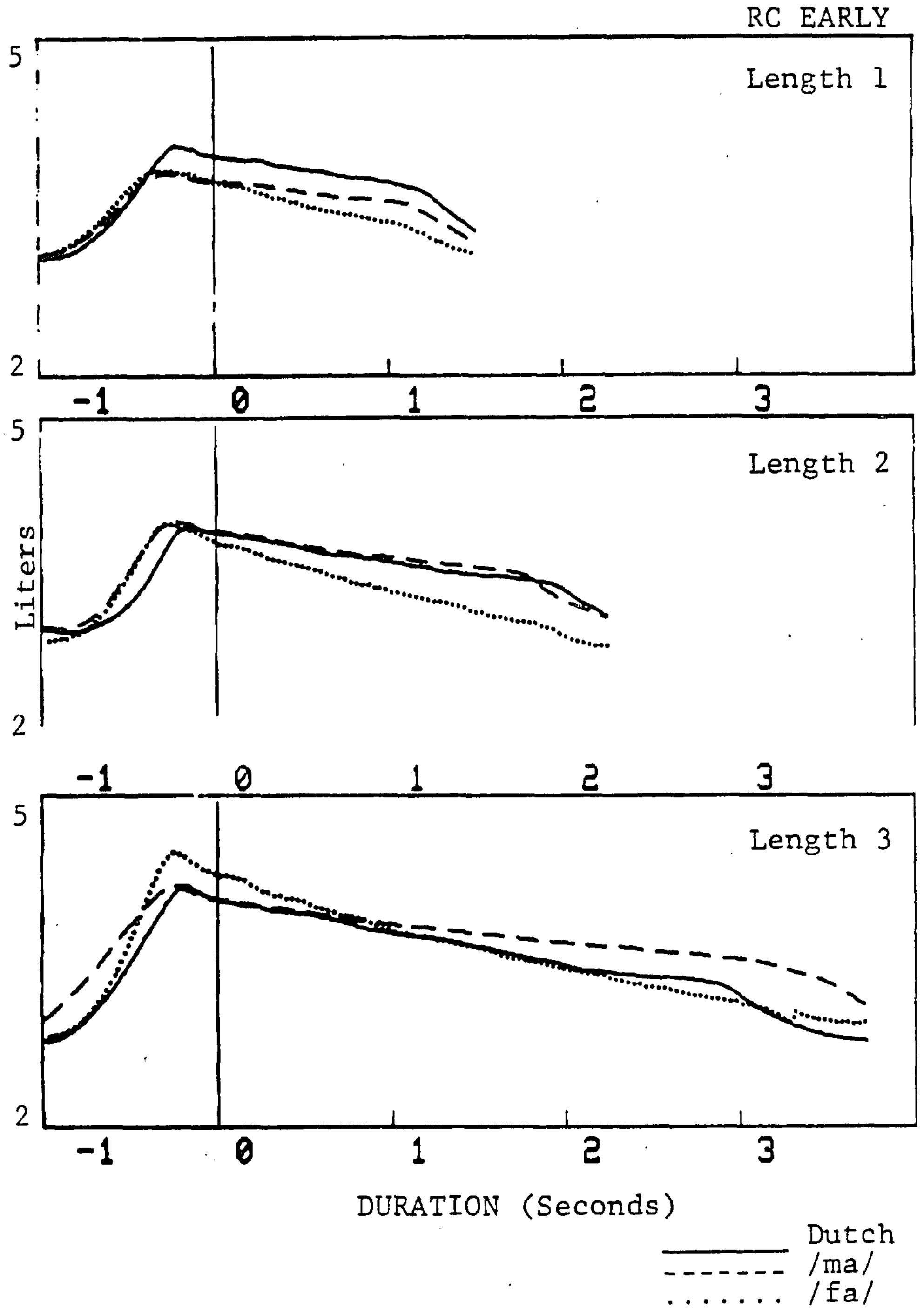


Figure 74. Comparison of lung volume changes across stress configurations for RC English.

Expiration: Respitrace curves for the different phonetic conditions are shown for each length condition in Figures 75-77. These comparisons are shown for the Early, Double and Late stress utterances, respectively. Unlike the length and stress comparisons, it is obvious from these figures that airflow rates (i.e., changes in lung volume over time) differ considerably across phonetic conditions. Airflow rates for the /fa/ utterances always exceed those for the /ma/ utterances, while, for the Dutch, air expenditure is more variable. The airflow rates for the reiterant utterances are shown in Table 14. An analysis of variance reveals a significant interaction of phonetic composition (i.e., /ma/ vs. /fa) with airflow rate ($F=697.91$, $p<.0001$). This is not a surprising result, and can be accounted for by the differences in resistance at the glottis and the configuration of the vocal tract for /ma/ and /fa/ utterances.

Subglottal Pressure: Because airflow rate varies with segmental composition, this allows us some insights into the mechanism for controlling subglottal pressure during sentence production. In other words, if the Ps-generating mechanism is a self-regulated one, there should be no differences observed in the pressure curves for /ma/ and /fa/ utterances. If not self-regulating, however, then Ps would be expected to evidence some differences due to phonetic context. For example, if the mechanism by which Ps is generated is a purely passive one, then pressure should fall in proportion to the rate of airflow.



<u>Figure 75</u>. Comparison of lung volume changes across phonetic conditions for RC Early stress utterances.

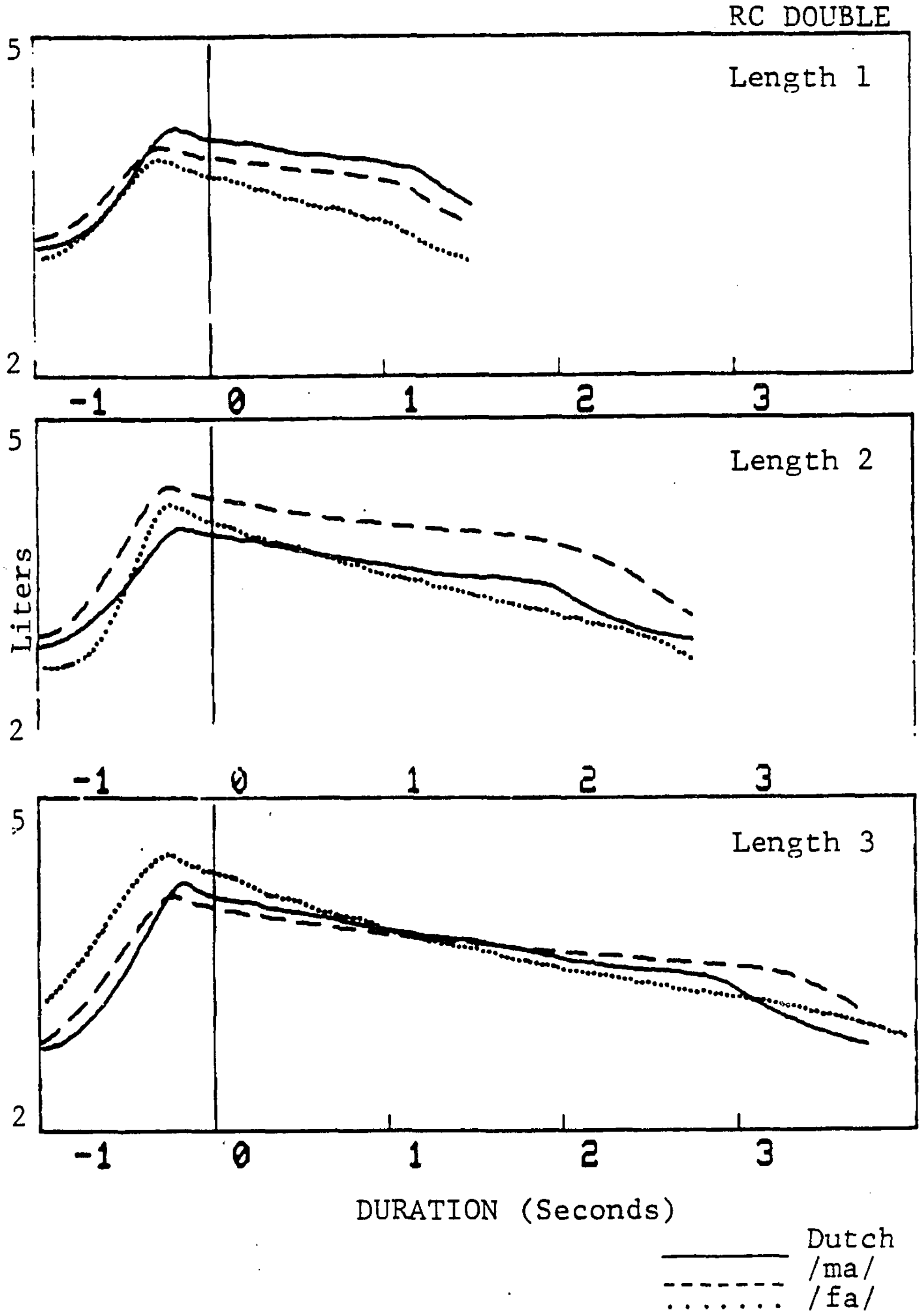


Figure 76. Comparison of lung volume changes across phonetic conditions for RC Double stress utterances.

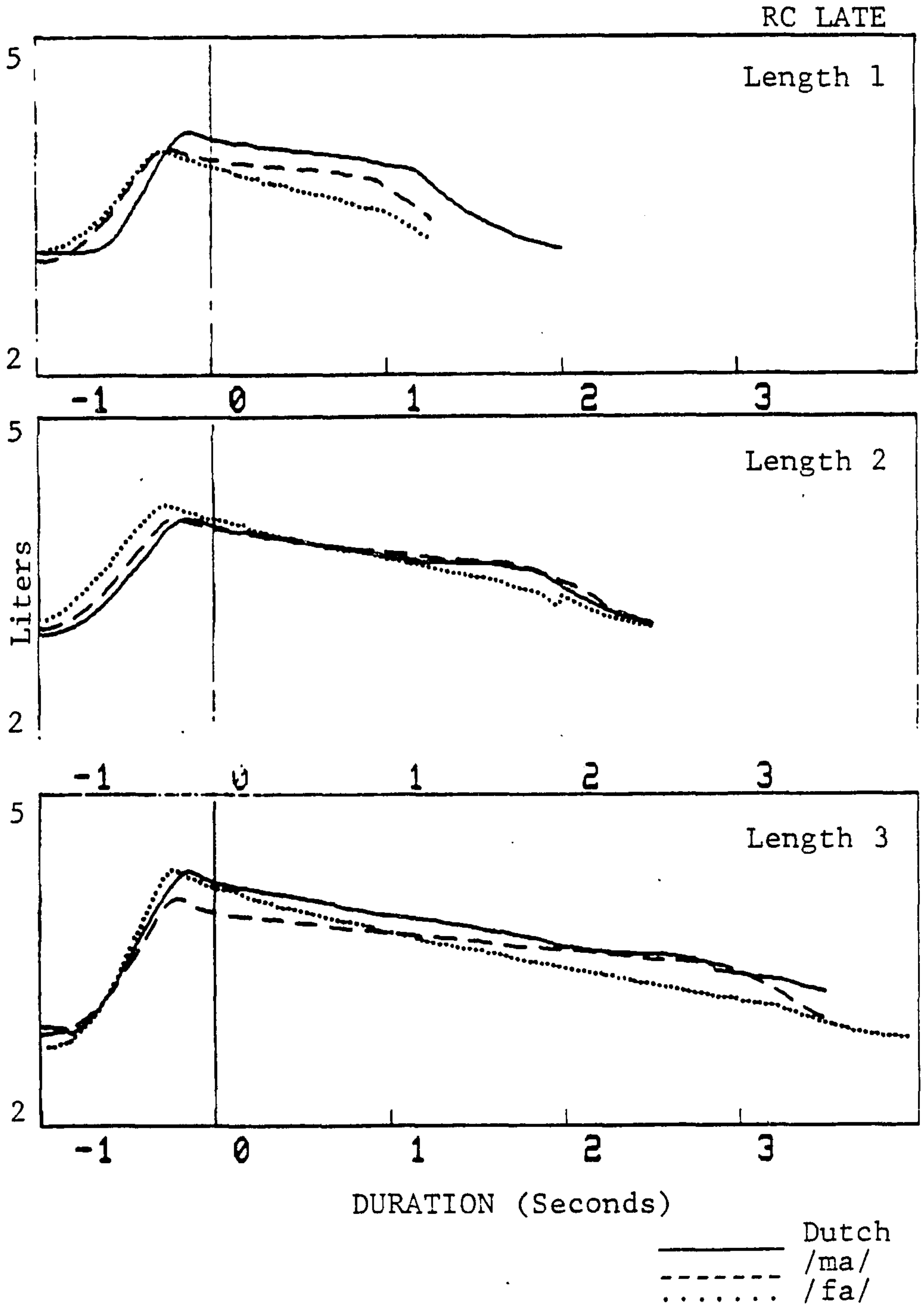


Figure 77. Comparison of lung volume changes across phonetic conditions for RC Late stress utterances.

Table 14

Rate of Expiration (Liters/Second)

	Early	Double	Late	Mean	
Length 1	.227	.287	.277	.266	
Length 2	.257	.238	.231	.242	RC /ma/
Length 3	.208	.208	.212	.209	
Mean	.231	.245	.237		
Length 1	.352	.456	.451	.420	
Length 2	.425	.444	.380	.420	RC /fa/
Length 3	.419	.388	.366	.393	
Mean	.399	.430	.401		

Corresponding subglottal pressure curves are shown in Figures 78-80. From these curves it is apparent that the differences in airflow as a function of phonetic composition are not reflected as corresponding differences in pressure. That is, apart from the presence of segmental effects in the form of a ripple at a roughly syllabic rate, particularly for the Dutch utterances, each curve within a given series of three phonetic conditions can be described by a similar function. Thus, the subglottal pressure cannot be described as a purely passive response to variations in airflow and lung volume.

In order to further quantify these data, the distributions of subglottal pressures and airflow rates for the /ma/ and /fa/ utterances were compared. For subglottal pressure, average levels over a fixed time interval were measured, rather than differences in Ps between two points in time, in order to neutralize any segmental effects. Because the effects of such variables as sentence length and stress configuration may be reflected in initial peak pressure values, these portions of the curves were eliminated from the measured interval. By calculating the averages over an interval of 600 msec, from 400 to 1000 msec after the occurrence of the first stressed syllable (shown as time 0 in Fig. 81), averaging values under these peaks could be avoided. At the same time, data from some of the shortest utterances could be included in the analysis.

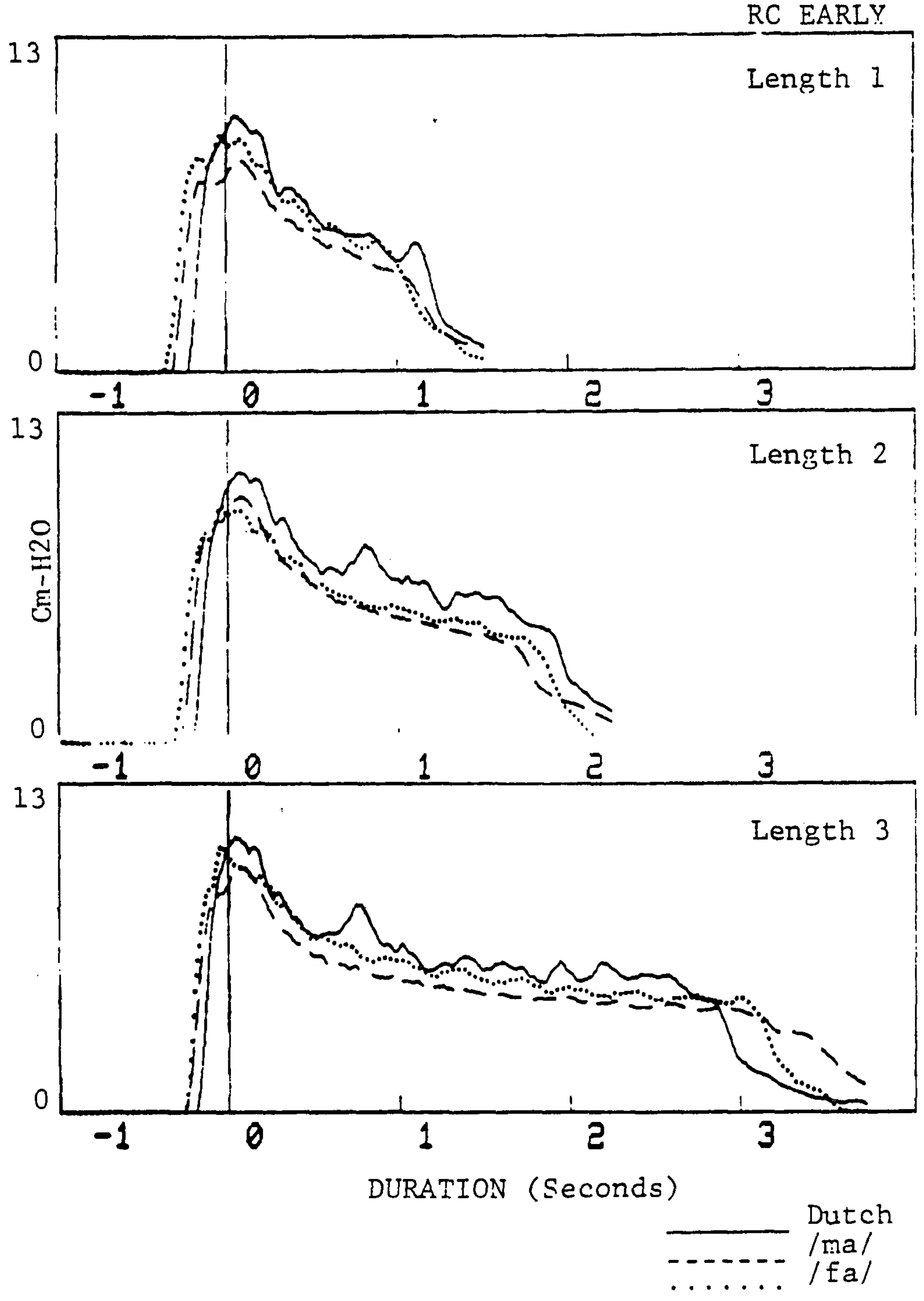


Figure 78. Comparison of Ps contours across phonetic conditions for RC Early stress utterances.

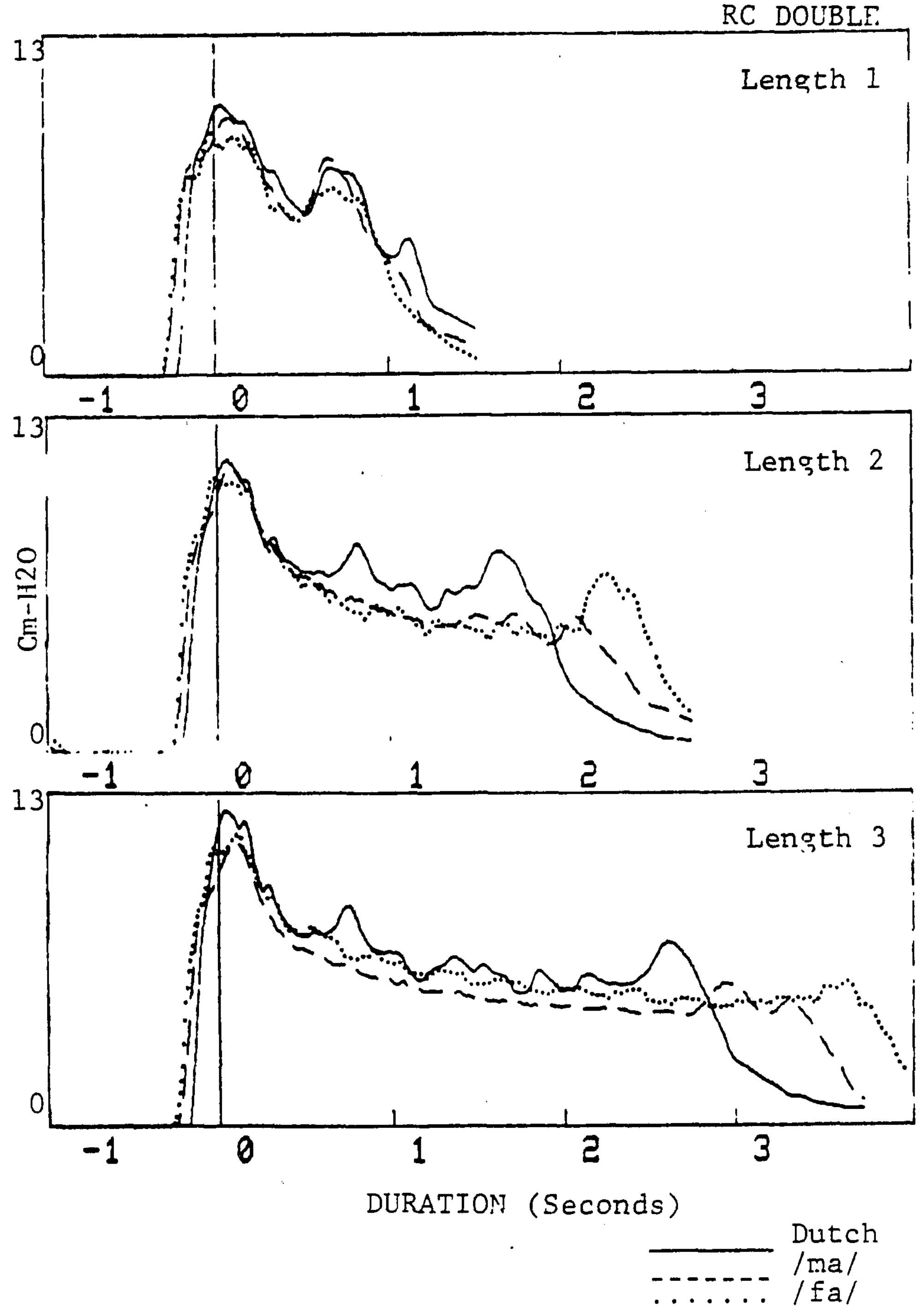


Figure 79. Comparison of Ps contours across phonetic conditions for RC Double stress utterances.

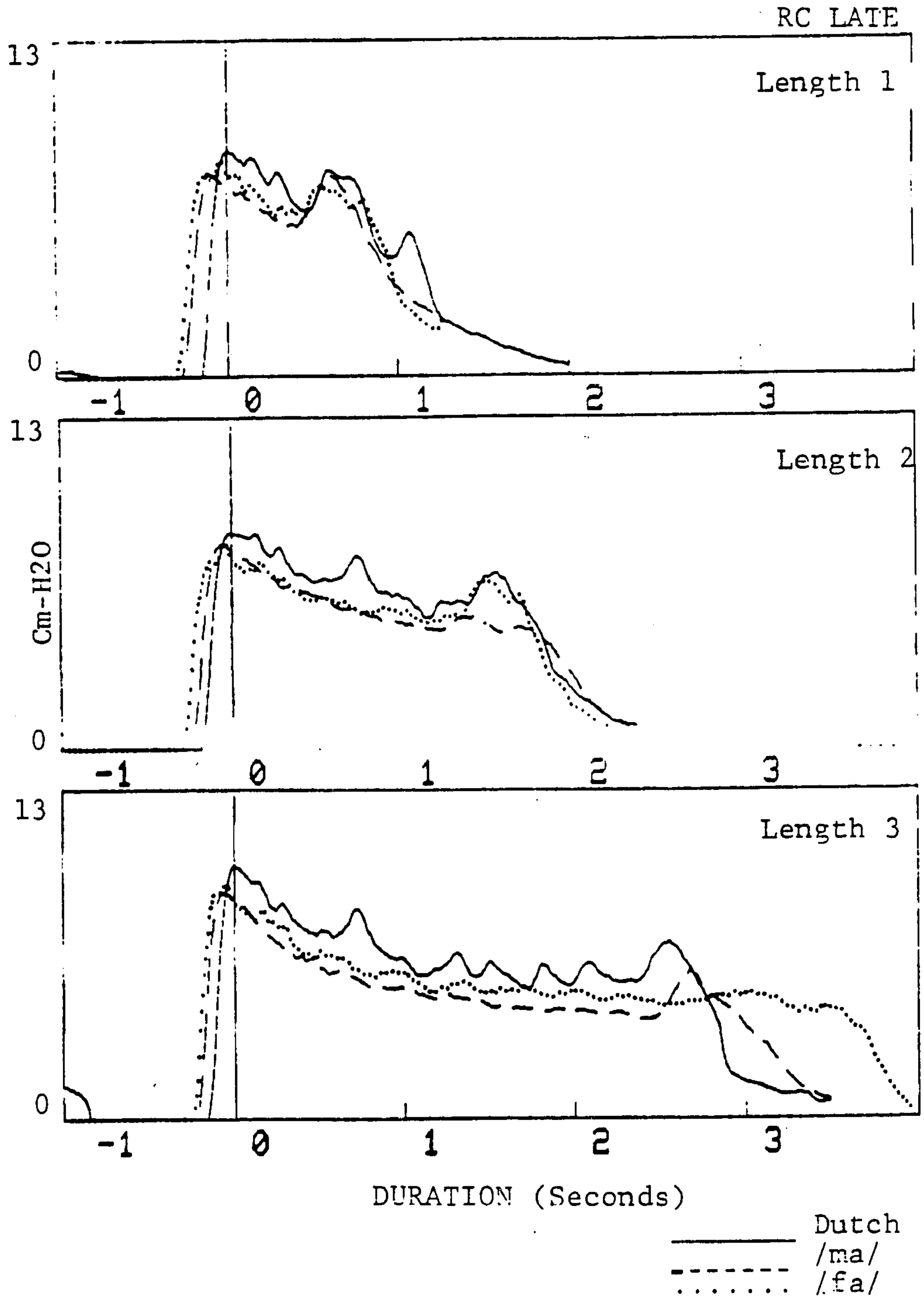


Figure 80. Comparison of Ps contours across phonetic conditions for RC Late stress utterances.

The same interval was used to calculate the change in lung volume over time. However, because the Respitrace curves do not show perturbations due to segmental effects, a simple difference in volume between the two points (i.e., 400 and 1000 msec beyond the first stressed syllable) was calculated in order to derive the rate of decline (i.e., airflow rate). The portions of the subglottal pressure and Respitrace curves over which these calculations were made are shown in Figure 81.

The distributions of Ps and airflow measures for all tokens of the /ma/ and /fa/ utterances are shown in Figure 82. The difference between the means of the distributions is statistically nonsignificant (t=1.09, p>.1). By contrast, the difference between the mean airflow rates for the /ma/ and /fa/ utterances is statistically significant (t=15.91, p<.001). Thus, Ps appears to remain stable despite the significant differences in airflow secondary to the phonetic structure of these utterances.

Cricothyroid Muscle Activity: Representative CT traces for Length 3 Dutch, /ma/ and /fa/ utterances are shown in Figures 83-85 for the Early, Double and Late stress conditions, respectively. Individual waveforms are shown separately in order to make the difference in CT activity across phonetic conditions more apparent. Of particular note is the figure showing cricothyroid activity for the /ma/ utterances (Panel 2) in all stress conditions. There are significant peaks of activity for the initiation of the utterance and for the stressed

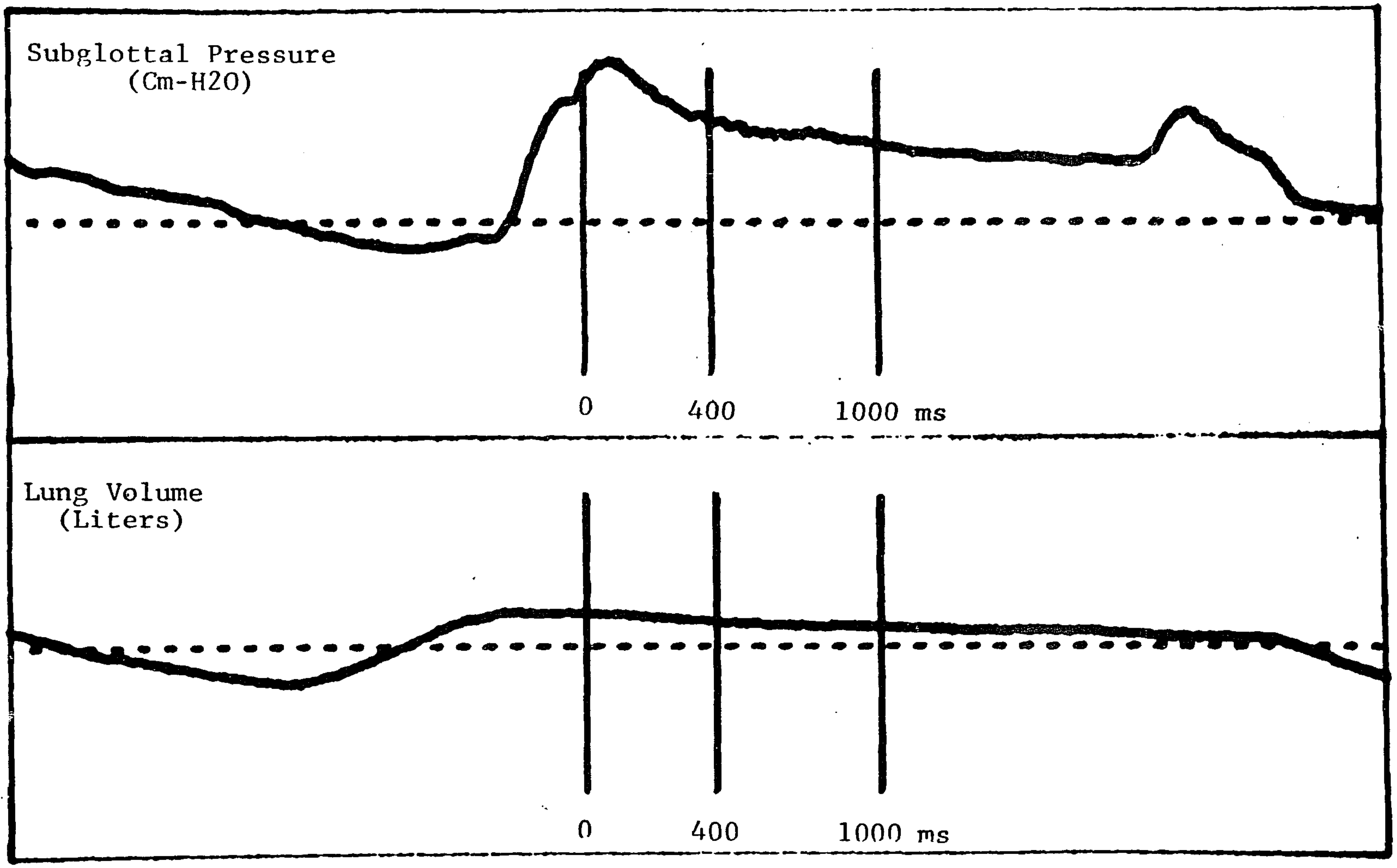


Figure 81. Intervals over which calculations of changes in Ps and airflow were made.

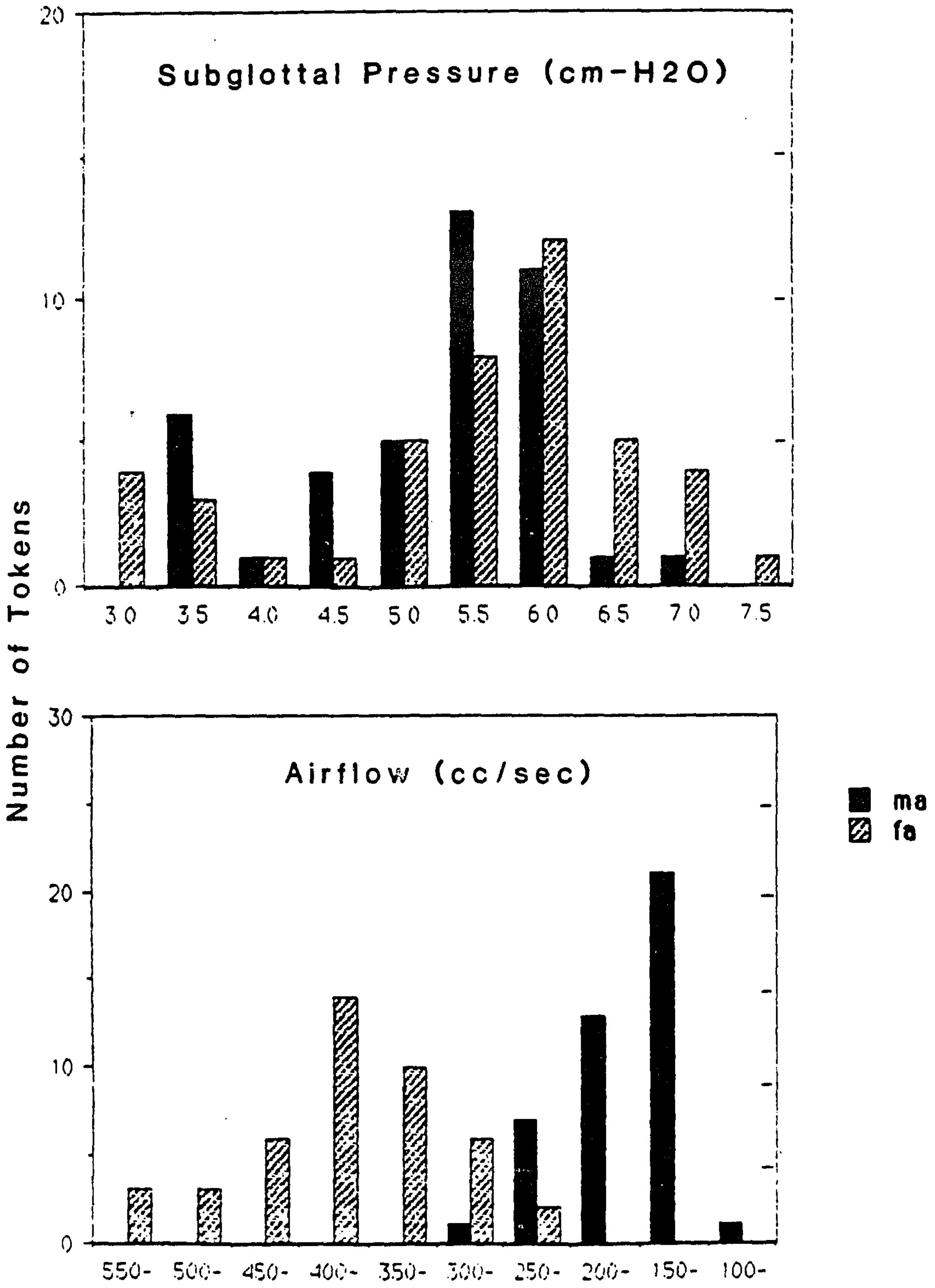


Figure 82. Distributions of Ps and airflow measures.

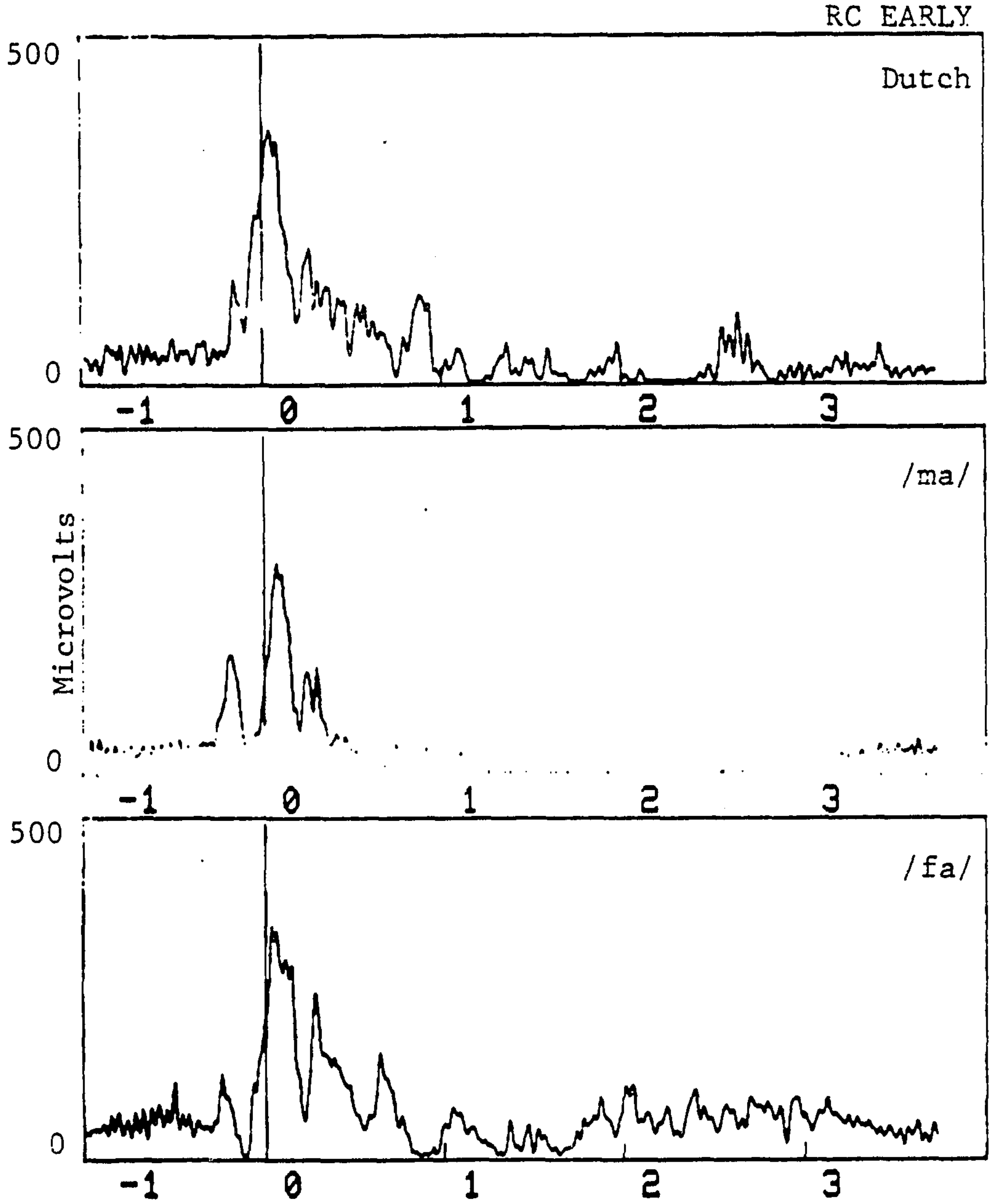


Figure 83. Comparison of CT muscle activity across phonetic conditions for RC Early stress utterances.

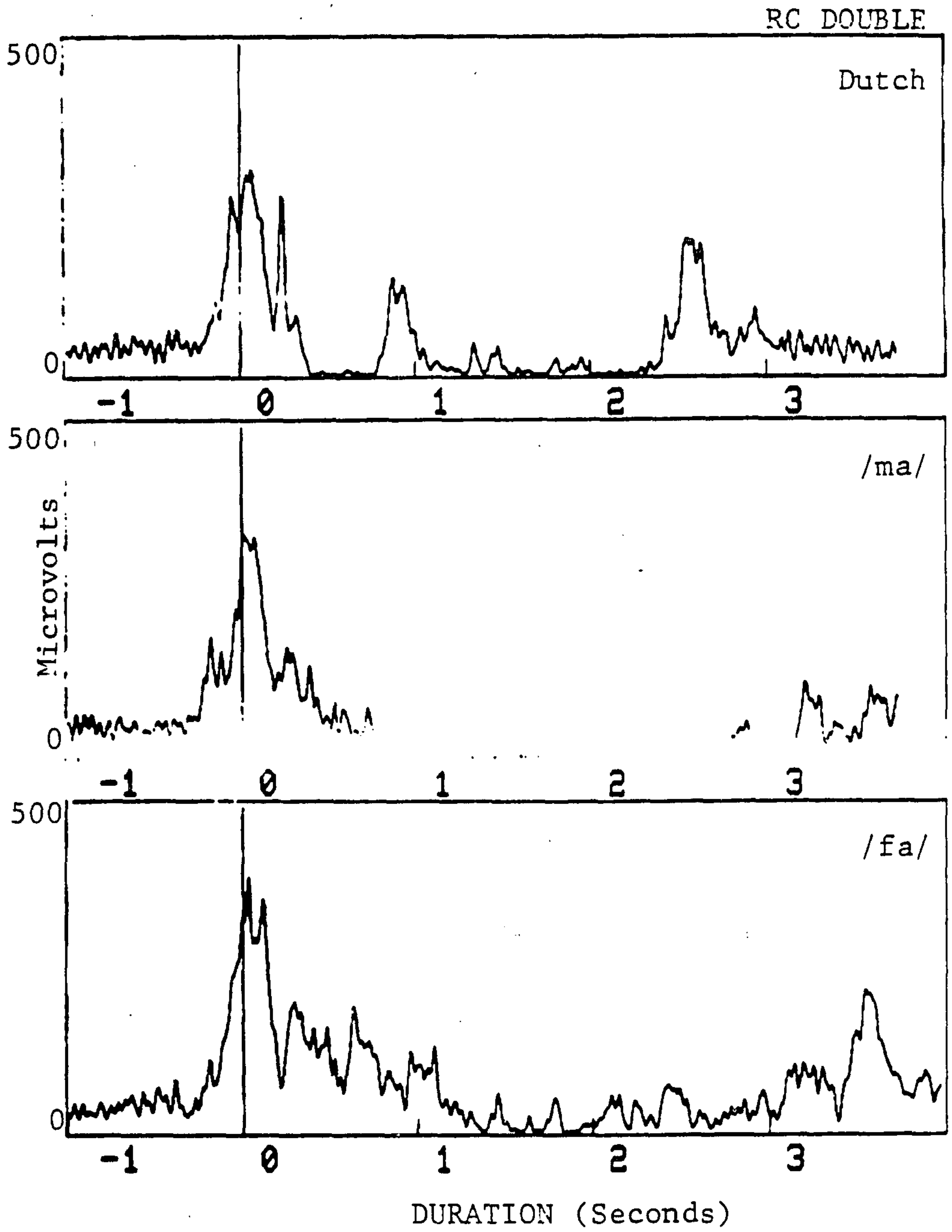


Figure 84. Comparison of CT muscle activity across phonetic conditions for RC Double stress utterances.

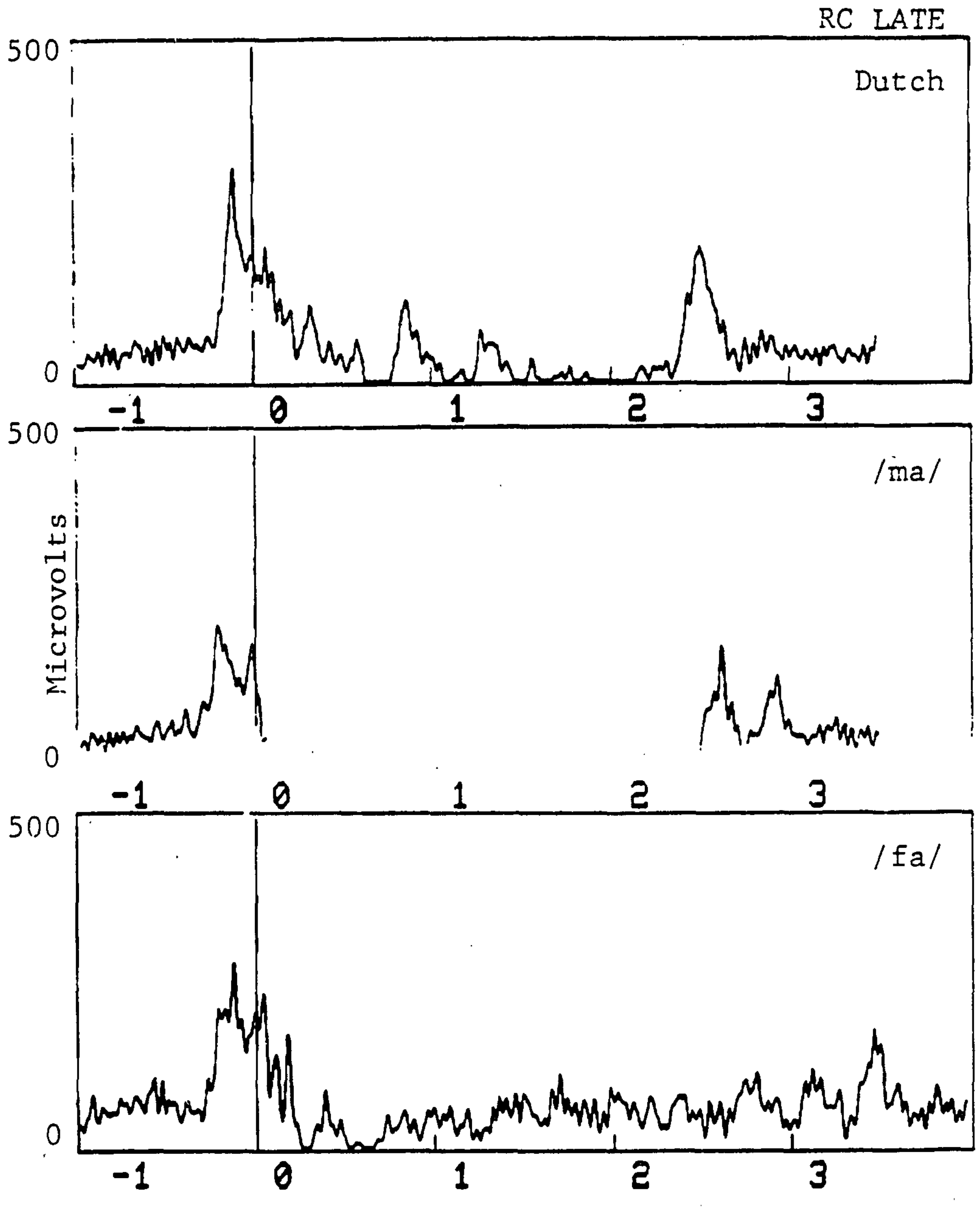


Figure 85. Comparison of CT muscle activity across phonetic conditions for RC Late stress utterances.

syllables, but a virtual suppression of activity during a significant portion of the remainder of the utterance. This drop in activity can be called 'suppression' because the CT activity falls to a level below what can be considered the inherent baseline level, as seen prior to the onset of an utterance. For the Dutch and /fa/ utterances, there is continuous activity beyond the initial peaks, presumably reflecting segmental adjustments not required in the /ma/ utterances, during which voicing is maintained continuously. However, there is the question of whether the differences in the level of cricothyroid activity for the various phonetic compositions are reflected in only local variations in F0 across phonetic conditions, or whether they affect the entire declination contour.

Fundamental Frequency: The corresponding F0 curves are shown in Figure 86. In this figure, the Length 3 F0 curves for the three phonetic conditions are superimposed, with each of the three stress conditions represented in a separate panel. Comparisons of Lengths 1 and 2 are not shown, as it was demonstrated earlier that length effects on the F0 contours are negligible.

This figure shows striking similarities in the shapes of these contours, irrespective of phonetic composition. There are obvious segmental effects, particularly for the /fa/ utterances, where the vocalic portion of each syllable begins at a higher frequency following the voiceless fricative. The overall relationship of these contours, however, is the same as that for the subglottal pressure. Moreover,

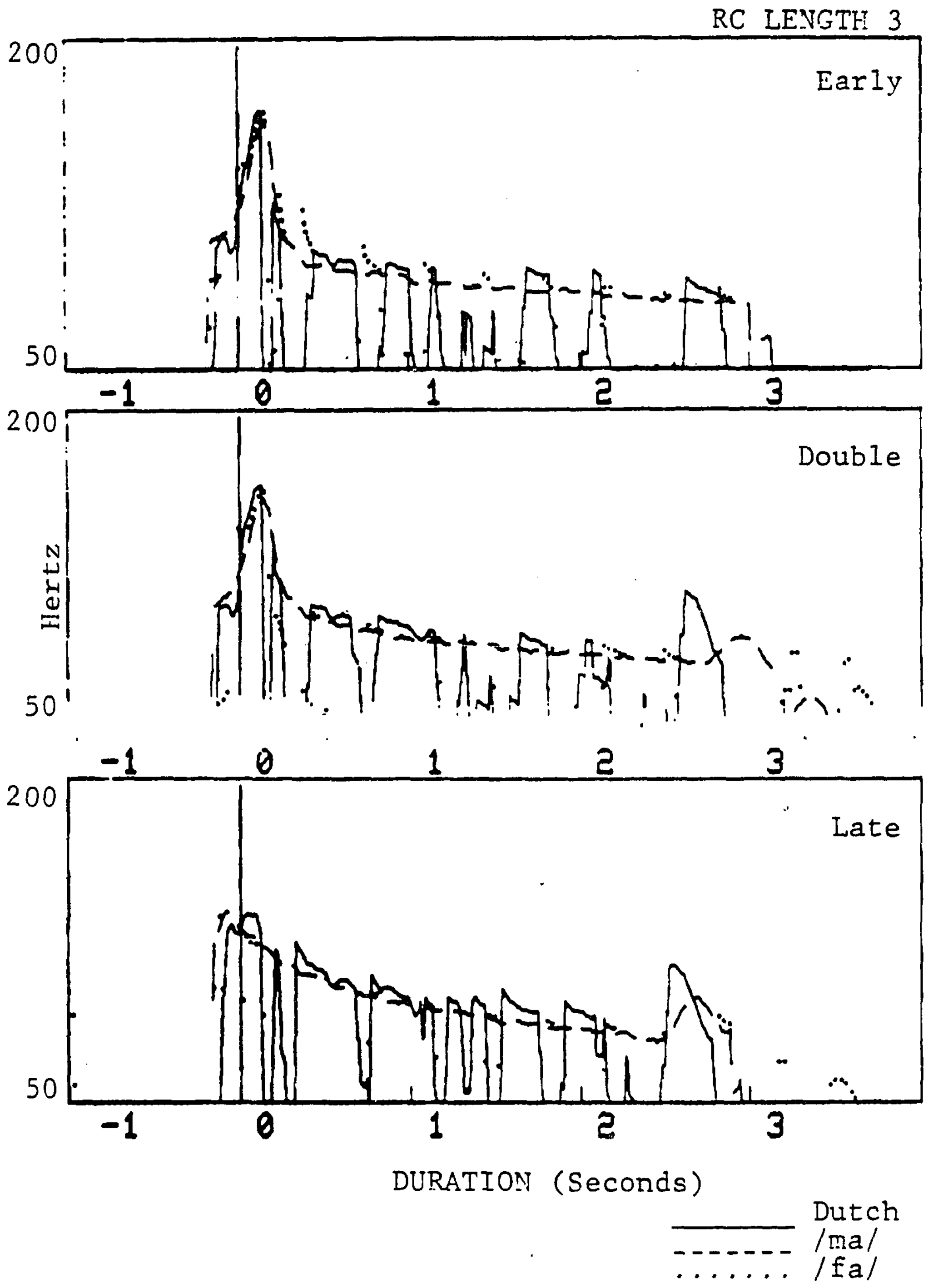


Figure 86. Comparison of F0 curves across phonetic conditions for RC Length 3 utterances.

the differences in the levels of CT activity for the three phonetic conditions are not reflected in the global F0 contours. That is, apart from the initial peaks, the different levels of CT activity can only be related to rapid segmental adjustments, such as devoicing gestures following the vocalic portion of each syllable in the /fa/ utterances.

Strap Muscle Activity: Strap (sternohyoid) muscle activity is displayed for Length 3 utterances across phonetic types for each stress condition in Figures 87-89. It is obvious from these figures that the pattern of activity varies across phonetic conditions. For the /fa/ utterances in particular, there is an almost rhythmic pattern of activity associated with each syllable. However, there is nothing in the activity of this muscle to suggest a role in fundamental frequency declination. That is, if F0 declination were the result of active larynx lowering and/or vocal fold shortening due to the contraction of SH, its activity would be expected to increase over the course of an utterance where F0 was decreasing. However, rather than increasing in magnitude as F0 declines, SH activity remains relatively constant or even decreases over the course of these utterances.

Summary

The effects of segmental composition on the variables under consideration here are in many ways as would be expected. That is, there are significant differences in airflow rates for the /ma/ and /fa/ utterances, as well as variations in the activity of both the

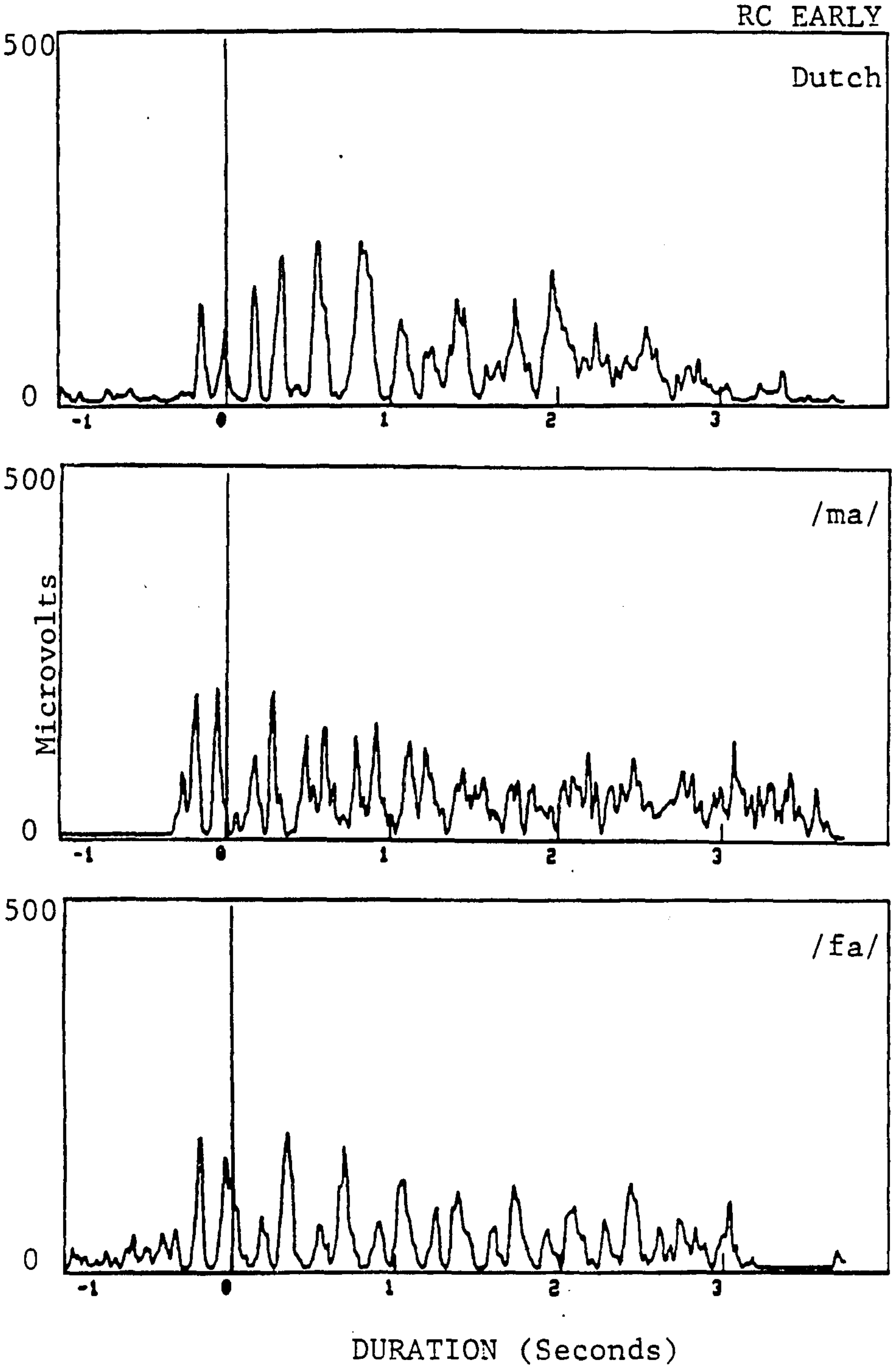


Figure 87. Comparison of strap muscle activity across phonetic conditions for RC Early stress.

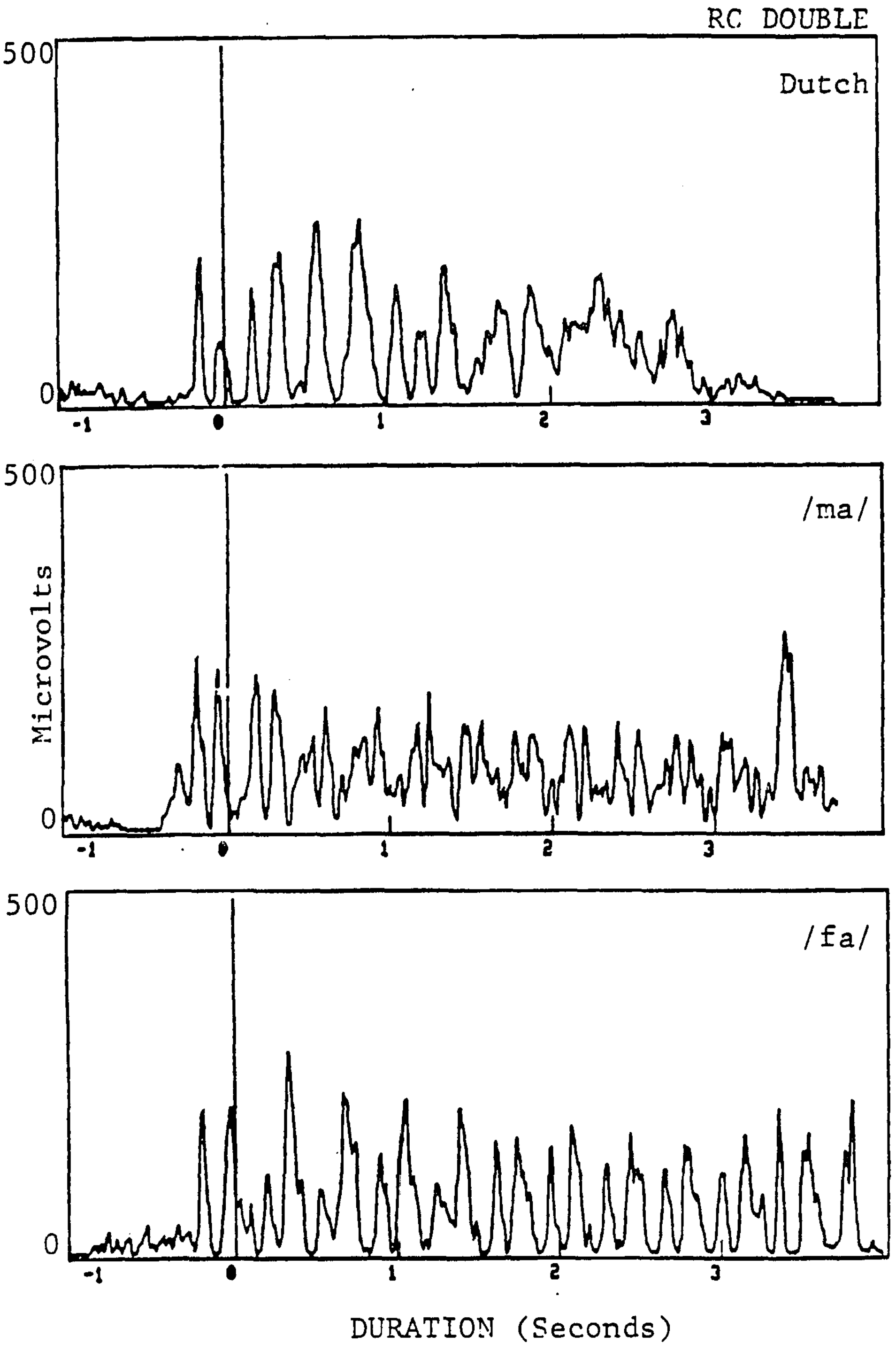


Figure 88. Comparison of strap muscle activity across phonetic conditions for RC Double stress.

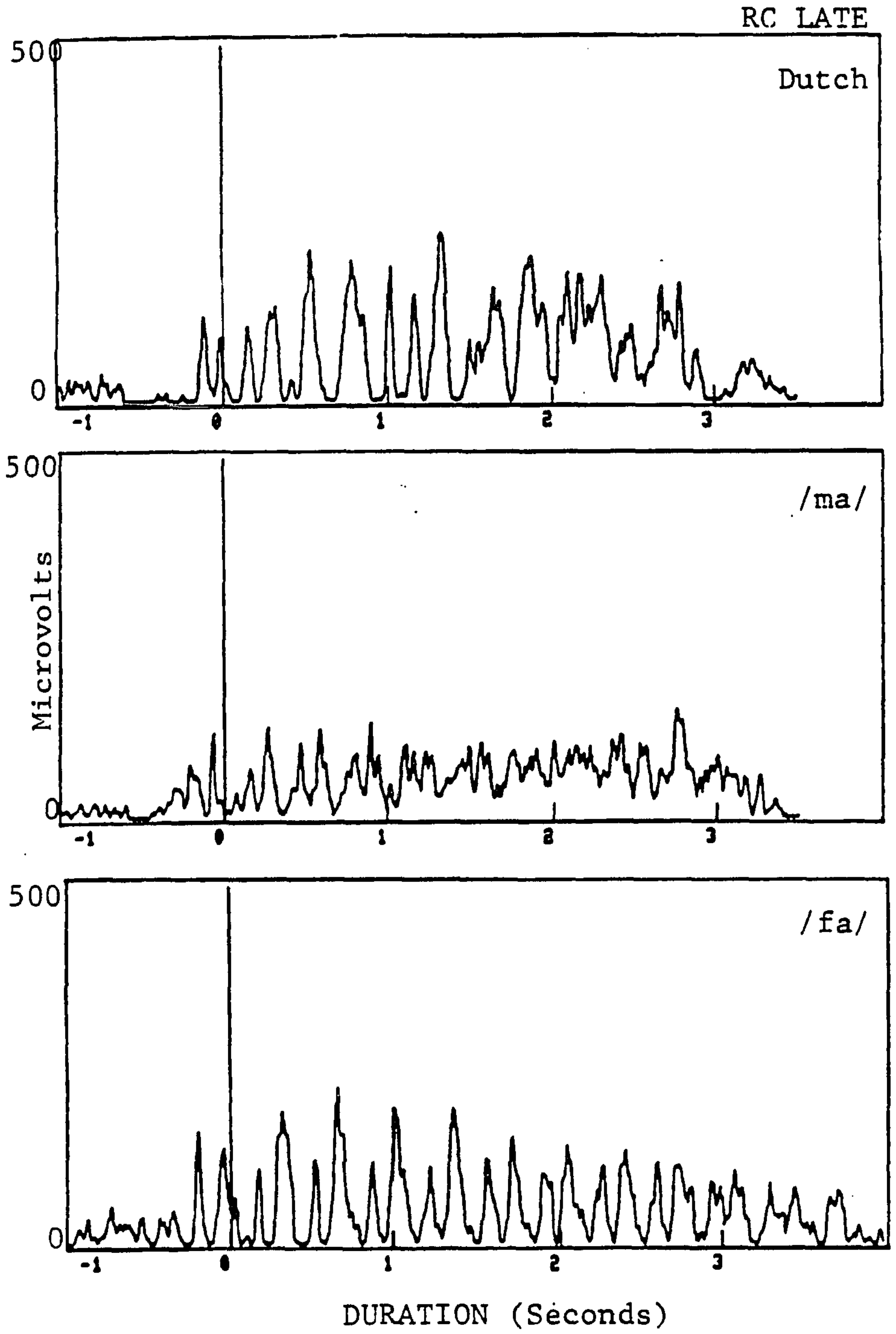


Figure 89. Comparison of strap muscle activity across phonetic conditions for RC Late stress.

cricothyroid and sternohyoid musculature. However, the finding of greatest interest concerns the absence of an effect of these variations on both the Ps and F0 contours. That is, Ps appears to maintain a dynamic stability across phonetic conditions, despite significant accompanying variations in airflow, while, for F0, the effects of variations in the level of cricothyroid activity are restricted to local, momentary variations in F0 at the level of the syllable which do not influence the trajectory of the declination in any significant way. When considered in conjunction with the preceding results on length and stress variations, it begins to appear that 1) there are independent local and global mechanisms involved in the realization of fundamental frequency declination and 2) the subglottal pressure is the mechanism that subserves the gradual decline in F0 over the course of an utterance.

4. Fundamental Frequency Declination and its Relationship to Subglottal Pressure

Because cricothyroid and sternohyoid muscle activity do not appear to be involved in fundamental frequency declination (i.e., that part of the decline occurring after the initial drop and, in the cases of Late and Double stress, before the final rapid increase in F0), the F0 and Ps curves were compared directly in order to determine whether the falling subglottal pressure alone could account for the decline in fundamental frequency.

For the initial comparisons, the data of both RC and EB will be considered. Figures 90-98 show superimposed F0 and Ps curves for each length and stress condition for RC, RC English and EB, respectively. For RC, only the /ma/ utterances are shown since the effect of phonetic composition on both the F0 and Ps contours was shown to be minimal.

From these figures it can be seen that, where there are abrupt and substantial increases in Ps, F0 shows abrupt and substantial increases as well. At the same time, however, the same relative increases occur for CT activity, samples of which are shown, together with F0 and Ps curves, for RC /ma/ Late stress utterances in Figure 99. However, it is not the behavior of peaks that is of interest here, as they have already been discussed. Rather, of interest is the behavior of the fundamental frequency declination in relation to the falling subglottal pressure beyond these peaks. And, for these portions of these utterances, F0 does in fact closely mirror the decline in Ps, not just in direction, but in relative rate of decline as well. This is particularly apparent for RC's /ma/ utterances.

In order to quantify further the relationship between F0 and Ps, and to establish the frequency-to-pressure ratio over this decline, it was first necessary to establish an appropriate method by which slope could be calculated for each. There were two technical matters to consider in deriving these measures. First, because the declination is considered here to be the portion of F0 decline that excludes initial peaks, it was necessary to eliminate any portion of the declination that could be considered the result of the relaxation of the CT muscle

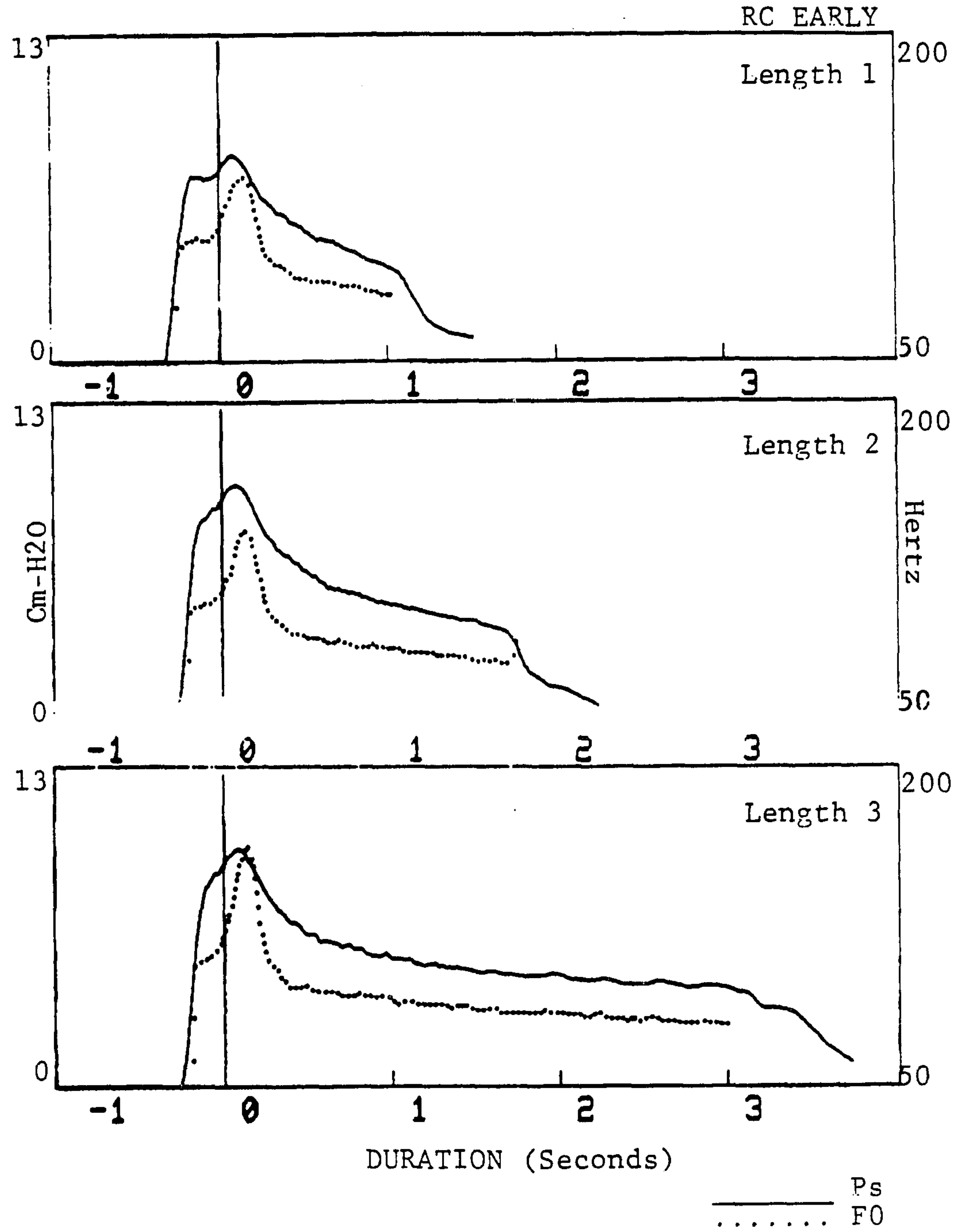


Figure 90. Comparison of Ps and F0 contours across RC Early stress utterances.

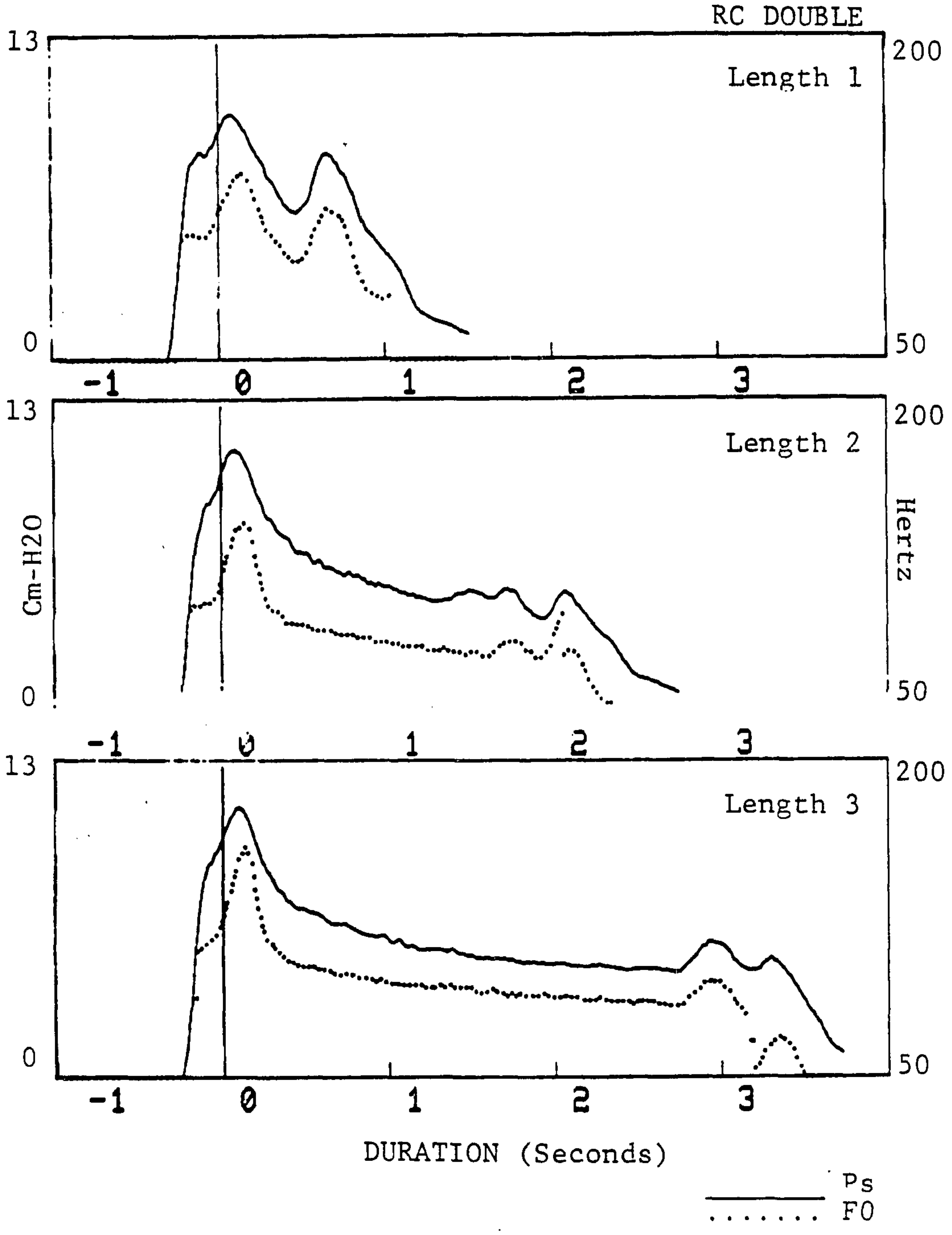


Figure 91. Comparison of Ps and F0 contours for RC Double stress utterances.

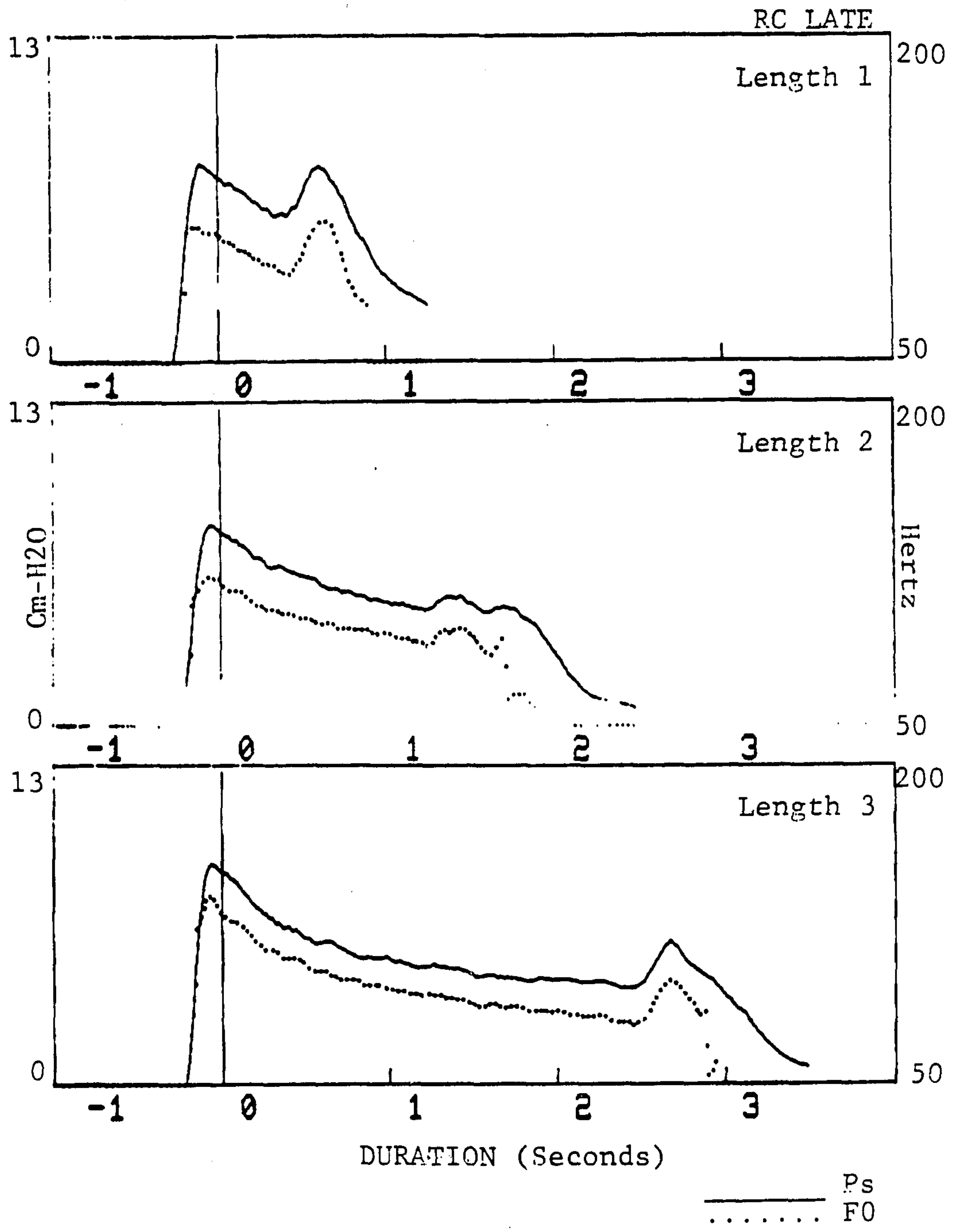


Figure 92. Comparison of Ps and F0 contours for RC Late stress utterances.

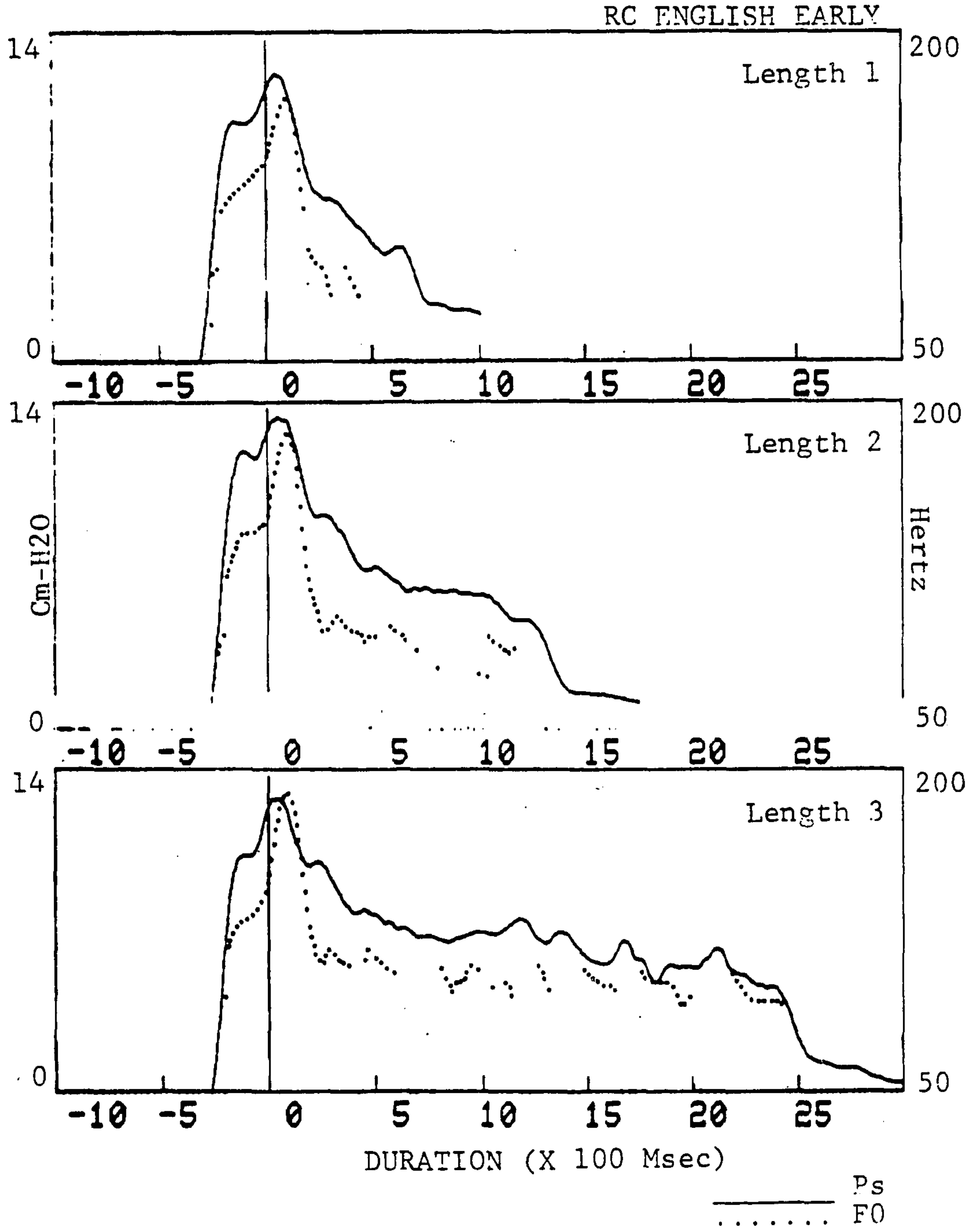


Figure 93. Comparison of Ps and F0 contours for RC English Early stress utterances.

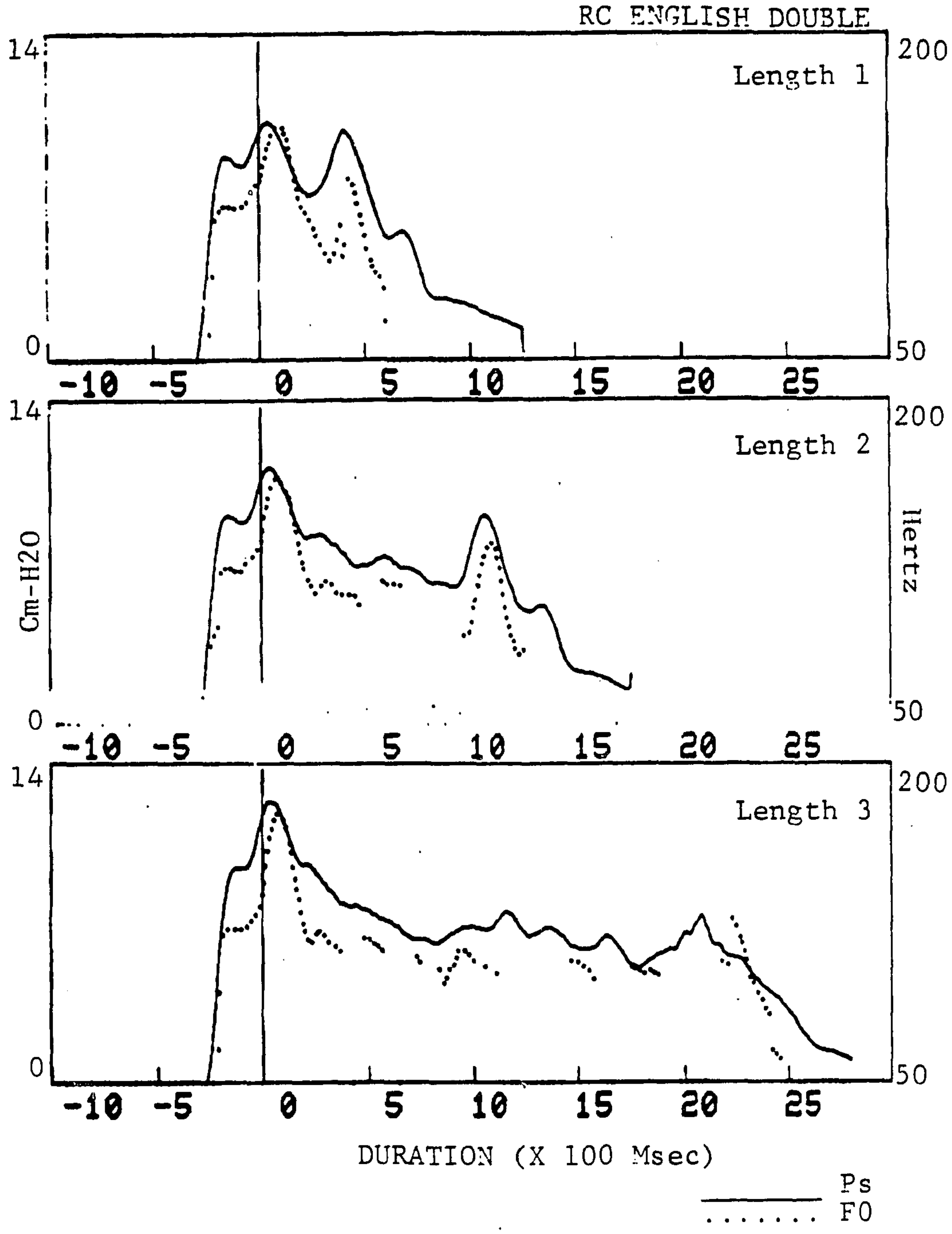


Figure 94. Comparison of Ps and F0 contours for RC English Double stress utterances.

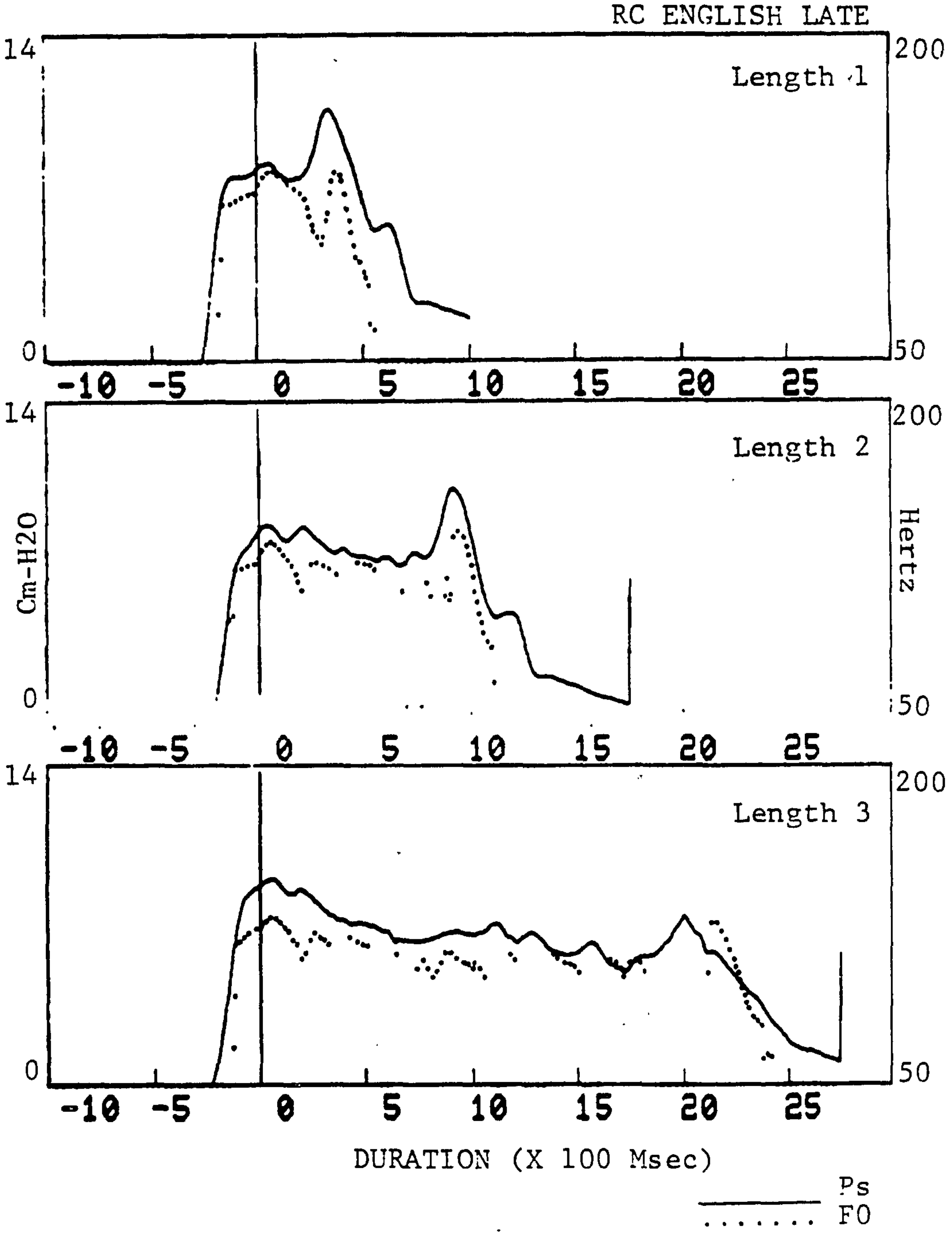


Figure 95. Comparison of Ps and F0 contours for RC English Late stress utterances.

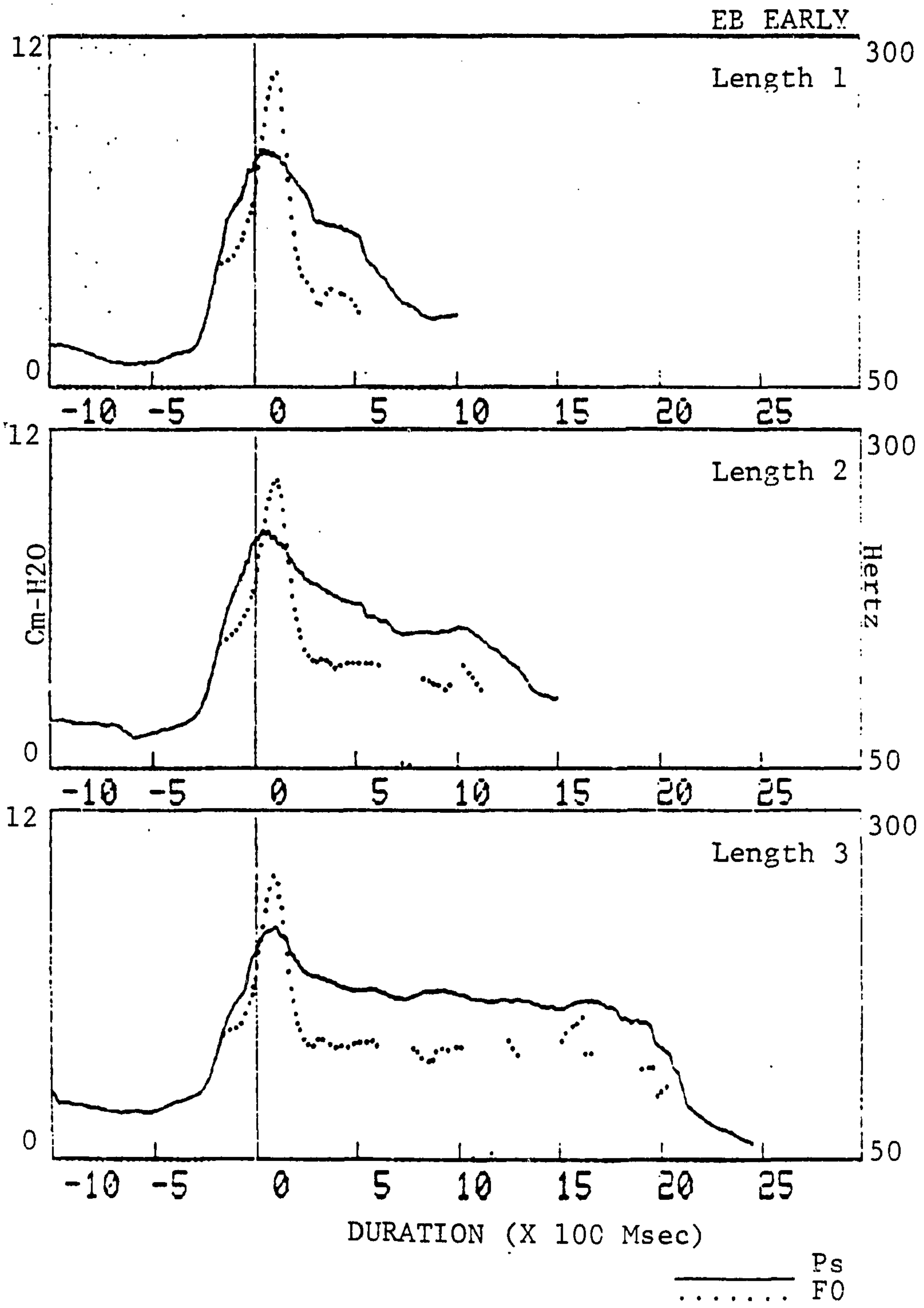


Figure 96. Comparison of Ps and F0 contours for EB Early stress utterances.

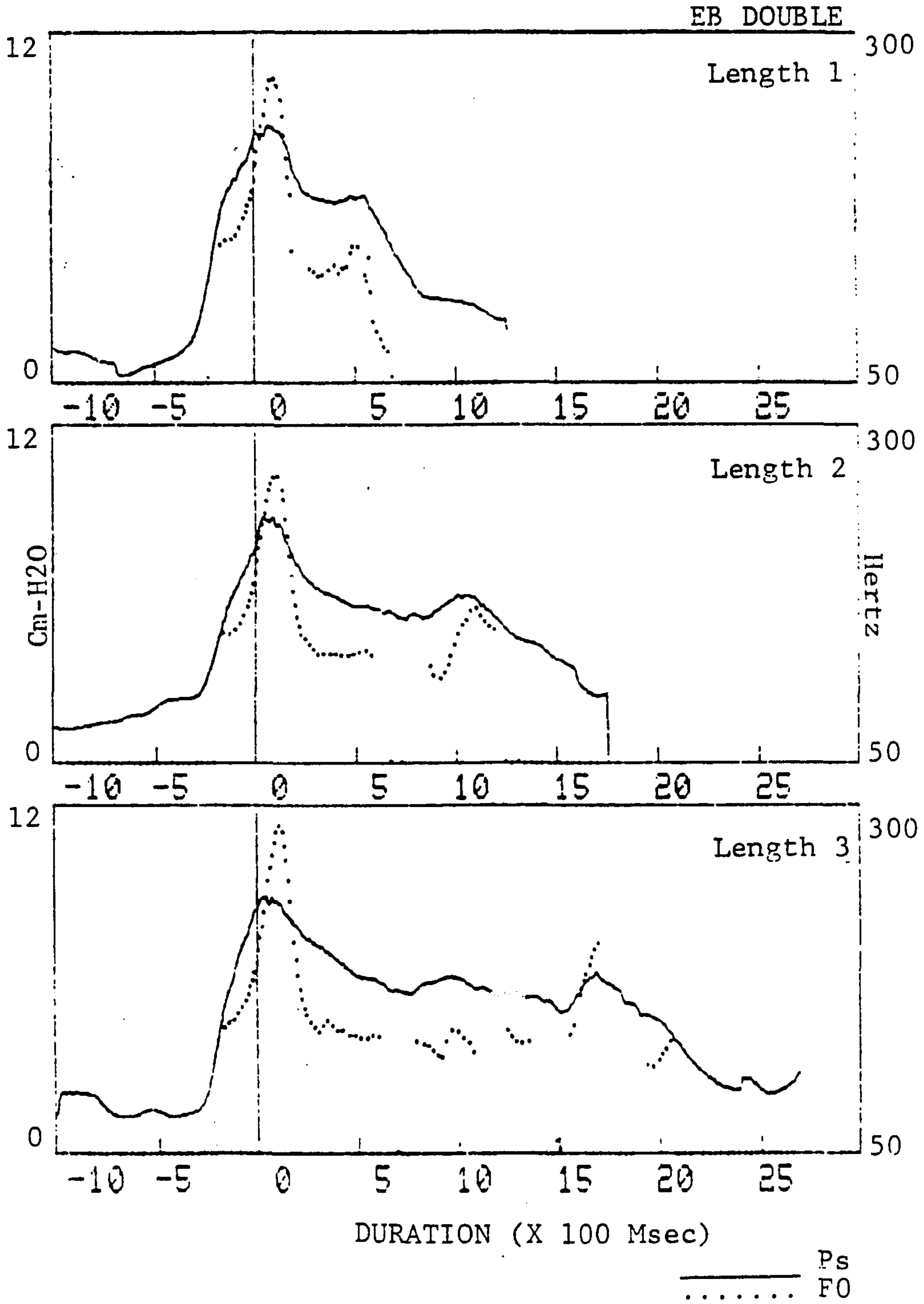


Figure 97. Comparison of Ps and F0 contours for EB Double stress utterances.

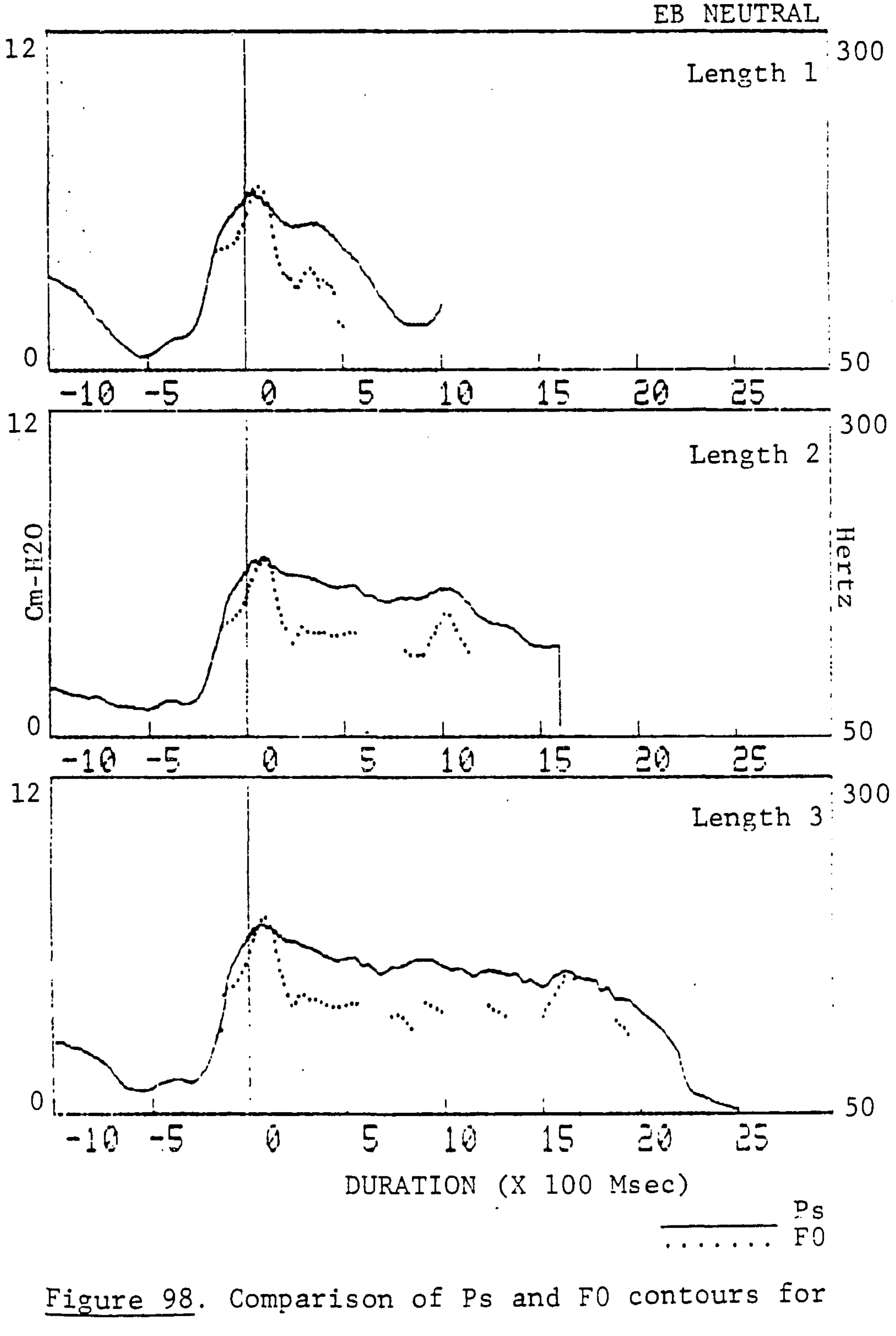


Figure 98. Comparison of Ps and F0 contours for EB Neutral stress utterances.

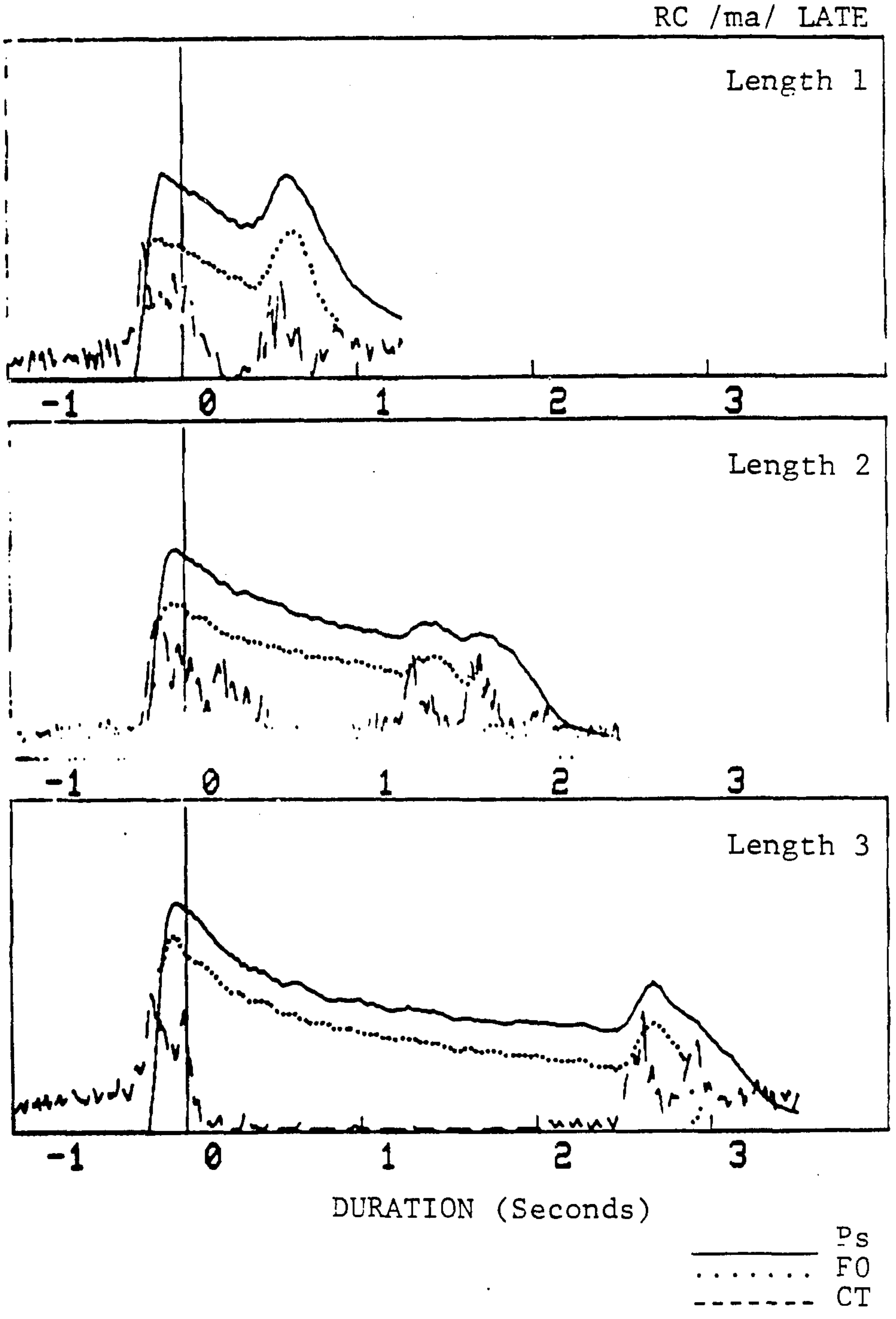


Figure 99. F0 declination in relation to falling Ps and CT muscle activity.

following its significant initial activity. The time at which the offset of CT activity occurred was thus used as a 'starting point' for calculating slope because it could be reasonably well assumed that the significant decreases in F0 that result from the relaxation of this muscle would be complete by this point in time. Offset time was defined as the time at which the EMG output (in microvolts) dropped to and remained below a level equivalent to the baseline plus 10 percent of the peak level. It was determined on a token-by-token basis in order to accommodate variability in the timing of CT activity.

The second consideration was finding a technique that would best capture the nature of the decline in both F0 and Ps. In other words, if their decline were strictly linear, then it would not matter over what portion of an utterance the slope was calculated (excluding, of course, initial and/or final peaks). On the other hand, if their decline followed a curve other than a linear one, then calculating the slope over different portions of the curves would yield different results. If the decline were exponential, for example, the slope would be greater for the earlier portions of utterances as compared to the average slope over the utterance as a whole.

Using RC's /ma/ data, two different analyses were thus performed. In the first, the amount of decline in both F0 and Ps was calculated for each token, either between the point at which CT activity ceased, following the initial peak, and the end of the utterance (for the Early stress condition), or between CT cessation and the minimum value just preceding the last peak (for the Double and Late stress conditions).

In the second analysis, the average duration of the Length 1 Early stress utterances from initial peak to endpoint (i.e., 1080 ms) was used as a durational reference, and the amount and rate of F0 and Ps decline for all utterances was calculated between the offset of CT activity and the average endpoint of the reference utterances. Thus, while the amount and rate of declination were determined over variable intervals in the first analysis, a relatively constant interval occurring over the same portion of each curve, regardless of absolute length, was used for the second analysis. Figure 100 is a schematic representation of the portions of the F0 and Ps curves over which these analyses were performed. The arrows denote the offset of cricothyroid activity, and the numbers 1 and 2 denote the respective endpoints for the first and second analyses.

These analyses could not be performed on Length 1 utterances of the Double and Late stress conditions. In the former condition, the interval between CT offset for the first peak and CT onset for the second peak was too short. In the latter condition, CT activity was never consistently suppressed, so that an offset time could not be obtained. In fact, it can be seen from Figure 100 that, even for the Length 2 and 3 Late stress utterances, the offset of CT activity occurs considerably later than that for the Early and Double stress utterances.

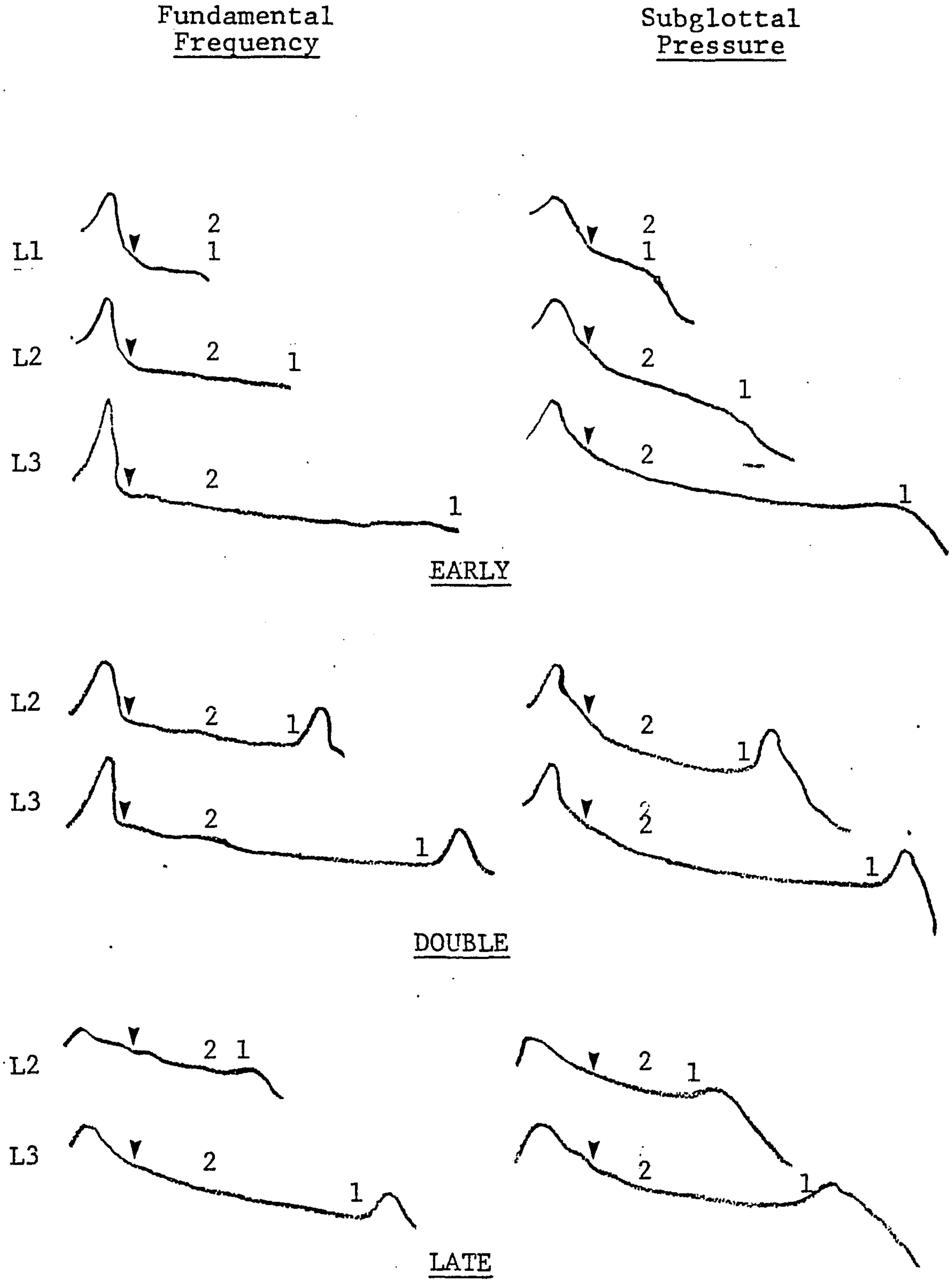


Figure 100. Schematic representations of Analyses 1 and 2 for corresponding F0 and Ps curves.

Table 15 shows the results of these analyses in terms of F0 slope (Hz/sec), Ps slope (cm-H2O/sec) and the frequency-to-pressure ratio (Hz/cm-H2O). Looking first at the effect of utterance length on slope for both F0 and Ps, the two analyses yield rather different results. That is, according to Analysis 1, there is a substantial decrease in the rate of change with increasing length. However, when slope is calculated over relatively fixed portions of these same utterances, as in Analysis 2, this length effect is substantially reduced, demonstrating a more constant rate of decline across lengths. (For Length 3 of the Late stress condition, there is some peculiarity in the F0 data that yields an abnormally high rate of decline.)

Regarding the frequency-to-pressure ratios, six of the seven values from Analysis 2 fall within the accepted range of 2-7 Hz/cm-H2O (the exception being, of course, the Length 3 Late utterance), while only four of the seven values from Analysis 1 fall within the accepted range. However, even those values that fall outside the range are considerably lower than those reported elsewhere (see, for example, Maeda, 1976).

The results of these analyses thus suggest that a strict linear model does not best characterize the declination of either F0 or Ps, as the slope calculated over an entire utterance is not equivalent to the slope calculated over some portion of it. Rather, both F0 and Ps appear to be decreasing nonlinearly, as is evident from the larger slopes for the earlier portions of utterances. Thus, treating F0 and Ps as linear functions would appear to obscure the nature of their

Table 15

Rate of Decline for Fundamental Frequency (Hz/Sec) and Subglottal Pressure (Cm-H2O/Sec) and Their Ratios (Hz/Cm-H2O)

	Analysis 1			Analysis 2		
	FO Rate	Ps Rate	FO/Ps Ratio	FO Rate	Ps Rate	FO/Ps Ratio
Early:						
Length 1	22.21	3.94	5.64	22.21	3.94	5.64
Length 2	14.39	2.47	5.83	19.70	3.73	5.28
Length 3	7.03	1.07	6.57	17.42	3.35	5.20
Double:						
Length 1	-	-	-	-	-	-
Length 2	15.37	2.38	6.46	22.52	4.12	5.47
Length 3	10.76	1.36	7.91	19.75	3.49	5.66
Late:						
Length 1	-	-	-	-	-	-
Length 2	20.79	2.57	8.09	17.32	2.61	6.64
Length 3	16.85	1.56	10.80	35.22	3.52	10.01

decline and to promote the impression of a length effect, whereby slope decreases inversely with utterance length, where perhaps none exists. Treating F0 and Ps as nonlinear functions, on the other hand, serves to reduce significantly any apparent length effect and demonstrates that the decline in Ps can indeed account for the decline in F0.

EXPERIMENT 2: F0 RESETTING AND ITS PHYSIOLOGICAL CORRELATES

This section reports the results of analyses performed on two-clause English utterances for five subjects. Where they apply, the analyses in this section will focus on the following: 1) the effect of syntactic structure on resetting and other acoustic variables, 2) the organization (or reorganization) of various acoustic aspects of the entire sentence as a function of varying the length of either clause, 3) the difference between clauses produced as part of two-clause utterances and the same clauses produced as isolated sentences, and 4) the resetting of F0 following a major syntactic boundary as a function of respiratory activity at that boundary.

1. The Effect of Syntactic Structure on the F0 Contour

If syntactic structure is a factor in the mental representation of an utterance, then one might expect it to be reflected in various aspects of the F0 contour. Of particular interest are those aspects which occur around the clause boundary, such as F0 resetting. Both RC and EB produced lexically matched pairs of sentences, one of which was composed of subordinate-main clauses and the other composed of conjoined main clauses (Appendix A). For these comparisons, RC produced one sentence of each syntactic type. EB produced a comparable pair of utterances which were further modified to create additional sentences by adding lexical items to create a total of three length variants for each clause. Each of the first clause variants was combined with each of the second clause variants to form nine pairs of utterances. Thus, possible first clauses (C1) were:

L1) When the lawyer called,/The lawyer called,

L2) When the lawyer called Reynolds,/The lawyer called Reynolds,

L3) When the lawyer called Reynolds on Monday,/The lawyer called Reynolds on Monday,

and possible second clauses (C2) were:

L1) the plans were discussed.

L2) the plans were discussed at some length.

L3) the plans were discussed at some length on the telephone.

For the conjoined main clause sentences, the conjunction 'and' connected the two clauses. Note that RC's original sentences exist here as a subset of EB's corpus of sentences.

For these analyses, F0 values were measured for the first stressed peak in the first clause (Peak 1A='lawyer'), the last stressed peak in the first clause (Peak 1N='called', 'Reynolds' or 'Monday'), and the first stressed peak in the second clause (Peak 2A='plans') for all tokens. Note that, because of the way EB's first clauses were constructed, Peak 1N values were not always derived from the same word.

Additional measures of resetting describe the relationship of clause-initial peaks (Reset A = Peak 1A-Peak 2A), and the relationship of peak F0 at the onset of a second clause to peak F0 at the offset of a first clause (Reset N = Peak 2A-Peak 1N). Thus, a value of 0 or less for Reset A indicates the presence of 'complete resetting' (i.e., Peak 2A $\geq$ Peak 1A), while a value of 0 or less for Reset N indicates the presence of at least 'partial resetting' (i.e., Peak 2A $\geq$ Peak 1N).

The averaged values for the key F0 variables are shown for both subjects for each syntactic condition in Table 16. For RC, the values for each syntactic condition represent one utterance collapsed across respiratory conditions (i.e., the presence or absence of an inspiration at the clause boundary). For EB, the values for each syntactic condition represent values collapsed across nine sentences where the length of either clause is varied.

For both subjects, averaged F0 values for the sentences composed of conjoined main clauses are somewhat higher than for those composed of subordinate-main clauses. For RC, an analysis of variance reveals a significant interaction of Peak 1A with syntactic condition ($F=33.95$, $p<.0001$), but the absence of any interaction for the other variables (Table 17). For EB, no dependent variable shows a significant interaction with syntactic condition (Table 17). Thus, because

Table 16

Syntax Comparisons

RC:

	N	1A	SD	1N	SD	2A	SD	ResA	SD	ResN	SD
SubMain	9	146	9.0	125	6.2	121	4.5	25	12.9	-4	10.4
Conjoin	10	152	6.3	128	3.9	125	7.0	27	3.0	-4	7.2

EB:

	N	1A	SD	1N	SD	2A	SD	ResA	SD	ResN	SD
SubMain	47	183	11.8	147	8.1	149	7.8	33	3.5	+2	7.9
Conjoin	42	184	14.1	150	8.2	150	6.5	34	11.7	0	6.5

Table 17

Interaction of FO Peaks with Syntactic Condition

	RC		EB	
	F	p	F	p
1A	33.95	<.0001	.47	>.1
1N	3.11	>.05	2.16	>.1
2A	4.37	>.05	.50	>.1
Reset A	3.86	>.05	.13	>.1
Reset N	.05	>.1	1.42	>.1

syntactic condition appears to have little influence on the F0 measures of interest here, the data for these two subjects were collapsed across syntactic condition for the remaining analyses.

2. The Effect of Varying Clause Length

The purpose of the analyses in this section is to determine the effect on all F0 variables of manipulating the length of either or both clauses. In particular, these data should reveal whether there is a reorganization of F0 as a function of length and provide insights into the question of whether it is the clause or the sentence that is the unit over which speech is organized.

It will be recalled that EB produced length variations for both Clause 1 and Clause 2, the durations of which are shown in Appendix M. These are coded, using the number 1, 2 or 3, to denote the length variant of each clause. For example, '3-1' denotes a Length 3 first clause and a Length 1 second clause utterance.

Subjects BK, JS and KM produced length variations only for second clauses, three of which were the same as those produced by EB - namely, subordinate-main clause sentences corresponding to EB's Length 2 first clauses (i.e., When the lawyer called Reynolds,) followed by second clauses of three lengths. In addition, three new sentences of the same syntactic structure and the comparable number of syllables (i.e., constant length first clauses with second clauses of three lengths) were produced by these speakers, for a total of six two-clause utterances (Appendix A). Averaged clause durations and their respective number of syllables are shown for these sentences in Appendix N. For the two sentence types produced by BK, JS and KM, first clauses contain the identical number of syllables, while the second clauses are one syllable longer for Lengths 2 and 3 of the second sentence type. For all subjects, there is a significant interaction of sentence type with the durations of both Clause 1 and Clause 2 (Table 18).

Clause 1

Table 19 shows the F0 values for the key variables for the sentences composed of each of the three length first clauses with one of the three length second clauses for EB. These values are represented schematically in Figure 101a, b and c. It can be seen from this figure that, for each Clause 2 length (i.e., a, b or c), the effect of Clause 1 length on the indicated peak values is very similar. Consequently, the data were collapsed further. The F0 values are shown in Table 4b and are represented schematically in Figure 101d. Each length schematic is labelled 1-N, 2-N and 3-N, respectively, to show the effect of varying the length of Clause 1 across all Clause 2 lengths. The corresponding collapsed duration measures are shown in Table 20.

Table 18

Interaction of Sentence Type with Clause Duration

	Clause 1		Clause 2	
	F	*p*	*F*	*p*
BK	5.52	<.05	7.57	<.01
JS	12.37	<.001	16.18	<.001
KM	113.90	<.0001	12.57	<.001

Table 19

Length Comparisons for FO Reset Data

	N	1A	SD	1N	SD	2A	SD	ResA	SD	ResN	SD
1-1	9	170	14.4	144	9.7	138	5.2	31	11.6	-5	8.1
2-1	12	190	11.3	151	7.2	151	7.1	39	10.6	0	8.5
3-1	9	186	9.8	147	6.3	150	6.2	36	7.2	+3	7.3
1-2	10	173	9.5	144	9.3	146	6.5	27	6.6	+2	7.1
2-2	9	198	13.3	156	7.3	157	5.8	41	13.9	0	6.8
3-2	10	187	6.7	147	8.3	152	3.5	36	7.5	+4	7.7
1-3	10	174	6.4	148	7.8	148	3.6	26	6.4	0	5.1
2-3	10	191	3.8	151	5.0	156	3.7	35	4.6	+5	5.2
3-3	10	180	8.8	149	7.7	151	5.2	29	6.3	+2	5.9

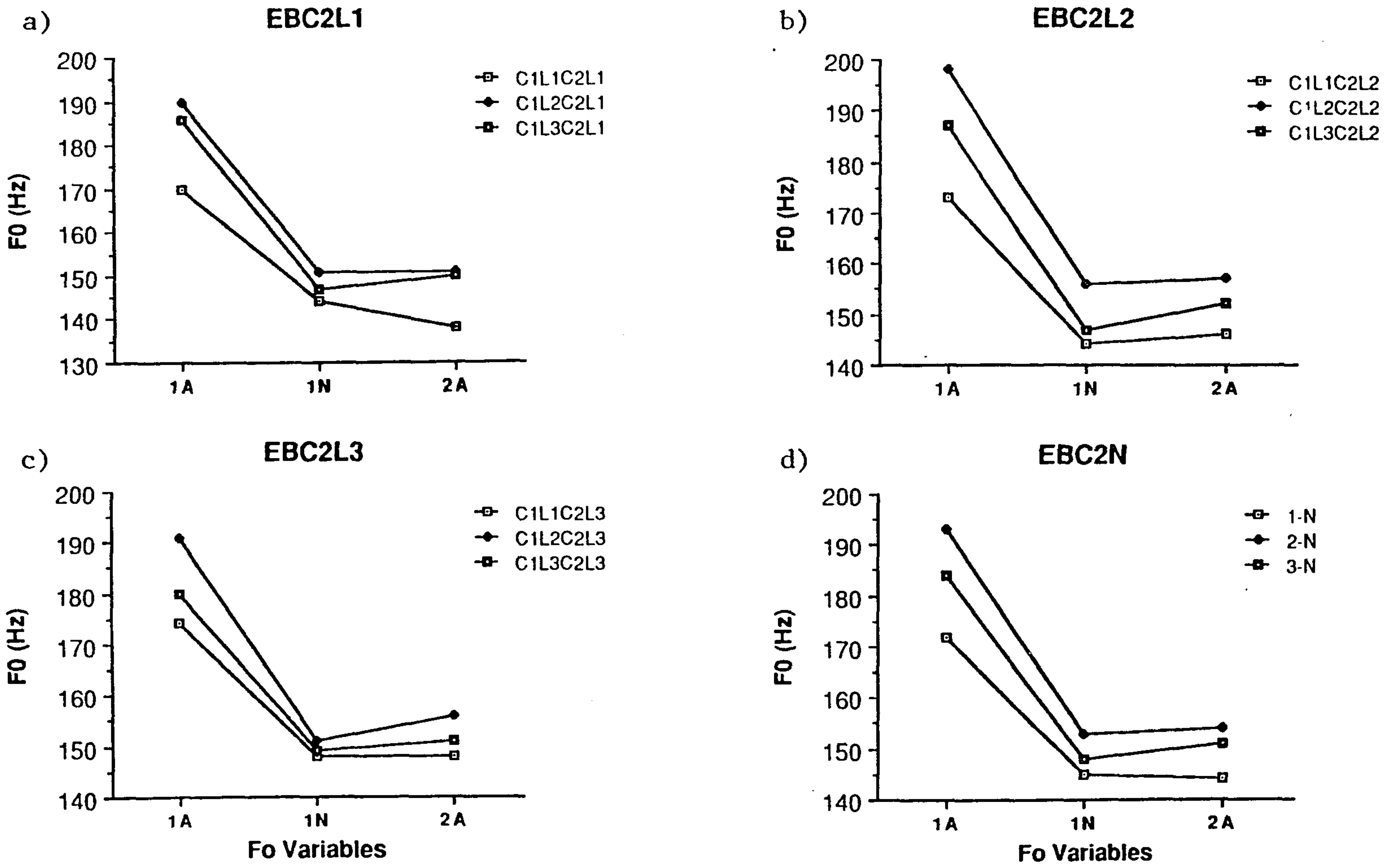


Figure 101. Key F0 variables for Clause 1 length comparisons for EB.

Table 20

FO Reset Data as a Function of Clause Length

Clause 1 Length Variations

	N	*1A*	*SD*	*1N*	*SD*	*2A*	*SD*	*ResA*	*SD*	*ResN*	*SD*
1-N	29	172	10.2	145	8.8	144	6.5	28	8.4	-1	7.3
2-N	31	193	10.5	153	6.8	154	6.2	38	10.2	+2	7.3
3-N	29	184	8.8	148	7.3	151	4.9	34	7.4	+3	6.8

Clause 2 Length Variations

	N	*1A*	*SD*	*1N*	*SD*	*2A*	*SD*	*ResA*	*SD*	*ResB*	*SD*
N-1	30	183	14.5	148	8.2	147	8.4	36	10.2	-1	8.5
N-2	29	186	14.0	149	9.5	151	6.7	34	11.1	+2	7.1
N-3	30	182	9.8	149	6.9	151	5.5	30	6.7	+2	5.7

Peaks 1A, 1N and 2A all show significant interactions with Clause 1 length (1A: F=31.67, p<.0001; 1N: F=6.94, p<.01; 2A: F=22.18, p<.0001). However, these F0 values are consistently out of order with respect to absolute Clause 1 length (Table 19). For Peak 1A, while values for the shortest clauses are lower than both Length 2 and length 3 peaks, those for Length 2 clauses are higher than Length 3 peaks. The same result is found for Peak 2A values. Thus, it is clear that the peak values are not adjusted solely on the basis of increasing (or decreasing) absolute utterance length.

The F0 values for Peak 1N may be related to different intonational patterns at the end of clauses of varying lengths and the fact that F0 is derived from different lexical items (i.e., 'called' for L1, 'Reynolds' for L2 and 'Monday' for L3). For this variable, Length 2 peak values are significantly higher than those for both Length 1 and Length 3 peaks (L1 vs. L2: t=3.955, p<.001; L2 vs. L3: t=2.747, p<.01), although Length 1 and Length 3 peaks are not significantly different from each other (t=1.413, p>.1).

Reset A (Peak 1A-Peak 2A) also shows a significant interaction with Clause 1 length (F=10.01, p=.0001). However, this result is most likely related to the interaction of Peaks 1A and 2A individually with the length of Clause 1. Reset N (i.e., Peak 2A-Peak 1N), on the other hand, shows little overall effect of Clause 1 length (F=2.65, p>.05), although the tendency is for this resetting measure to increase as Clause 1 gets longer (-1 Hz for L1, 2 Hz for L2, 3 Hz for L3). The difference between Length 3 and Length 1 values for Reset N is also significant (t=2.159, p<.05).

Clause 2

Fundamental Frequency: Figure 102a, b and c is a schematic representation of sentences composed of each Clause 1 length with the three length second clauses for EB. Here again, the trends are similar for each Clause 1 length condition. That is, with the exception of the Length 1 first clause utterances (i.e., 'called'), utterance-initial peaks (1A) are out of order with respect to absolute sentence length (see Table 19). The collapsed data, derived from Table 19c, are shown schematically in Figure 102d. These are labelled N-1, N-2, and N-3 and show the effect of varying the length of Clause 2 across all Clause 1 lengths. The corresponding duration measures are shown in Table 21.

Figure 103 shows schematics for the Clause 2 length comparisons for subjects BK, JS and KM, collapsed across sentence types, along with the Clause 2 length comparisons for EB (from Fig. 102d). The values from which these schematics are derived for the former three subjects are shown in Table 22. The corresponding collapsed duration measures are shown in Table 23.

Table 24 shows the interaction of each FO variable with the length of Clause 2 for subjects EB, BK, JS and KM. For all subjects, Peak 1A values are unaffected by Clause 2 length. In fact, for the latter three subjects, these values are actually quite stable across lengths (see Table 22). Even EB shows far less variability in this measure

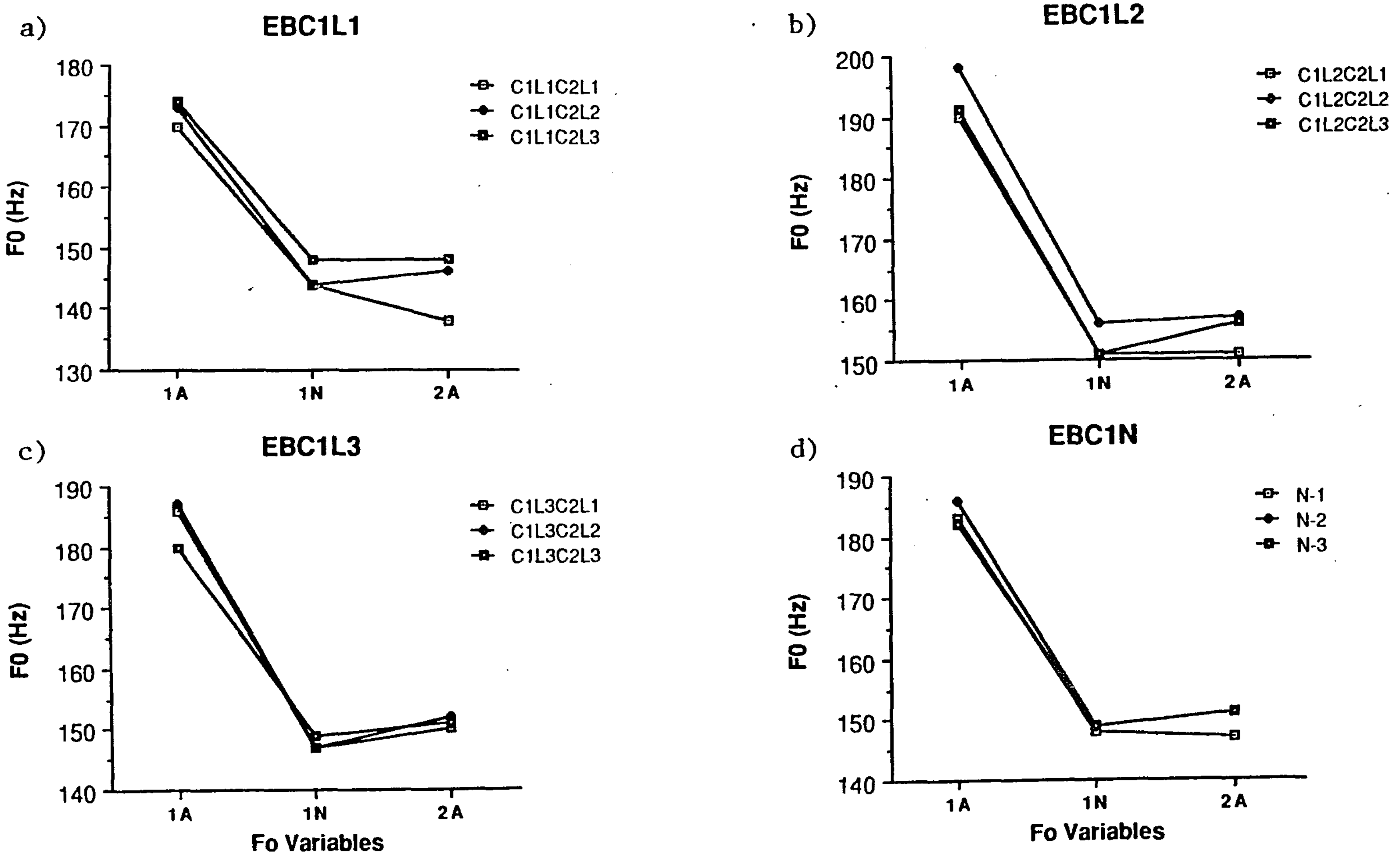


Figure 102. Key F0 variables for Clause 2 length comparisons for EB.

Table 21

Two-Clause Utterance Clause Durations (EB)

Collapsed Across Clause 2 Length

	Clause 1		Clause 2	
	Mean	SD	Mean	SD
1-N	853	56.43	1412	502.39
2-N	1277	121.61	1451	485.19
3-N	1614	260.02	1424	425.05

Collapsed Across Clause 1 Length

	Clause 1		Clause 2	
	Mean	SD	Mean	SD
N-1	1243	295.61	854	47.03
N-2	1261	337.40	1446	189.41
N-3	1239	335.57	1988	53.74

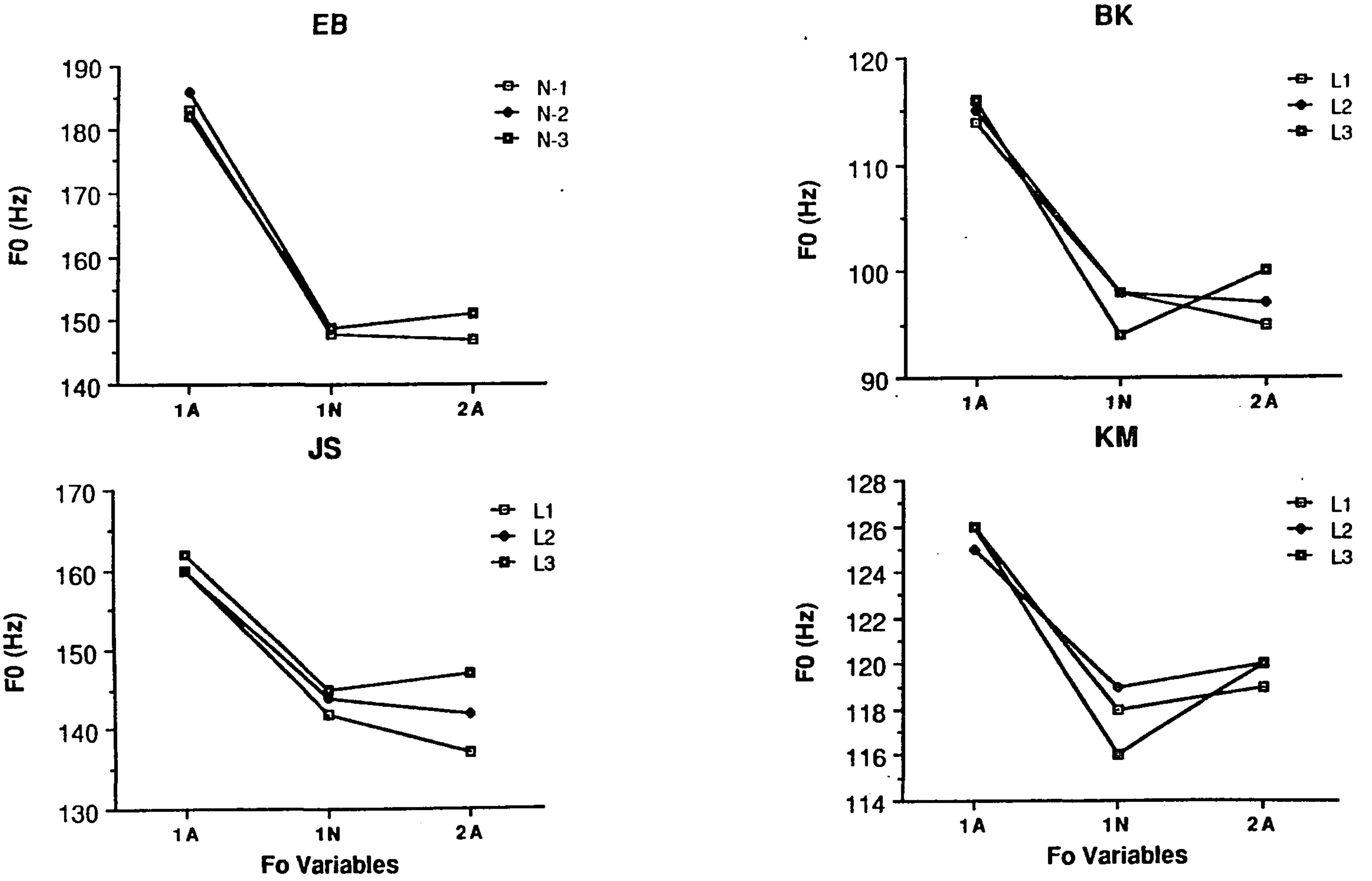


Figure 103. Key F0 variables for Clause 2 length comparisons for all subjects.

Table 22

F0 Reset Data as a Function of Clause 2 Length

	N	1A	SD	1N	SD	2A	SD	ResA	SD	ResB	SD
BK:											
L1	20	114	6.8	98	5.7	95	3.9	19	6.2	-3	4.7
L2	20	115	7.1	98	4.1	97	3.0	18	6.1	0	3.8
L3	20	116	7.0	94	22.4	100	4.1	16	7.1	+6	21.5
JS:											
L1	20	160	12.7	142	5.9	137	3.9	23	11.8	-6	4.7
L2	20	160	12.7	144	6.5	142	6.0	17	10.2	-2	6.2
L3	20	162	11.0	145	7.4	147	7.5	15	10.3	+2	11.2
KM:											
L1	20	126	4.3	118	3.3	119	2.8	7	5.0	+1	4.1
L2	20	125	3.4	119	4.4	120	2.2	5	4.3	+1	4.5
L3	20	126	3.5	116	2.9	120	2.7	6	4.4	+3	3.6

Table 23

Two-Clause Utterance Clause Durations (BK, JS, KM)

		Clause 1		Clause 2	
		Mean	SD	Mean	SD
	L1	828	30.04	896	28.09
BK	L2	854	42.76	1209	67.02
	L3	881	49.01	1625	74.83
	L1	1428	36.48	853	54.56
JS	L2	1461	71.34	1312	176.49
	L3	1448	30.35	1617	98.04
	L1	1981	32.90	809	43.63
KM	L2	2039	39.80	1309	51.13
	L3	1943	35.65	1599	117.89

Table 24

Interaction of Clause 2 Duration with Key FO Variables

	EB		*BK*		*JS*		*KM*	
	F	*p*	*F*	*p*	*F*	*p*	*F*	*p*
1A	1.85	>.1	.90	>.1	.91	>.1	.75	>.1
1N	.74	>.1	.64	>.1	1.20	>.1	6.29	<.01
2A	8.54	<.001	7.56	<.01	16.56	<.001	1.09	>.1
Reset A	2.79	>.05	7.87	=.001	.80	>.1	1.78	>.1
Reset N	1.71	>.1	2.57	>.05	4.65	<.02	5.04	<.01

when the length of Clause 1 is kept constant and the length of Clause 2 varies than when the length of Clause 1 varies and the length of Clause 2 is kept constant (Table 20).

For subjects EB, BK and JS, Peak 1N is unaffected by Clause 2 length. For KM, however, while the interaction is significant, its meaning is unclear, since Peak 1N values are out of order with respect to length. That is, for both sentence types, values for Length 2 sentences are highest, while those for Length 3 sentences are lowest. Moreover, the greatest averaged difference between any two lengths is only 2 Hz.

For EB, BK and JS, Peak 2A values increase with increasing Clause 2 length. The interaction of Peak 2A with Clause 2 length is significant, as is the increase in F0 for Peak 2A between the shortest and longest second clauses (EB: $t=2.182$, $p<.05$; BK: $t=3.952$, $p<.001$; JS: $t=5.29$, $p<.001$). For KM, peak 2A values remain virtually unchanged with increasing Clause 2 length, with a nonsignificant difference of 1 Hz separating the shortest and longest length clauses ($t=1.150$, $p>.1$). An analysis of variance shows no interaction of Peak 2A and Clause 2 length for this subject.

It thus appears that, for three of the four subjects, the presence of resetting, when defined by the frequency relationship among the Clause 2-initial peaks (i.e., Peak 2A), is a function of the length of the second clause. When defined in terms of the relationship of Clause 2-initial peaks to Clause 1-final peaks (i.e., Reset N), however, Peak

2A values alone cannot predict this resetting measure. For example, only subject JS shows an interaction in the same direction for the two variables, while, for the other three subjects, there is a significant interaction with Clause 2 length for one variable, but not for the other (Table 24). This result is probably attributable to the less systematic behavior of the Peak 1N values described above (and shown in Table 21) than to any systematic adjustments in Peak 2A.

Reset A will similarly reflect the lawful or erratic behavior of either of the two variables used to derive this measure. Of the three subjects who show a significant interaction of Peak 2A with Clause 2 length (i.e., EB, BK, JS), only JS shows the same interaction for Reset A (Table 24). Subject KM, who shows no significant interaction of either Peak 1A or Peak 2A with Clause 2 length, similarly fails to show an interaction for Reset A.

Taken together, these results suggest that the influence of varying the length of a second clause is limited to the height of the first F0 peak in that clause (i.e., Peak 2A) and does not extend beyond the preceding clause boundary. This argues against the sentence as the unit over which an F0 contour is organized. It is interesting that, although resetting shows an effect only of second clause length for all subjects (significant interactions aside), whether defined as the relationship of second clause-initial peaks (Peak 2A) or by the relative differences between these peaks and those occurring at the end of the first clause (Reset N), partial resetting, whereby Peak 2A values are higher than Peak 1N values, only occurs an average of 50% of the time across subjects. Thus, it is not the clause boundary alone that triggers F0 resetting. Moreover, there is no sentence for any subject for which there is complete resetting (i.e., Peak 2A = Peak 1A), suggesting that a speaker does not begin a new clause as if it were a new sentence. However, it is apparent that KM's Peak 1A and 2A values are closer than either of the other subjects'. It is also the case that the range of F0 fall over the first clauses (i.e., Peak 1A-Peak 1N) is significantly smaller for KM (8.5 Hz) than for either BK (16.8 Hz) or JS (17.7 Hz) (BK vs. KM: t=9.474, p<.001; JS vs. KM: t=6.328, p<.001). However, the latter two subjects do not differ from each other in terms of range of F0 variation (t=.534, p>.1). Thus, the tendency towards a greater degree of resetting may result from a speaker's tendency to produce a reduced range of F0 declination in a first clause, rather than from a strategy concerning the initiation of a second clause.

Respiration: Figure 104 shows the averaged Respitrace curves for the three lengths, collapsed across sentence type, superimposed for BK, JS and KM, respectively. It can be seen that BK (Panel 1) takes a new inspiration at the clause boundary for the longest sentences.

Because depth of inspiration was shown to vary inconsistently as a function of anticipated utterance length in Experiment 1, initial inspiratory depth was also derived for these subjects for the two-clause utterances. It can be seen from Table 25 that BK is the only subject who shows a systematic increase in depth of inspiration as a function of utterance length. This results provides further evidence

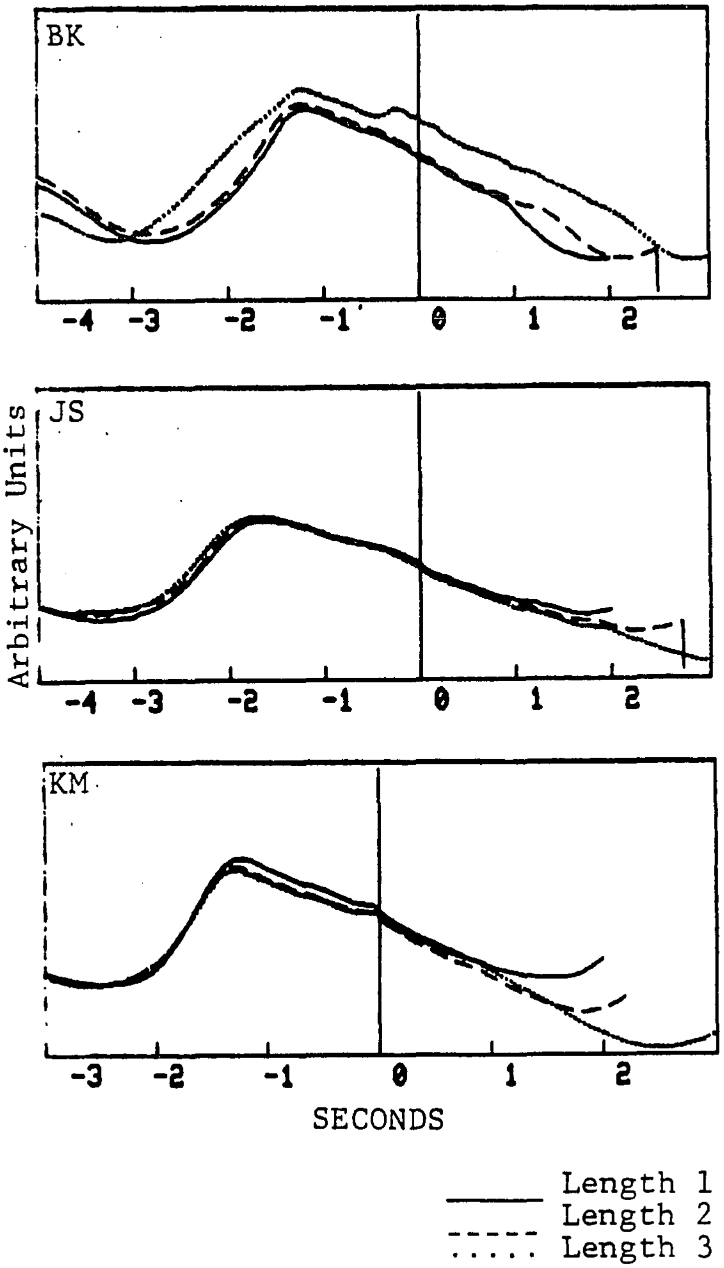


Figure 104. Depth of inspiration across lengths for two-clause utterances.

Table 25

Depth of Initial Inspiration (Arbitrary Units) as a Function of Two-Clause Utterance Length

	Depth	*SD*	*N*	
Length 1	1312	417.2	20	
Length 2	1373	354.3	20	*BK*
Length 3	1615	415.8	20	
Length 1	978	237.0	19	
Length 2	943	215.0	20	*JS*
Length 3	985	205.1	19	
Length 1	1268	379.2	19	
Length 2	1243	336.4	18	*KM*
Length 3	1176	333.9	19	

that, even when the length of an utterance is known in advance, a speaker does not necessarily make inspiratory adjustments accordingly (see also Horii & Cooke, 1978).

3. Single Clause Utterances

Subject EB produced both main clauses of the three conjoined main clause utterances as isolated sentences, while subjects BK, JS and KM produced the second clauses of the two conjoined-main clause sentence types as isolated sentences. Utterance-initial F0 peaks were derived for all tokens of the single clause utterances and compared to the corresponding peaks in the comparable clauses in the two-clause utterances. These are shown in Table 26 as 'Peak F0' along with the values with which they are being compared (i.e., Peak 1A or Peak 2A) and the differences between them. Single clause sentence durations and the corresponding second clause durations are shown in Table 27. For EB, second clause F0 values and durations are collapsed across syntactic conditions and all Clause 1 length variants. For BK, JS and KM, the data are collapsed across the two sentence types.

Comparison with First Clauses

For the single clause utterances corresponding to the first clauses (subject EB), averaged Peak F0 values are lower than those for Peak 1A by an average amount of 14 Hz, although the value of Peak 1A for the Length 2 first clauses clearly accounts for most of the

Table 26

Peak FO Values for Single-Clause Utterances as a Function of Length

First Clauses

	Peak FO	1A	Diff	
Length 1	168	173	- 5	EB
Length 2	165	196	-31	
Length 3	178	183	- 5	

Second Clauses

	Peak FO	2A	Diff	
Length 1	174	147	+27	EB
Length 2	166	152	+14	
Length 3	176	152	+24	
Length 1	102	95	+ 7	BK
Length 2	106	97	+ 9	
Length 3	112	100	+12	
Length 1	148	137	+11	JS
Length 2	158	142	+16	
Length 3	162	147	+15	
Length 1	124	119	+ 5	KM
Length 2	124	120	+ 4	
Length 3	126	120	+ 6	

Table 27

Clause Durations (in Msec) for One- and Two-Clause Utterances

		Single Clauses			*Second Clauses*		
		Mean	*SD*	*N*	*Mean*	*SD*	*N*
	L1	1006	17.46	5	854	47.03	30
EB	*L2*	1510	47.04	5	1446	189.41	30
	L3	2028	69.79	5	1988	53.74	30
	L1	1193	75.90	20	1168	98.30	20
BK	*L2*	1748	165.40	20	1707	170.50	20
	L3	2426	182.78	20	2388	135.10	20
	L1	1249	66.50	20	1214	93.30	20
JS	*L2*	1952	60.97	20	1935	81.70	20
	L3	3087	186.98	20	3065	103.60	20
	L1	945	58.39	20	939	59.00	20
KM	*L2*	1457	32.14	20	1477	47.60	20
	L3	2216	65.81	20	2216	78.90	20

difference (Table 26). There is a significant interaction of single clause sentence length with Peak F0 ($F=4.12$, $p<.05$), although the values are out of order with respect to absolute sentence length. Furthermore, while Length 3 peaks are significantly higher than Length 2 peaks ($t=3.109$, $p<.02$), they are not significantly higher than Length 1 peaks ($t=2.003$, $p>.05$).

Comparison with Second Clauses

Fundamental Frequency: Analyses of variance reveal no interaction of Peak F0 and sentence type (i.e., 'plans' or 'Paris') for BK ($F=.91$, $p>.1$), JS ($F=3.1$, $p>.05$) or KM ($F=.04$, $p>.1$). As a result, these values have been averaged across sentences and are compared to the corresponding Peak 2A data that have been similarly collapsed for second clauses (Table 26).

For all four subjects, Peak 2A values are lower than the corresponding Peak F0 values for the same clauses produced as isolated sentences. EB's Peak F0 data do not show a systematic length effect in that Length 2 peaks are lower than Length 1 peaks and Length 3 peaks are not significantly higher than those for Length 1 ($t=.258$, $p>.1$). Similarly, an analysis of variance reveals the absence of a significant interaction of Peak F0 and utterance length ($F=2.39$, $p>.1$). Thus, the trends for EB's Peak F0 data are both qualitatively and quantitatively different than his Peak 2A data.

The Peak F0 data for BK, JS and KM, on the other hand, show very similar trends when compared to their Peak 2A data. There is a significant interaction of Peak F0 with utterance length for BK (F=22.88, p<.0001) and JS (F=17.74, p<.0001), but not for KM (F=1.82, p>.1). In addition, for all three subjects, F0 peaks for Length 3 utterances are significantly higher than those for Length 1 utterances (BK: t=6.871, p<.001; JS: t=5.297, p<.001; KM: t=2.031, p<.05). It should be recalled that, for KM, this result differs from his Peak 2A data, for which no length comparisons were significant. However, for the single clause utterances, the greatest average difference between any two lengths is only 2 Hz. Thus, although statistically significant, it is difficult to imagine that this represents a true length effect, particularly when compared to differences of 10 and 14 Hz for BK and JS, respectively.

The results in this section suggest that, for peak F0, the initial portions of clauses composing two-clause utterances are not equivalent to the same clauses produced as isolated sentences. For the one subject who produced first clauses as sentences (EB), sentence-initial peak F0 was consistently lower for the single sentence condition. However, Peak F0 was consistently higher for second clauses produced as isolated sentences than the corresponding Peak 2A values for the four subjects who produced them. In addition, the relationship of F0 peaks as a function of Clause 2 length was maintained for the isolated sentences of three subjects (BK, JS, KM) while, for the fourth subject (EB), it was not.

Respiration: Depth of inspiration measures were made for the single clause utterances for BK, JS and KM in order to assess further the degree to which this measure reflects planning for an utterance of a given length. As can be seen from Figure 105 and from Table 28, only KM shows a tendency toward increased initial inspiratory depth as a function of utterance length. For BK, the absence of a length effect contrasts with that for his two-clause utterances, for which there was a tendency for utterance length to be reflected in his initial inspiration.

4. The Effect of Respiratory Patterns on Resetting

This section reports the effects on F0 resetting of directly manipulating respiratory behavior at the clause boundary. It was hypothesized that, where a new inspiration occured, Peak 2A would be higher than if the entire sentence were produced on a single expiration. This hypothesis is based, in part, on the assumption that, following an inspiration, when subglottal pressure becomes negative, the pressure will be 'reset', possibly to a higher value than if it were steadily declining over the entire utterance. A similar 'resetting' of cricothyroid muscle activity might be expected as well, although, unlike Ps, which must remain positive throughout an expiration, CT activity appears to be characterized normally by continuous on-off adjustments, as was shown in the data from Experiment 1. However, while the results of Experiment 1 show a significant peak

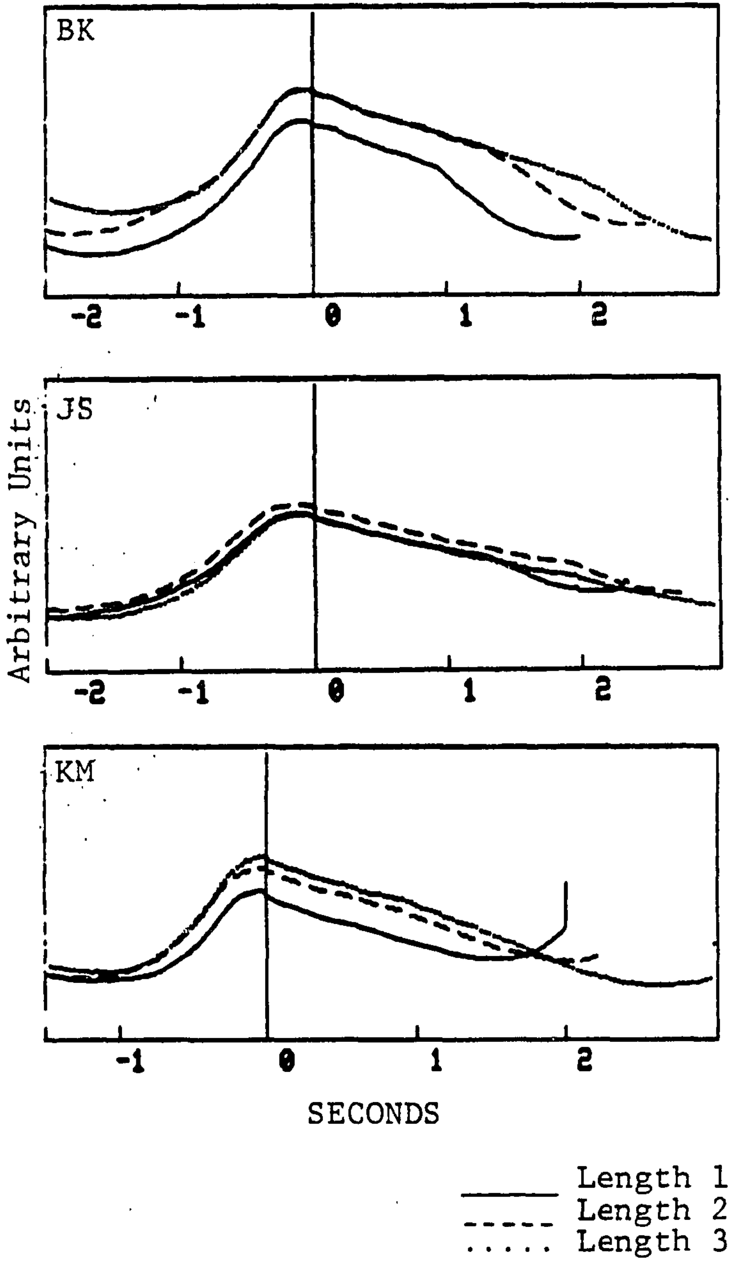


Figure 105. Depth of inspiration across lengths for single clause utterances.

Table 28

Depth of Initial Inspiration (in Arbitrary Units) as a Function of Single-Clause Utterance Length

	Mean	*SD*	*N*	
Length 1	1058	247.3	20	
Length 2	1194	366.5	19	*BK*
Length 3	1001	232.5	20	
Length 1	946	221.0	20	
Length 2	985	270.3	18	*JS*
Length 3	973	254.4	20	
Length 1	927	251.2	18	
Length 2	1147	303.9	18	*KM*
Length 3	1061	305.8	20	

of CT activity at utterance initiation, the acoustic resetting results described above show that second clause-initial F0 peaks rarely approach utterance-initial peaks. This would suggest the absence of significant CT activity following the clause boundary. However, there remains the question of whether the presence of a new inspiration at the clause boundary has the same effect on Ps and CT activity, and therefore, on F0, as at the start of a new utterance.

Subjects RC, BK, JS and KM participated in this part of the experiment, although Ps and CT data were collected only for RC. In addition to the two utterances differing in syntactic structure (i.e., subordinate-main clauses and conjoined main clauses), RC also produced a slightly altered version of the former sentence, where 'Reynolds' was changed from being the object of Clause 1 to modifying the subject of Clause 2 (i.e., When the lawyer called up, Reynolds' plans were discussed.). Subjects BK, JS and KM produced Length 1 utterances of each of the sentence types previously discussed.

Each subject produced five repetitions of each utterance under each of two conditions. In Condition 1, the sentences were produced on a single expiration. In Condition 2, the subjects were instructed to produce a new inspiration before beginning the second clause. As was done earlier, the data for the matched pair of subordinate-main and conjoined main clause utterances are combined for RC, as are the data for the two sentence types produced by BK, JS and KM.

Fundamental Frequency: The dependent variables considered here are the same as those already discussed at length in the results above. Figures 106-109 show sample F0 contours of one sentence for each respiratory condition for all subjects. Despite the differences in pause durations and respiratory activity, it is obvious that the subjects produced the same general F0 contours across conditions.

Table 29 shows the values for the key F0 variables for all subjects under each respiratory condition (i.e., -Inspiration or Inspiration). For RC, these values are averaged across syntactic condition and sentence type. For BK, JS and KM, F0 values are averaged across the two sentence types.

Not surprisingly, for all variables, there is a significant subject effect (Table 30). As for the effect of inspiring at the clause boundary (Table 30), only Peak 1N fails to show an effect of respiratory condition. All other measures show a significant interaction with respiratory condition.

Figure 110 shows schematic representations of the dependent F0 values averaged across sentence types. For all subjects, Peak 1A values are higher for utterances produced on a single expiration than those for which a new inspiration occurs at the clause boundary, although the difference is significant only for BK (RC: $t=.699$, $p>.1$; BK: $t=3.751$, $p<.01$; JS: $t=2.042$, $p>.05$; KM: $t=1.816$, $p>.05$). This difference is most striking for subjects BK and JS, and suggests, perhaps, a kind of length effect for utterance-initial F0 whereby the

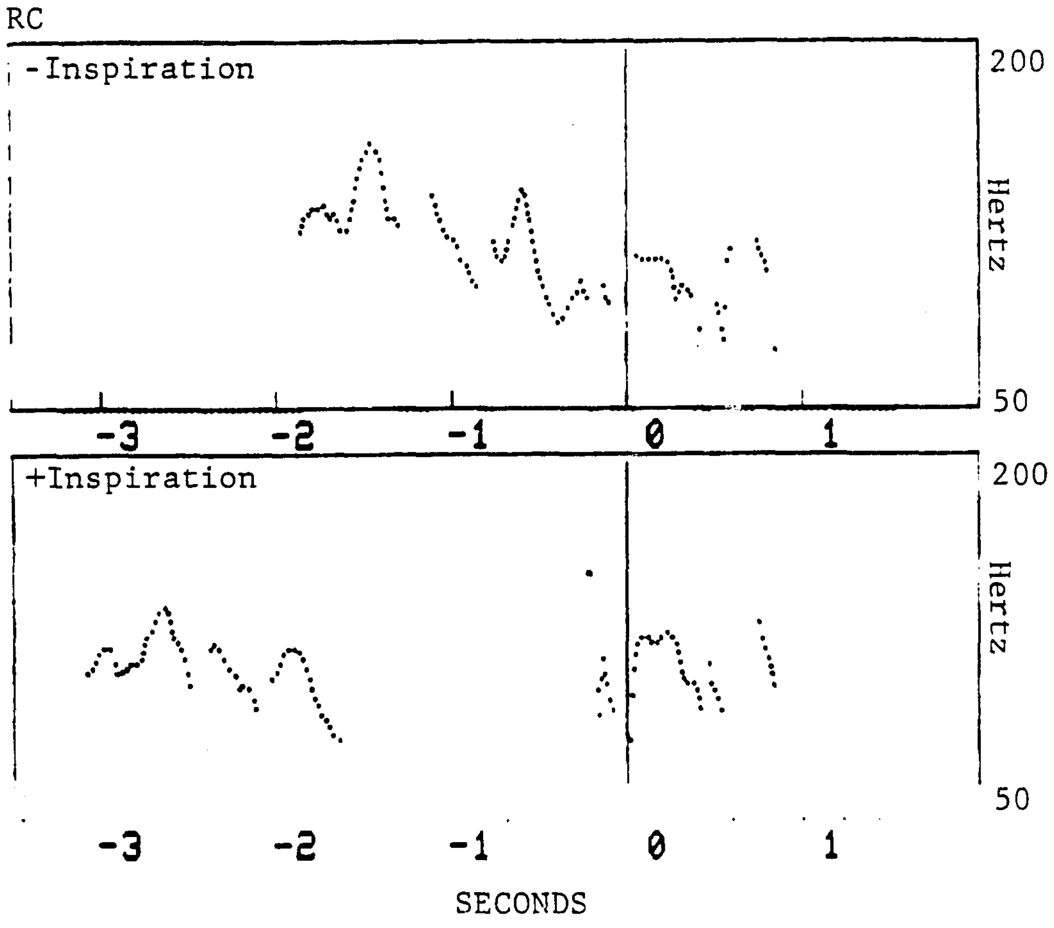


Figure 106. Fundamental frequency contours for two respiratory conditions for RC.

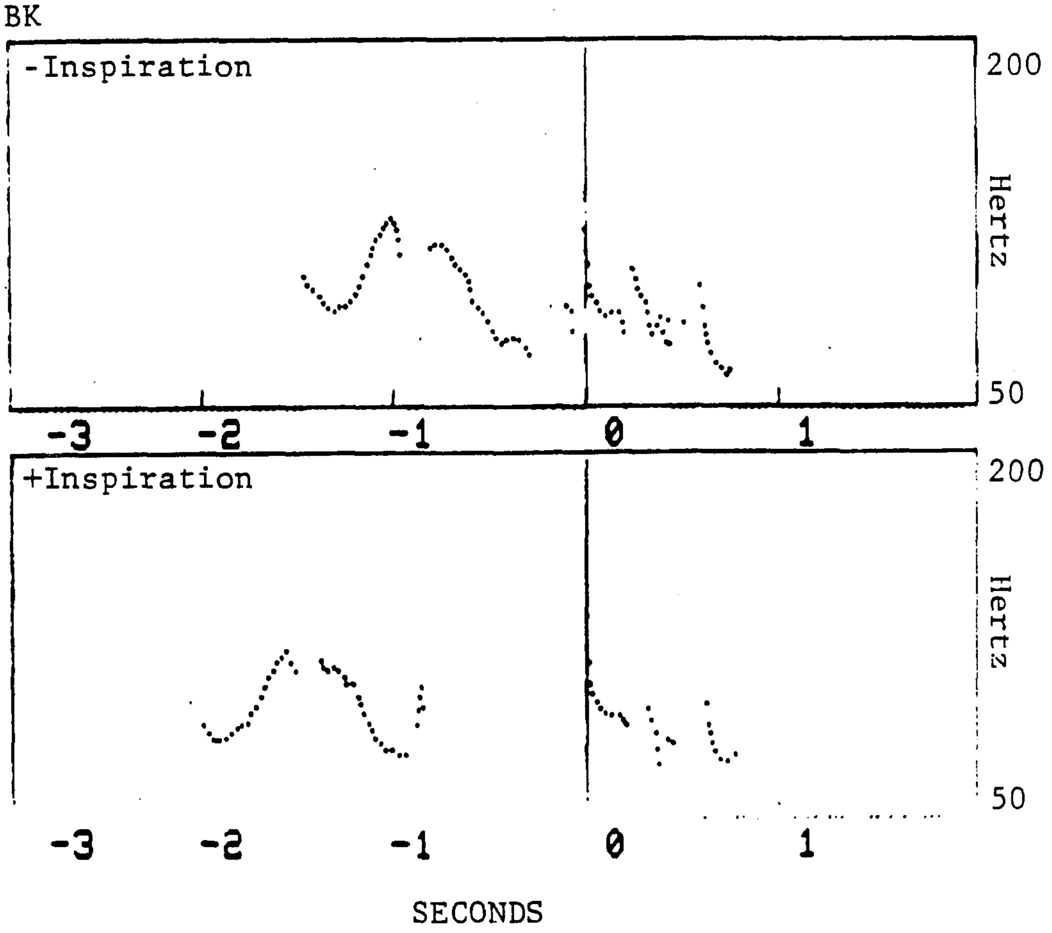


Figure 107. Fundamental frequency contours for two respiratory conditions for BK.

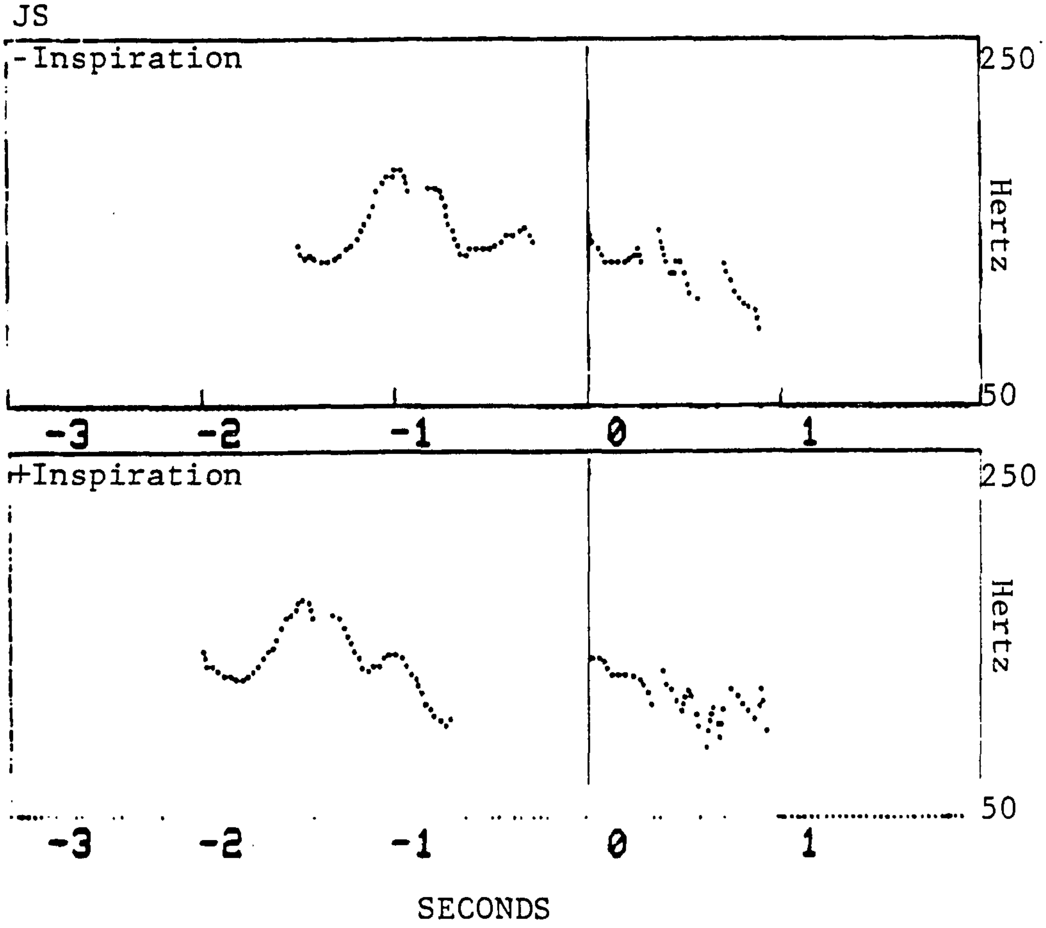


Figure 108. Fundamental frequency contours for two respiratory conditions for JS.

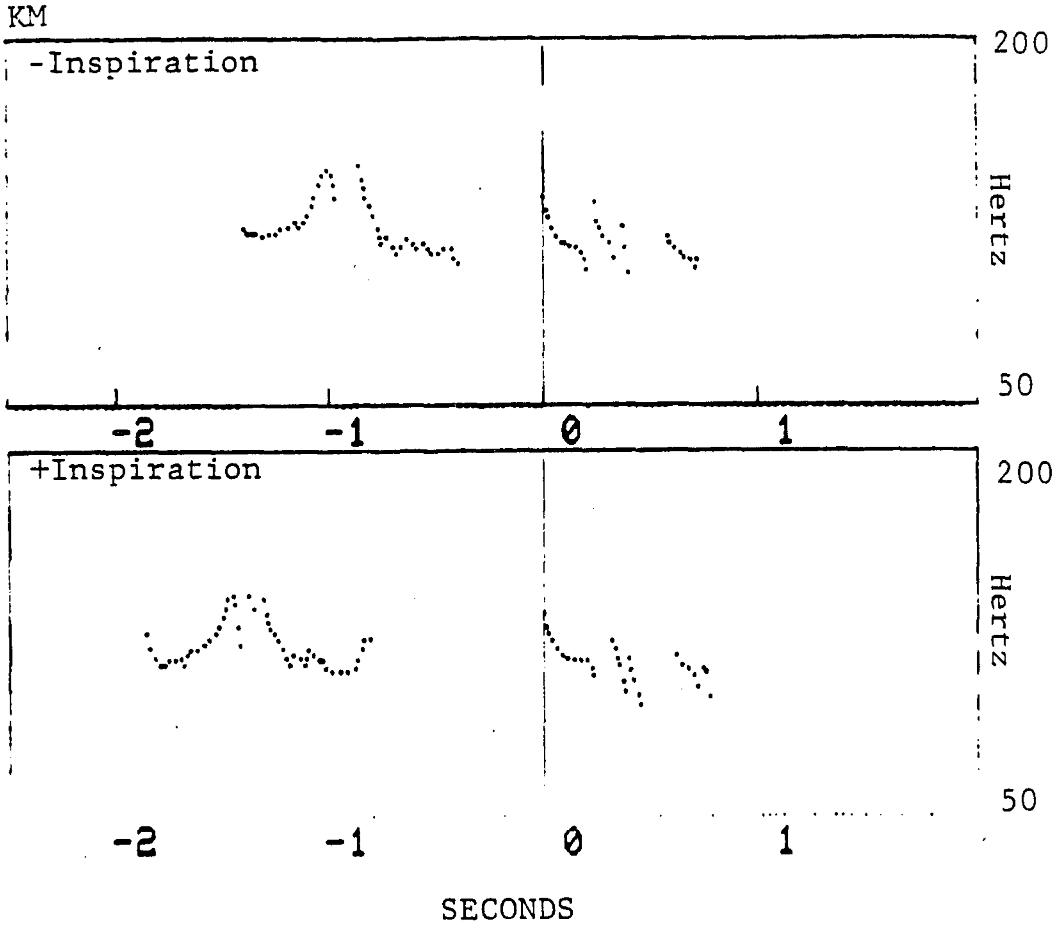


Figure 109. Fundamental frequency contours for two respiratory conditions for KM.

Table 29

Peak FO as a Function of the Presence or Absence of an Inspiration at the Clause Boundary

	N	1A	SD	1N	SD	2A	SD	ResA	SD	ResN	SD
RC:											
-IN	13	149	4.6	127	5.1	119	3.4	30	6.6	-8	6.7
+IN	14	147	9.3	124	4.5	126	4.1	21	6.7	+2	2.7
BK:											
-IN	10	116	4.7	96	3.7	91	4.3	25	4.6	-5	3.9
+IN	10	106	7.0	98	4.0	96	4.0	10	5.4	-2	5.6
JS:											
-IN	10	168	10.9	147	10.7	135	7.6	32	14.9	-12	10.5
+IN	10	158	11.0	149	9.2	132	7.2	26	12.5	-17	10.9
KM:											
-IN	10	124	2.2	118	2.9	116	2.6	8	3.0	-2	3.7
+IN	10	122	2.7	116	2.5	120	2.4	2	3.3	+4	1.6

Table 30

Interaction of Key F0 Variables with a) Subject and b) Respiratory Condition

a)

	1A	1N	2A	Reset A	Reset N
F	209.53	258.73	261.92	35.67	20.83
p	<.0001	=.0001	<.0001	<.0001	<.0001

b)

	1A	1N	2A	Reset A	Reset N
F	13.18	.49	9.54	26.29	8.56
p	<.001	>.1	<.01	<.001	<.01

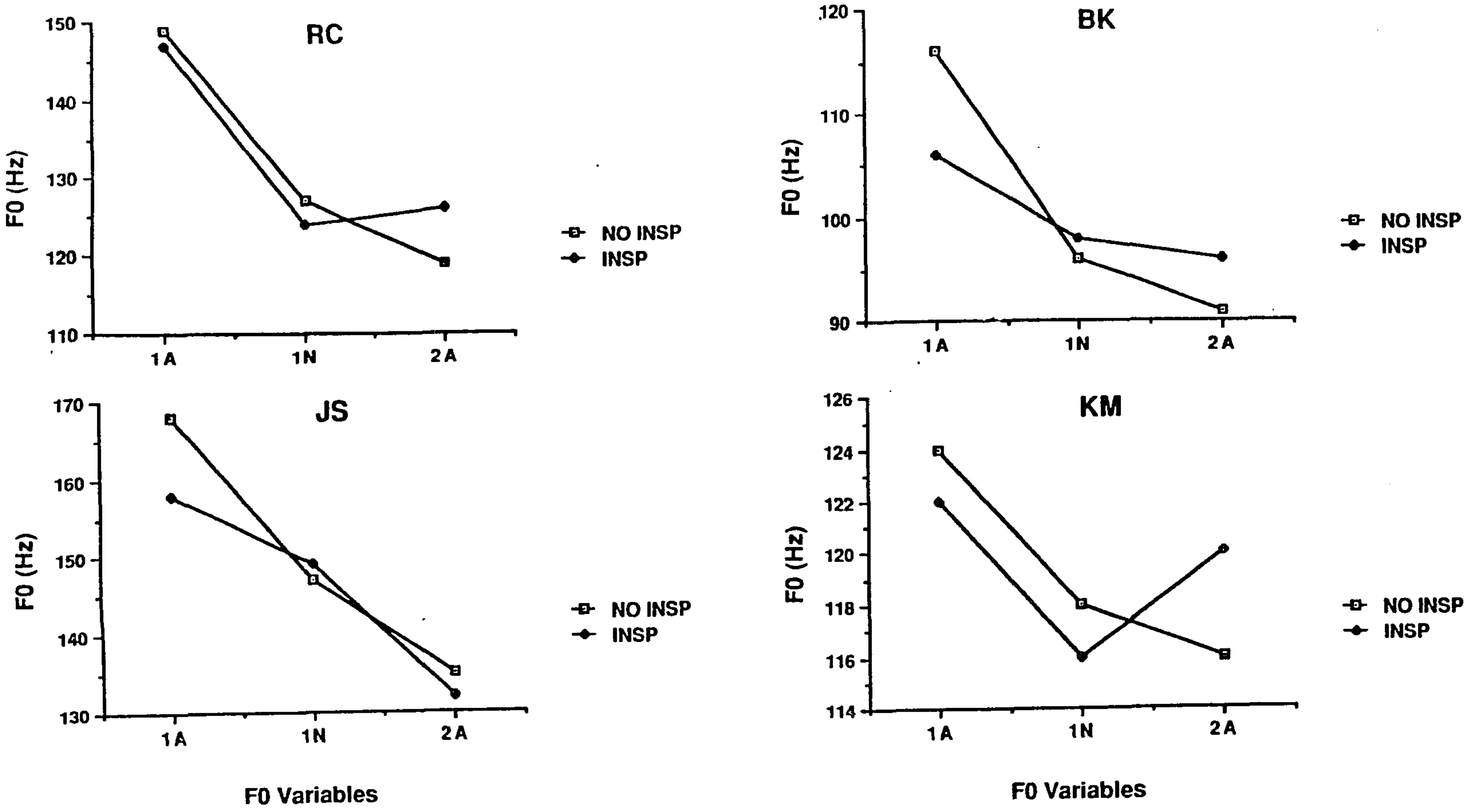


Figure 110. Key F0 variables as a function of presence or absence of inspiration at the clause boundary.

anticipated length of the breath group (i.e., the unit produced on a single expiration) is taken into account by the speaker, regardless of its identity as a clause or the sentence.

Values for Peak 2A, for which it was expected that the effects of inspiring at the clause boundary would be most evident, are significantly higher following an inspiration for three of the four subjects (RC: t=4.808, p<.001; BK: t=2.692, p<.02; KM: t=3.575, p<.01). For JS, Peak 2A values are an average of 3 Hz lower following an inspiration. Similarly, he is the only subject for whom Reset N values are lower in the 'inspiration' condition. Neither of these differences, however, is significant (Peak 2A: t=.906, p>.1; Reset N: t=1.045, p>.1). It is interesting that Reset N values indicate no partial resetting (i.e., Peak 2A < Peak 1N) for any subject in the 'no inspiration' condition, while even in the inspiration condition, only two subjects show partial resetting (i.e., Peak 2A > Peak 1N). Once again, this result suggests that a clause boundary does not automatically trigger resetting, although resetting is more likely following an inspiration at the boundary.

Reset A values are higher for all subjects when no inspiration occurs at the clause boundary. The difference is significant for three subjects (RC: t=3.513, p<.01; BK: t=6.687, p<.001; JS: t=.976, p>.1; KM: t=4.254, p<.001). Note, however, that for KM, the difference between Peaks 1A and 2A in the inspiration condition is not significant (Peak 1A-Peak 2A=2 Hz; t=1.751, p>.1), as compared to the no inspiration condition (Peak 1A-Peak 2A=8 Hz; t=7.428, p<.001).

However, it is probably unwise to assign too much significance to this result, as it was shown earlier that this subject typically produces a restricted range of F0 declination across first clauses. Thus, even very small increases in Peak 2A for KM will result in complete resetting - a result that may or may not reflect statistical artifact.

Subglottal Pressure: Figure 111 shows the averaged Respitrace and Ps curves for one sentence across the two respiratory conditions (i.e., -pause/-inspiration and pause/inspiration) for subject RC. In the absence of both a pause and an inspiration at the clause boundary (Panel 1), there is a continuous, although choppy, subglottal pressure curve throughout both clauses and across the intervening boundary as well. On the other hand, where both a pause and inspiration occur (Panel 2), there is a concommitant drop in the subglottal pressure during the inspiration which then increases significantly as expiration resumes.

A comparison of Ps values at Peak 2A yields results that correspond to those for F0 for this subject (top of Table 31). That is, subglottal pressure is significantly higher when a pause and inspiration occur than when they do not ($t=6.471$, $p<.001$). Moreover, when the ratio of frequency change per centimeter of water is calculated for Peak 2A between Conditions 1 and 2 (bottom of Table 31), these ratios fall within the accepted range of 2-7Hz/Cm-H2O (Baer, 1979; Hixon, Klatt, and Mead, 1971; Ladefoged, 1963), suggesting that the relationship between the increase in Ps and that in F0 could be more than a correlational one. However, before the behavior of F0 is

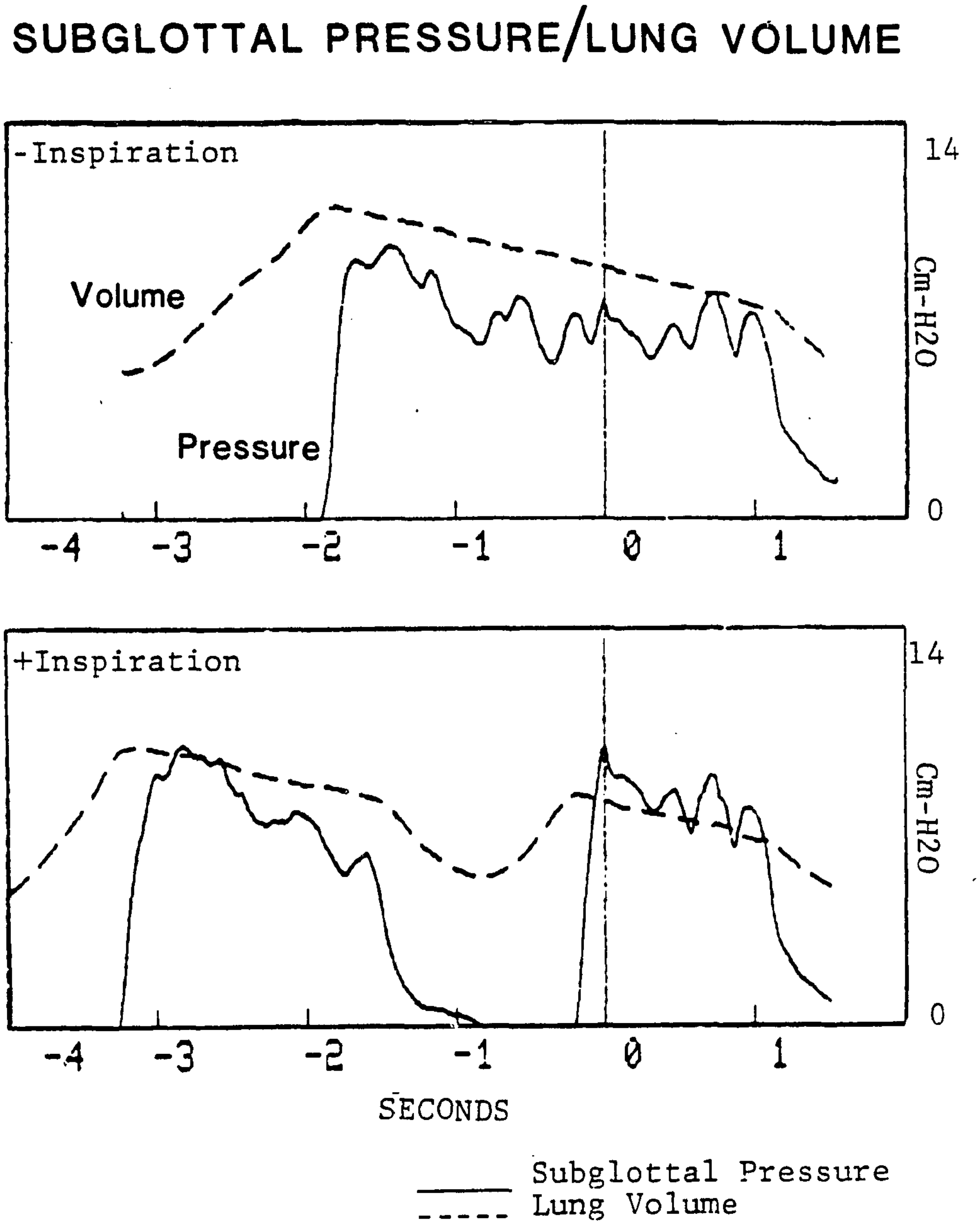


Figure 111. Subglottal pressure and lung volume as a function of respiratory behavior at clause boundary.

Table 31

Peak 2A Values and Ratio of Hz/Cm-H2O as a Function of Respiratory Condition for Subject RC

	Fundamental Frequency		Subglottal Pressure	
	-Insp	+Insp	-Insp	+Insp
Sentence 1	116	123	7.8	9.9
Sentence 2	117	122	8.6	10.1
Sentence 3	118	130	9.1	11.6

	Condition 1-Condition 2		Ratio
	FO	Ps	Hz/Cm-H2O
Sentence 1	7	2.1	3.33
Sentence 2	5	1.5	3.33
Sentence 3	12	2.5	4.80

attributed solely to the presence or absence of an increase is Ps, the contribution of laryngeal muscle activity must be determined.

Cricothyroid Muscle Activity: Figure 112 shows the cricothyroid muscle activity for the two conditions for the same sentence shown above. It appears that there is no systematic resetting of CT activity as a function of inspiring at the clause boundary. In fact, there is more CT activity following the clause boundary in the first condition, where no inspiration occurs, than in the second condition where an inspiration precedes the boundary. It would thus appear that CT contributes little, if anything, to F0 resetting for this subject, and that its relative inactivity can account for the absence of complete resetting.

Respiration: Depth of inspiration measures were derived for RC, BK, JS and KM in order to address two issues raised earlier in this section. First, does the depth of an initial inspiration reflect planning for the number of clauses to be encompassed by a breath group? That is, does a speaker take in a smaller amount of air when he knows that he will inspire again at the clause boundary (i.e., one breath group = one clause) or, conversely, does he take in a larger amount of air when he knows he will not (i.e., one breath group = two clauses = a sentence)? Table 32 and Figure 113 show the amount of the initial inspiration, collapsed across all sentence types, for the two respiratory conditions for all subjects. For subjects BK, JS and KM, there is a tendency for the inspiration to be greater when the entire sentence is produced on a single expiration, suggesting that it is the

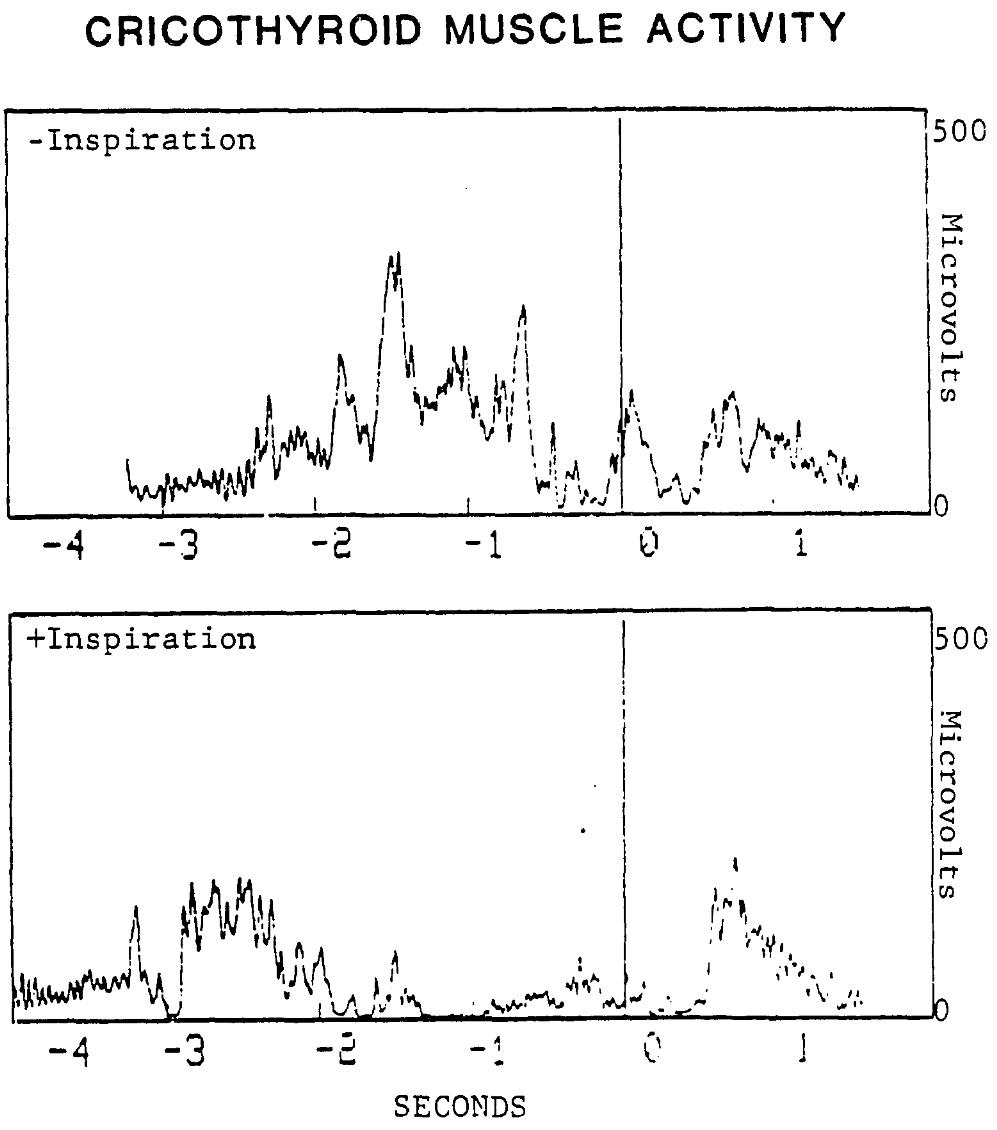


Figure 112. Cricothyroid muscle activity as a function of respiratory behavior at clause boundary.

Table 32

Depth of Initial Inspiration (in Arbitrary Units) as a Function of Respiratory Activity at the Clause Boundary

	Mean	SD	N	
-Inspiration	3190	644.6	13	RC
+Inspiration	3261	474.0	14	
-Inspiration	1676	578.6	9	BK
+Inspiration	1366	349.7	10	
-Inspiration	998	249.6	10	JS
+Inspiration	887	112.3	10	
-Inspiration	1005	198.6	10	KM
+Inspiration	884	153.3	9	

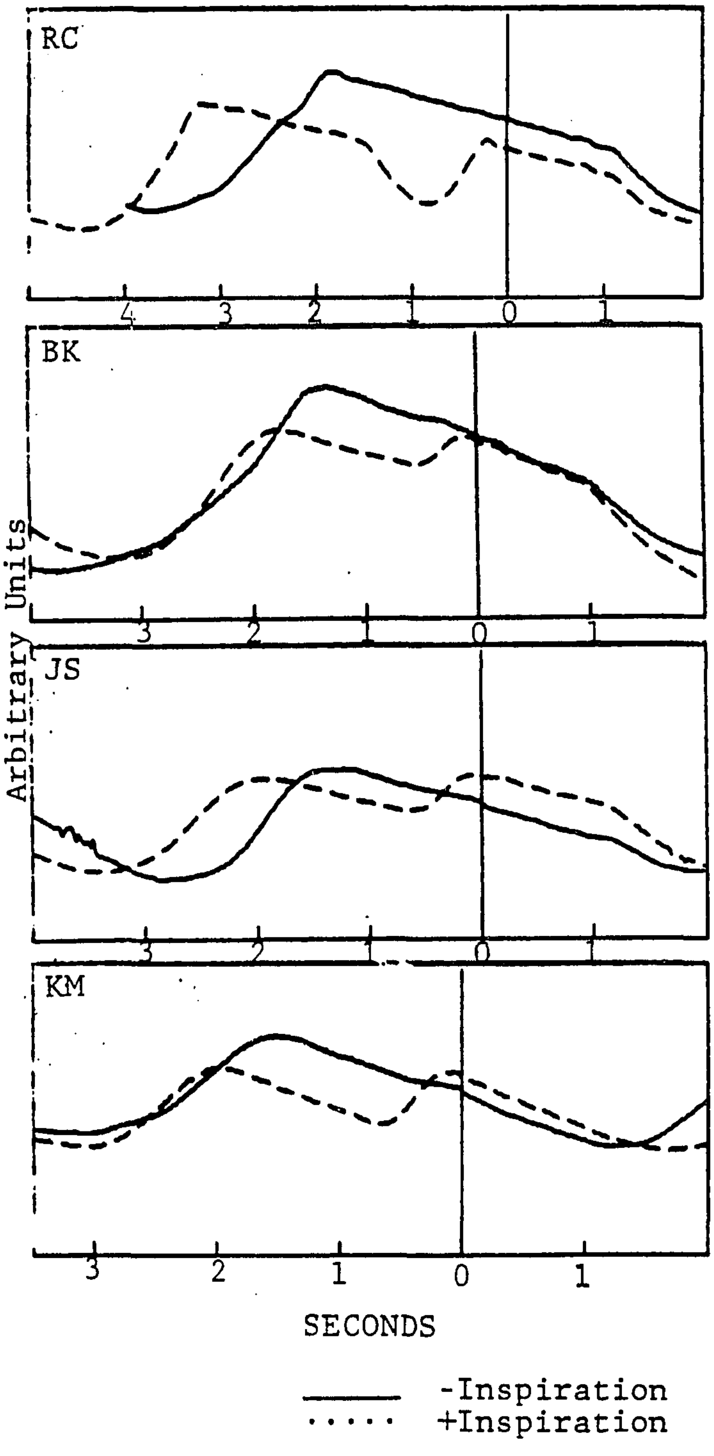


Figure 113. Depth of inspiration as a function of preceding and succeeding respiratory behavior.

size of the breath group that is taken into account. However, the differences between the two conditions are not statistically significant (BK: t=1.431, p>.1; JS: t=1.282, p>.1; KM: t=1.474, p>.1) and an analysis of variance confirms the absence of a significant interaction (F=2.18, p>.1). Thus, the notion that a speaker adjusts his initial inspiration on the basis of the size of the breath group is not supported by these data.

The second issue concerns the question of whether the start of a new clause is equivalent to the start of a new sentence. In the earlier part of this section, this question was addressed by comparing single clause utterance-initial F0 values (Peak F0) to comparable second clause-initial values (Peak 2A). The same question is addressed here by comparing the depth of an utterance-initial inspiration to the inspiration occurring at the clause boundary within the same utterance. The dashed contours in Figure 113 shows the averaged Respitrace curves for the subordinate-main clause sentences for which there was an inspiration at the clause boundary. For all subjects, the initial inspiration is significantly greater than the inspiration occurring at the clause boundary (Table 33), suggesting that the clause and the sentence are not equivalent.

Horii and Cooke (1978) also found that intrasentence inspirations were shallower than those occurring at sentence initiation. However, they attribute this result to the durations of the accompanying pauses, with intersentence pauses being longer than intrasentence pauses. This finding cannot be addressed directly here, as the subjects in this

Table 33

Depth of Inspiration (in Arbitrary Units) at Sentence and Second Clause Initiation and Significance of Differences

	Initial Inspiration			Second Inspiration			Significance of Difference	
	Mean	*SD*	*N*	*Mean*	*SD*	*N*	*t*	*p*
RC	3261	474.00	14	1905	396.80	14	8.208	<.001
BK	1366	349.70	10	338	168.26	10	8.377	<.001
JS	887	112.30	10	380	78.03	10	11.724	<.001
KM	884	153.29	9	574	146.97	10	4.499	<.01

study read single sentences with instructions to return to resting respiratory levels in between, while Horii and Cooke's subjects read lengthy paragraphs with no instructions regarding respiratory behavior. Thus, the results of the two studies are not entirely comparable. However, it is also possible that the depth of these respective inspirations is a function of where one is in one's vital capacity when an inspiration begins. In other words, while these subjects knew they were to breathe at the clause boundary, and therefore had every opportunity to plan for it, they were also at a higher point in their vital capacities than when they began their utterance-initial inspirations, as is apparent from Figure 113. Thus, it is difficult in this case to conclude definitively whether a shallower inspiration at the clause boundary is a function of physiological or linguistic demands, or, both.

Summary

The results in this part of Experiment 2 have shown that, for three of the four subjects, when there is a new inspiration at the clause boundary, peak F0 values at the beginning of the second clause are higher than when the entire utterance is produced on a single expiration. However, even in the presence of a new inspiration, the likelihood that there will be even partial F0 resetting, whereby second clause-initial F0 values are higher than those at the end of the first clause, cannot be predicted. Of the two subjects who showed resetting following a new inspiration, the largest average amount was only 4 Hz.

Thus, taken together with the acoustic data in the first part of this experiment and RC's Ps and CT data presented here, these data suggest that, even following a new inspiration, a speaker does not increase his CT activity to a level remotely comparable to that which occurs at utterance initiation, and that, as a result, the differences observed between the two respiratory conditions can be attributed to the corresponding levels in subglottal pressure.

FOOTNOTES

[2] Utterance durations are also influenced by phonetic composition. For RC, Dutch utterances are, on average, the shortest, followed by the /ma/ and /fa/ utterances, respectively. There are two explanations for this result. The first is that the vowel /a/ is inherently longer than other vowels. The second, and probably most significant, is that, in producing these reiterant utterances, syllables were occasionally added inadvertently, thus causing the /ma/ and /fa/ utterances to have slightly longer durations.

[3] When the Double and Late stress tokens are averaged after being realigned at the _final_ stress peaks, final CT activity is actually greater than is indicated by these figures.

CHAPTER 5

DISCUSSION

The general purpose of this study has been to gain insights into some of the organizing principles of speech production by examining various aspects of sentence intonation, with a view to separating those aspects of a fundamental frequency contour which are variable from those which remain stable, and to identify their physiological substrates. In addition, it has been a goal of this study to demonstrate that theories of sentence intonation involving notions of high-level cognitive processing and extensive speaker preplanning are unnecessarily elaborate.

This study has had a relatively restricted goal as compared to the substantial number of recent examinations of declination and related issues. For example, it has not addressed the issue of invariant acoustic endpoints except to suggest that they might be the result of mechanisms other than those previously proposed. And, although this is largely a study of fundamental frequency declination, no attempt has been made to measure the amount of declination over a wide variety of experimental conditions. Of issue is not what speakers do under every possible set of circumstances, but, rather, what they do when they are given every opportunity to plan for the details of the execution of a sentence. That a speaker might evidence patterns in his spontaneous speech that differ from those which occur when producing utterances under relatively constrained experimental conditions (eg., Umeda, 1982) is a distinct possibility, but one which is irrelevant to the purposes and conclusions of this study.

ACOUSTIC RESULTS

The results of previous studies that have examined utterance-initial peak F0 as a function of utterance length have been inconclusive in that, while some have found systematic length adjustments (eg., Bruce, 1982; Cooper & Sorenson, 1981), others have not (eg., Maeda, 1976). On this point, the results of the present study show variability from speaker to speaker. In the first experiment, subject RC showed systematic increases in peak F0 for the Dutch, /ma/ and /fa/ utterances, but a less consistent relationship with length for the English utterances. Subject EB's peak F0 data, on the other hand, showed no systematic length adjustments. In Experiment 2, Clause 2-initial peak F0 (i.e., Peak 2A) showed systematic length effects in the data of three of the four subjects, including EB, who produced sentences in which the length of the second clause was varied. On the other hand, no subject who produced two-clause utterances of varying lengths varied sentence-initial peak F0 in any systematic fashion, although, again with the exception of EB, three of the four who produced single clause utterances did produce this length effect. Thus, not only do different speakers appear to adjust initial F0 (sentence or clause) differentially on the basis of length, but the same speaker used different strategies under various linguistic conditions.

If length-dependent adjustments are to be considered instances of 'hard' preplanning in the sense that they are essential to the realization of an utterance (Liberman & Pierrehumbert, 1982), then they should be evident, if not for all speakers, then certainly in all utterances spoken by the same speaker. It is clear, however, that the results on peak F0 in these experiments do not support a notion of planning on this level. On the other hand, to the extent that they appear to represent optional strategies, both within and across speakers, these results might be considered instances of Liberman and Pierrehumbert's category of 'soft' preplanning.

A related, and perhaps more interesting question, is the influence of length-dependent peak F0 adjustments on the trajectory of the F0 contour as a whole. Cooper & Sorenson's (1981) Topline Rule predicts that, when initial F0 is adjusted according to utterance length, there will be a reorganization of the trajectory of the entire declination because the Topline Rule uses initial and final F0 peaks to compute the values of intermediate peaks, and thereby the slope of the declination. In the present experiment, however, the data for RC, for whom the most systematic adjustments of initial F0 occurred, provided no evidence to suggest that these adjustments influenced the overall trajectory of the declination contour. Rather, their sphere of influence was limited to the peaks themselves and, to some extent, the frequency from which the declination began. This latter frequency is probably similar to what Cooper and Sorenson (1981) have referred to as the 'key point.' However, Cooper and Sorenson do not actually identify this frequency in their data, but predict it as being the point at which half the total drop in F0 occurs (i.e., initial peak-endpoint). Whether this is an accurate formula for deriving this point is open to question, as, in this study, initial peak frequencies were never used in those analyses for which calculations of declination range were made. However, while there appears to be a relationship between initial peak F0 and the point from which the declination begins, even substantial differences in the former do not appear to influence the overall trajectory of the declination, as Cooper and Sorenson suggest. Thus, while increasing the height of peak F0 may serve some communicative function, it does not appear to be an essential aspect of the realization of an intonation contour. Nor is it likely that the 'key point' is a reflection of preplanning as much as it is a function of temporal constraints whereby, within a brief time interval, the frequency to which F0 falls is a function of the frequency from which the initial peak starts. Moreover, this time interval appears to be determined by the cessation and/or relaxation of initial physiological activity. Thus, except to the extent that a speaker 'plans' to produce a stressed syllable, and perhaps makes some initial length adjustments, there is no evidence for preplanning insofar as one actually calculates the point at which the declination will begin.

By suggesting that initial F0 values must be known in advance of the execution of an utterance, Cooper and Sorenson's planning hypothesis also fails to account for any unsystematic variability. For example, EB's peak F0 data in the first experiment show neither systematic length adjustments for peak F0 nor invariant peak frequencies as a function of length. Yet, the trajectory of the overall F0 contour appears to be insensitive to these peak frequencies in the same way that RC's F0 contours were.

The variations in initial peak F0 that result from the manipulation of early utterance stress do not appear to be qualitatively dissimilar from the length-dependent adjustments in F0 insofar as there is little evidence to suggest a reorganization of the declination depending on the presence or absence of early stress. This finding lends support to Bruce's (1982) conclusion that the declination is not reorganized on the basis of the placement of sentence prominence, per se, although the relative height of F0 for a syllable receiving emphatic stress may influence that of the subsequent declination. On the other hand, the findings of Thorsen (1979b) and Eady and Cooper (1986) that, following prominence early in an utterance, there is a lowering of the declination relative to those utterances without it, is not supported by the results of the present study. However, as was suggested earlier, these discrepancies may be a function of whether one looks for these effects in the topline or baseline declination.

As for the effect of manipulating late utterance stress, the general finding of this study is that it has little significant influence on the trajectory of the declination up to the vicinity of that stress peak, although, for a subset of the data for one subject (i.e., RC's Dutch, /ma/ and /fa/ utterances), the F0 curves for the

Late stress utterances evidenced a more gradual decline from the initial peaks than they did for the other two stress configurations. However, because the Double stress utterances, for which there were both early and late stress peaks, showed declination contours that were quite similar to those for the Early stress utterances, the differences in the contours of the Late stress utterances are probably a function of the absence of stress early in the utterance and not influenced by the presence of late stress.

Although this study did not attempt to characterize or model F0 declination in any detail, an attempt was made to determine whether a linear model which fits a straight line to the declination, either through the F0 peaks (eg., Cooper & Sorenson, 1981) or the F0 minima (eg., Maeda, 1976), is appropriate. A common finding of studies that assume linearity is that slope decreases inversely with utterance length. The same result was found in this study when slope was calculated simply as the difference in frequency over time over entire utterances, excluding initial and/or final peaks. However, a strict linear model of declination also predicts that the slope will be invariant, whether measured over an entire utterance or some part of it. This result, however, was not found. Without exception, when slope was calculated over fixed intervals occurring during the early portion of the declination, the rate of decline was greater than when calculated over the entire utterance. Moreover, across utterances of varying lengths, the slopes were more nearly equivalent - that is, as measured over entire utterances, the length effect virtually disappeared.

The finding that the early portion of an utterance evidences a steeper slope than the utterance as a whole argues rather strongly against modelling F0 declination as a linear function. It is, however, compatible with theories that model declination as a second-order linear system, or an exponential function (eg., Fujisaki et al., 1982; Liberman & Pierrehumbert, 1982). Of course, this result applies only to those portions of the declination that evidence the most gradual decline in F0 and excludes the rapid drop in F0 from initial peaks.

The results of Experiment 2 suggest that the term 'F0 resetting' may be a misnomer, as initial F0 values for second clauses were often lower than those at the end of a preceding clause, whether or not an inspiration preceded the clause boundary. Moreover, these data provide virtually no evidence for 'complete resetting' in that second clause-initial values do not even begin to approach sentence-initial F0. This result contrasts with Collier's (1987), who found considerable evidence of complete resetting in the two-clause Dutch utterances produced by one speaker. It is, of course, always possible that language or experimental material may account for conflicting results. However, while the effects of the former cannot be ruled out in this case, the effects of the latter can in that the same second clauses produced as isolated sentences in this study yielded significantly higher initial F0's than when produced as part of two-clause utterances. Thus, it cannot be the material itself that results in lower second clause-initial values (relative to sentence-initial values) than those found by Collier.

These data thus suggest that beginning a new clause is not the equivalent of beginning a new sentence, even if the two are identical in every way but for their syntactic environments. However, the acoustic differences between the two may not be related to differences in their mental representations as clauses or sentences. Rather, the absence of resetting in two-clause utterances may instead be related to differences in the degree of physiological effort that may distinguish sentence-initial from clause-initial events. This is not to say, however, that there is no interaction between syntax and sentence intonation. In fact, the finding that there is a length effect for second clause-initial F0 values (i.e., Peak 2A) that does not extend beyond the clause boundary suggests that the clause, as opposed to the sentence, per se, has reality as a unit of organization in speech.

PHYSIOLOGICAL RESULTS

Where peak F0 was shown to vary with utterance length in Experiment 1, the same relationship was observed for peak cricothyroid muscle activity and peak subglottal pressure. However, in both Experiments 1 and 2, the depth of the initial inspiration did not serve as a consistent index of anticipated utterance length. Rather, the inspiratory data in this study, in conjunction with previous studies (Horii & Cooke, 1978), suggest that speech is produced within in a part of the vital capacity range that is probably habitual and which allows for a certain amount of normal variation in the production of sentences.

One of the larger controversies in previous physiological investigations of F0 control concerns the relative contribution of the 'larynx versus the lungs' in sentence intonation. The data presented here suggest that it is the activity of the cricothyroid muscle (and perhaps other laryngeal muscles that contribute to increasing the tension/and or length of the vocal folds) that is involved in the implementation of F0 increases that occur for utterance-initial and other prominent (i.e., accented or emphasized) syllables. For while there are significant accompanying increases in Ps which also show high correlations with those for F0, these increases in Ps alone are unable to account for the magnitude of the F0 increases. This result contradicts Lieberman's (1967) claim that momentary increases in subglottal pressure represent the primary physiological correlate of stressed syllables, with the CT being active only in cases of utterance-final rises that accompany marked breath groups (i.e., questions). On the other hand, it supports the claim of Fromkin and Ohala (1968) that increases in Ps account for only a small fraction of the increases in F0.

It is also likely that the rapid drop in F0 from utterance-initial peaks is a function of the rate at which CT relaxation and/or inhibition occurs, despite the fact that the offset of CT activity occurs in combination with a relatively rapid decrease in Ps from its initial peak (Titze & Durham, 1987). As with the initial increases in Ps, the fall-off in pressure is more gradual than the corresponding fall-off in F0, and is probably insufficient to account for the magnitude and rate of the initial decline in F0 from these peaks. The point at which the declination actually begins (i.e., Cooper & Sorenson's 'key point'), however, can be explained as a combined effect of CT suppression and the onset of the more gradual Ps decline. Thus, positing elaborate mental computations on the part of the speaker in order to calculate this point is unecessary.

Following these initial peaks, there is no evidence from this study to suggest that the relaxation of the cricothyroid muscle and/or the activity of the strap muscles contributes to F0 declination. Thus, because the effects of CT and strap muscle activity were shown to be negligible, it follows that the nature of the F0 decline over these intervals is the manifestation of the expiratory (i.e., pressure) forces that act upon the vocal folds. In fact, the subglottal pressure contours evidence the same significant characteristics as those observed for F0. That is, for both RC and EB, despite variability in initial peak Ps, whether due to length or stress effects, the overall trajectories of the Ps curves were stable. Similarly, where the rate of Ps decline was calculated over entire utterances for RC, the slope was shown to vary inversely with utterance length. However, when calculated over fixed time intervals during the early part of its decline (following initial peaks), the apparent length effect virtually disappeared. Thus, a strict linear model also fails to characterize the fall in subglottal pressure over time.

While these particular analyses were performed only on reiterant /ma/ utterances, it was also demonstrated, both for RC and EB, that the CT and strap muscle activity accompanying utterances with different phonetic compositions had no discernible effect upon their F0 contours. Thus, it is reasonable to assume that subglottal pressure is the major force in determining the trajectory of the declination of any utterance for which the influence of other known forces can be assumed to be negligible.

Admittedly, this represents only a first step in quantifying the relationship of these variables in F0 declination, and may indeed characterize only a particular class of utterances spoken under a particular set of circumstances. However, the results of this study support the conclusions of some of the studies that precede it (i.e., Atkinson, 1973; Collier, 1975; Lieberman, 1967) in demonstrating that the drop in Ps is sufficient to account for F0 declination. At the same time, it refutes the hypotheses of those studies which suggest that declination is the exclusive province of the muscles of the larynx (eg., Fromkin & Ohala, 1968; Vanderslice, 1967) or the myoelastic properties of the vocal folds (Titze, 1980). Moreover, the results of this study show that the stability of the declination derives from the stability of the subglottal pressure. Thus, if we are to address the issue of control in declination, it should be at this level.

It has been known for some time that the respiratory system acts in such a way as to stabilize subglottal pressure (eg., Draper et al., 1960; Mead et al., 1968). The data presented here not only confirm the results of these earlier studies, but provide evidence that this control is dynamic in nature. Furthermore, this stability is maintained even when the system must respond to perturbations in the form of varying airflow requirements, such as those which occur over utterances composed of the syllable /fa/. The finding that the rate of pressure decline is independent of the rate of airflow therefore suggests that subglottal pressure is a highly regulated variable in sentence production and not the passive consequence of unchecked expiratory forces (i.e., lung deflation). Declination, then, could be considered the <u>passive</u> consequence of a <u>regulated</u> subglottal pressure.

In light of these data, the decision as to whether one measures declination by the top- or baseline becomes a far less trivial and arbitrary one. Topline declination would appear to be the composite result of local, momentary increases in F0, which vary with the local inflectional structure of an utterance and the implementation of phonological rules responsible for the realization of tonal variations of this kind. Baseline declination, on the other hand, appears to be insensitive to local inflectional variation, characterizing instead a global 'reference level' of vocal fold vibration over the course of an utterance upon which these local 'perturbations' are superimposed (Cohen et al., 1982; Fujisaki et al., 1982). That the topline and baseline may be acoustically similar or distinct is thus irrelevant in that they are implemented by physiological mechanisms that appear to be differentially involved in sentence intonation (Collier, 1987). Topline declination is most likely the province of the cricothyroid muscle, while baseline declination represents the behavior of a stable subglottal pressure.

On the other hand, as was shown in Experiment 1, both significant cricothyroid and Ps activity accompany utterance-initial syllables. In investigating the physiological correlates of F0 resetting, measures of the activity of these two variables were made in order to determine the degree to which either or both could account for the amount of resetting observed. The results for one subject for whom these data were collected showed that the differences in clause-initial peak F0 could be attributed to concomittant differences in the peak subglottal pressure. The cricothyroid muscle, on the other hand, was shown to be relatively inactive following the clause boundary, both with and without an inspiration, so that its contribution was likely to have been small.

This result is incompatible with Collier's (1987) finding that, following an inspiration at the clause boundary, both F0 and CT activity decrease, while Ps levels remain virtually unchanged. That is, although there is no reason to suppose that Ps must be reset to a higher level following a clause boundary relative to its value just before it, the data for both RC and for the three subjects for whom there are only acoustic data for the two respiratory conditions suggest that the difference in F0 between these conditions is small enough to be accounted for by a difference in Ps, but too small to reflect a significant difference in CT activity. This would also explain why there is so little evidence for resetting in general. In other words, since F0 increases (or decreases) by only a relatively small amount per centimeter of water increase (or decrease) in Ps, the finding that, where it occurs, F0 resetting in two-clause utterances is rather small suggests that the subglottal pressure is the controlling factor. On the other hand, the fact that RC's data showed no significant increases in CT activity following the clause boundary could explain why second clause-initial F0 values do not approximate those found at utterance onset, or even those for the same clauses produced as isolated sentences. Thus, Collier's claims that an increase in Ps "does not explain the observed F0 difference" (p.412) and that "the onset frequency of the declination line is dependent upon the level of CT activity" (p.415) are not supported by the data in the present study. Nor does it become necessary to explain the presence or absence or resetting in terms of a speaker's intention. Rather, these data suggest the conclusion that F0 resetting depends on the level of the subglottal pressure, the setting of which does not depend on speaker control.

CONCLUDING REMARKS

Planning accounts of declination were developed, at least in part, because of the failure of earlier 'passive' models (eg., Lieberman, 1967) to substantiate their predictions. However, despite their alleged failings, the strength of these passive models was that they attempted to relate the ultimate form of the declination to the organization of the physiological mechanisms underlying speech production. Planning accounts, on the other hand, have, in essence, turned things around by assuming that the form of the physiological events is a passive response to a cognitive controller. In other words, implicit in their arguments is the claim, as Fowler (1984) describes it, that the declination takes on a particular form simply because a speaker puts it there in that form. This implies that physiological systems are marshalled in a way that subserves a mental representation of an acoustic end, so that both variability and stability in the declination are seen as evidence of speaker control. Thus, the origins of declination, according to these theories, lie somewhere in the realm of higher cognitive processing.

The results of the experiments in this study have shown that such a formulation is unnecessarily elaborate and, most likely, erroneous. Instead, the origins of declination, as was suggested by the earlier studies, are to be found in the organization of the physiological systems subserving speech. Where this study goes beyond these earlier works is in better characterizing the nature of this organization so that both the variability and stability of the declination can be more readily explained without invoking elaborate look-ahead planning mechanisms and complicated mental computations to be performed by a speaker.

Appendix A1

Experimental Stimuli

Experiment 1

Part 1

L1: Je weet dat jan nadenkt.
L2: Je weet dat jan erover nadenkt te betalen.
L3: Je weet dat jan erover nadenkt ons daarvoor met genoegen te betalen.

Part 2

Subject RC
L1: You knew I did it.
L2: You knew I didn't want you to do it.
L3: You knew I didn't want you to do this job so much ahead of time.

Subject EB
L1: You knew I didn't.
L2: You knew I didn't want you to do it.
L3: You knew I didn't want you to do it so much ahead of time.

Appendix A2

Experimental Stimuli

Experiment 2

Subject RC

S1: When the lawyer called up Reynolds, the plan was disussed.
S2: The lawyer called up Reynolds and the plan was discussed.
S3: When the lawyer called up, Reynolds' plan was discussed.

Subject EB

Two-Clause Utterances: Subordinate-Main Clauses

1-1: When the lawyer called, the plans were discussed.
1-2: When the lawyer called, the plans were discussed at some length.
1-3: When the lawyer called, the plans were discussed at some length on the telephone.

2-1: When the lawyer called Reynolds, the plans were discussed.
2-2: When the lawyer called Reynolds, the plans were discussed at some length.
2-3: When the lawyer called Reynolds, the plans were discussed at some length on the telephone.

3-1: When the lawyer called Reynolds on Monday, the plans were discussed.
3-2: When the lawyer called Reynolds on Monday, the plans were discussed at some length.
3-3: When the lawyer called Reynolds on Monday, the plans were discussed at some length on the telephone.

Two-Clause Utterances: Conjoined Main Clauses

1-1: The lawyer called and the plans were discussed.
1-2: The lawyer called and the plans were discussed at some length.
1-3: The lawyer called and the plans were discussed at some length on the telephone.

2-1: The lawyer called Reynolds and the plans were discussed.
2-2: The lawyer called Reynolds and the plans were discussed at some length.
2-3: The lawyer called Reynolds and the plans were discussed at some length on the telephone.

3-1: The lawyer called Reynolds on Monday and the plans were discussed.
3-2: The lawyer called Reynolds on Monday and the plans were discussed at some length.

3-3: The lawyer called Reynolds on Monday and the plans were discussed at some length on the telephone.

Single Clause Utterances: First Clauses
L1: The lawyer called.
L2: The lawyer called Reynolds.
L3: The lawyer called Reynolds on Monday.

Single Clause Utterances: Second Clauses
L1: The plans were discussed.
L2: The plans were discussed at some length.
L3: The plans were discussed at some length on the telephone.

Subjects BK, JS, KM

Two-Clause Utterances
S1L1: When the lawyer called Reynolds, the plans were discussed.
S1L2: When the lawyer called Reynolds, the plans were discussed at some length.
S1L3: When the lawyer called Reynolds, the plans were discussed at some length on the telephone.

S2L1: In the throes of rebuilding, Paris was destroyed.
S2L2: In the throes of rebuilding, Paris was destroyed in a fire.
S2L3: In the throes of rebuilding, Paris was destroyed in a fire set by arsonists.

Two-Clause Utterances
S1L1: The plans were discussed.
S1L2: The plans were discussed at some length.
S1L3: The plans were discussed at some length on the telephone.

S2L1: Paris was destroyed.
S2L2: Paris was destroyed in a fire.
S2L3: Paris was destroyed in a fire set by arsonists.

Appendix B

Utterance Durations (In Msec)

	Early	SD	Double	SD	Late	SD	
L1	1333(14)	39.0	1337(15)	58.1	1203(14)	72.7	
L2	2053(15)	95.0	2495(14)	304.8	2026(15)	106.4	RC
L3	3353(15)	279.0	3701(15)	555.6	3300(14)	406.2	
Mean	2246(44)	860.6	2511(44)	1051.9	2176(43)	878.1	
L1	959(5)	40.8	985(5)	28.0	890(5)	32.4	
L2	1556(5)	68.2	1642(5)	39.6	1425(6)	51.0	RC English
L3	2665(5)	20.3	2565(5)	26.0	2566(5)	62.8	
Mean	1727(15)	733.0	1731(15)	663.1	1627(16)	702.4	

	Early	SD	Double	SD	Late	SD	Neutral	SD	
L1	812(16)	70.7	884(8)	91.6	803(9)	29.2	775(10)	47.6	
L2	1375(10)	67.9	1443(15)	84.4	1347(3)	76.5	1441(9)	74.7	EB
L3	2283(7)	95.3	2233(18)	127.1	2121(6)	199.8	2348(10)	197.7	
Mean	1490(33)	580.3	1520(41)	544.3	1424(18)	617.1	1521(29)	678.2	

Appendix C

Initial Peak Fundamental Frequency

(In Hertz)

	Early	SD	Double	SD	Late	SD	
L1	139	5.0	138	5.7	118	6.6	
L2	145	8.1	148	4.5	121	3.3	RC
L3	168	5.4	165	5.4	139	4.9	
Mean	151	13.9	150	12.4	126	10.4	

L1	172	4.1	160	2.2	138	8.4	
L2	189	5.6	164	6.7	135	3.3	RC English
L3	190	4.1	178	8.0	130	3.9	
Mean	184	9.5	167	9.9	134	6.1	

	Early	SD	Double	SD	Late	SD	Neut	SD	
L1	278	35.2	273	39.2	186	19.0	180	22.7	
L2	249	28.6	237	36.6	210	2.9	175	11.7	EB
L3	274	21.5	269	37.1	189	10.4	188	11.6	
Mean	267	32.7	260	40.6	195	16.7	181	16.6	

Appendix D

F0 x Length Correlations

	r	p	N	
Dutch	.627	<.001	45	
/ma/	.754	<.001	42	
/fa/	.746	<.001	44	
				RC
Early	.874	<.001	44	
Double	.848	<.001	44	
Late	.809	<.001	43	
Early	.735	=.001	15	
Double	.794	<.001	14	RC English
Late	-.525	<.05	16	
Neutral	.235	>.1	29	
Early	-.086	>.1	33	
Double	.073	>.1	41	EB
Late	.053	>.1	18	

Appendix E

Initial Peak Subglottal Pressure (In Cm-H2O)

	Early	SD	Double	SD	Late	SD	
L1	9.27	.9	9.96	.5	8.44	.7	
L2	10.00	.6	11.14	.7	8.29	.6	RC
L3	10.47	.7	11.71	.9	9.55	1.0	
Mean	9.91	.9	10.94	1.0	8.76	.9	
L1	12.58	.5	10.47	.5	8.76	.7	
L2	13.66	.5	11.36	.5	8.85	.5	RC English
L3	13.28	.7	12.72	.7	9.34	.6	
Mean	13.17	.7	11.52	1.1	8.98	.6	

	Early	SD	Double	SD	Late	SD	Neutral	SD	
L1	12.03(15)	1.7	12.41(7)	.9	9.56(7)	1.9	8.48(8)	1.1	
L2	11.88	1.9	11.49(13)	1.9	-	-	8.88	1.1	EB
L3	12.49	1.3	12.13(17)	1.4	8.73(4)	.9	9.56(9)	.9	
Mean	12.13(32)	1.7	12.01(37)	1.5	9.15(11)	1.8	8.97(26)	1.2	

Appendix F

Ps x Length Correlations

	r	*p*	*N*	
Dutch	.474	<.001	45	*RC*
/ma/	.485	<.001	42	
/fa/	.614	<.001	44	
Early	.471	=.001	44	
Double	.570	<.001	44	
Late	.448	<.01	43	
Early	.308	>.1	15	*RC English*
Double	.853	<.001	14	
Late	.388	>.1	16	
Neutral	.428	<.05	26	*EB*
Early	.133	>.1	32	
Double	.100	>.1	37	
Late	-.280	>.1	11	

Appendix G

Initial Peak Cricothyroid Muscle Activity (in Microvolts)

	Early	SD	Double	SD	Late	SD	
L1	309	73.6	313	63.9	249	26.3	
L2	337	46.6	313	62.9	245	38.1	RC
L3	397	47.9	391	46.7	315	97.7	
Mea	348	66.6	339	67.5	270	70.3	
L1	414	50.2	329	78.5	331	101.0	
L2	437	55.2	397	54.2	257	71.9	RC English
L3	501	20.8	374	63.7	222	79.4	
Mean	451	56.2	367	68.0	270	90.6	

	Early	SD	Double	SD	Late	SD	Neut	SD	
L1	187	21.1	188	9.4	118	50.1	81	22.7	
L2	170	23.2	170	27.1	146	18.5	83	35.1	EB
L3	176	34.6	173	21.5	114	25.1	103	32.8	
Mean	178	25.4	177	22.7	126	39.3	89	32.7	

Appendix H

CT x Length Correlations

	r	p	N	
Dutch	.595	<.001	45	RC
/ma/	.347	<.05	42	
/fa/	.553	<.001	44	
Early	.513	<.001	44	
Double	.558	<.001	44	
Late	.411	<.01	43	
Early	.688	<.01	15	RC English
Double	.085	>.1	14	
Late	-.457	>.05	16	
Neutral	.270	>.1	29	EB
Early	-.181	>.1	33	
Double	-.146	>.1	41	
Late	-.049	>.1	18	

Appendix I

Depth of Initial Inspiration

(In Liters)

	Early	SD	Double	SD	Late	SD	
L1	.96	.3	1.03	.2	1.15	.3	
L2	1.03	.1	1.41	.3	1.23	.2	RC
L3	1.56	.3	1.72	.3	1.50	.3	
Mean	1.18	.4	1.39	.4	1.29	.3	
L1	1.17	.1	1.07	.1	1.07	.2	
L2	1.46	.3	1.01	.2	1.03	.2	RC English
L3	1.32	.1	1.41	.1	1.06	.2	
Mean	1.32	.2	1.16	.2	1.05	.2	

Appendix J

Depth x Length Correlations

	r	*p*	*N*	
Dutch	.529	<.001	45	
/ma/	.745	<.001	42	
/fa/	.671	<.001	44	
				RC
Early	.678	<.001	44	
Double	.707	<.001	44	
Late	.490	<.001	43	
Early	.102	>.1	15	
Double	.753	=.001	14	*RC English*
Late	-.015	>.1	16	

Appendix K

Ps x FO Correlations

	r	p	N	
Dutch	.835	<.001	45	RC
/ma/	.700	<.001	42	
/ta/	.820	<.001	44	
Early	.676	<.001	44	
Double	.768	<.001	44	
Late	.667	<.001	43	
Early	.406	>.1	15	RC English
Double	.900	<.001	14	
Late	.329	>.1	16	
Neutral	.725	<.001	26	EB
Early	.201	>.1	32	
Double	.431	<.01	37	
Late	-.054	>.1	11	

Appendix L

CT x FO Correlations

	r	p	N	
Dutch	.635	<.001	45	RC
/ma/	.651	<.001	42	
/fa/	.733	<.001	44	
Early	.591	<.001	44	
Double	.526	<.001	44	
Late	.511	<.001	43	
Early	.459	>.05	15	RC English
Double	.440	>.1	14	
Late	.128	>.1	16	
Neutral	.786	<.001	29	EB
Early	.527	<.001	33	
Double	.432	<.01	41	
Late	.755	<.001	18	

Appendix M

Ps x CT Correlations

	r	*p*	*N*	
Dutch	.359	<.02	45	*RC*
/ma/	.518	<.001	42	
/fa/	.678	<.001	44	
Early	.511	<.001	44	
Double	.222	>.1	44	
Late	.322	<.05	43	
Early	.059	>.1	15	*RC English*
Double	.282	>.1	14	
Late	-.179	>.1	16	
Neutral	.486	<.02	26	*EB*
Early	-.041	>.1	32	
Double	.249	>.1	37	
Late	.249	>.1	11	

Appendix N

EB Two-Clause Utterance Clause Durations (in Msec)

	Clause 1		Clause 2	
	Mean	SD	Mean	SD
1-1	896	28.09	828	30.04
2-1	1209	67.02	854	42.76
3-1	1625	74.83	881	49.01
1-2	853	54.56	1428	36.48
2-2	1312	176.49	1461	71.34
3-2	1617	98.04	1448	30.35
1-3	809	43.63	1981	32.90
2-3	1309	51.13	2039	39.80
3-3	1599	117.89	1943	35.65

Appendix O

Two-Clause Utterance Clause Durations

(in Msec)

		Clause 1			Clause 2		
		Mean	SD	Syl	Mean	SD	Syl
BK	S1L1	1115	83.7	7	1223	103.6	5
	S1L2	1110	107.6	7	1830	93.8	8
	S1L3	1087	73.9	7	2325	130.7	13
	S2L1	1034	81.2	7	1112	53.0	5
	S2L2	1072	52.1	7	1584	137.6	9
	S2L3	1065	58.8	7	2450	113.1	14
JS	S1L1	1390	43.7		1136	33.6	
	S1L2	1390	29.2		1874	63.1	
	S1L3	1383	33.8		3091	93.3	
	S2L1	1428	37.8		1293	59.5	
	S2L2	1431	44.9		1995	44.5	
	S2L3	1408	36.2		3040	111.8	
KM	S1L1	1037	39.9		890	32.0	
	S1L2	1042	33.9		1455	36.1	
	S1L3	1041	24.3		2211	59.8	
	S2L1	1175	56.0		988	31.5	
	S2L2	1138	49.2		1500	48.6	
	S2L3	1166	50.0		2222	97.4	

REFERENCES

Askenfelt, A., Gauffin, J., Sundberg, J. and Kitzing, P. (1980). A comparison of contact microphone and electroglottograph for the measurement of vocal fundamental frequency. *Journal of Speech & Hearing Research*, *23*, 258-273.

Atkinson, J. E. (1973). *Aspects of intonation in speech: Implications from an experimental study of fundamental frequency*. Unpublished doctoral dissertation, University of Connecticut.

Baer, T. (1979). Reflex activation of laryngeal muscles by sudden induced subglottal pressure changes. *Journal of the Acoustical Society of America*, *65*, 1271-1275.

Baer, T., Gay, T. & Niimi, S. (1976). Control of fundamental frequency, intensity, and register of phonation. *Haskins Laboratories Status Report on Speech Research*, *SR-45/46*, 175-185.

Baer, T., Lofqvist, A. & McGarr, N.S. (1983). Laryngeal vibrations: a comparison between high-speed filming and glottographic techniques. *Journal of the Acoustical Society of America*, *73*, 1304-1308.

Bless, D.M., Hunker, C.J. & Weismer, G. (1981). Comparison of non-invasive methods to obtain chest-wall displacement and aerodynamic measures during speech. *Transcripts of the Tenth Symposium Care of the Professional Voice. Part 1: Instrumentation in Voice Research*. The Julliard School. The Voice Foundation. New York. 43-51.

Boyce, S. & Menn, L. (1979). Peaks vary, endpoints don't: Implications for linguistic theory. *Proceedings from the Fifth Annual Meeting of the Berkeley Linguistics Society*.

Breckenridge, J. (1977). Declination as a phonological process. *Bell System Technical Memorandum*, Bell Laboratories, Murray Hill, NJ.

Bruce, G. (1982). Developing a Swedish intonation model. *Working Papers*, *22*, Lund University, Department of Linguistics, 51-116.

Cohen, A., Collier, R., & 't Hart, J. (1982). Declination: Construct or intrinsic feature of pitch? *Phonetica*, *39*, 254-273.

Collier, R. (1974). Laryngeal muscle activity, subglottal air pressure, and the control of pitch in speech. *Haskins Laboratories Status Report on Speech Research*, *SR-39/40*, 137-170.

Collier, R. (1975). Physiological correlates of intonation patterns. *Journal of the Acoustical Society of America*, *58*, 249-255.

Collier, R. (1987). F0 declination: The control of its setting, resetting and slope. In T. Baer, C. Sasaki & K. Harris (Eds.), *Laryngeal Function in Phonation and Respiration*, Little, Brown and Company, 403-421.

Cooper, W.E. & Sorenson, J.M. (1977). Fundamental frequency contours at syntactic boundaries. *Journal of the Acoustical Society of America*, *62*, 683-692.

Cooper, W. E. & Sorenson, J. M. (1981). *Fundamental frequency in sentence production*. New York: Springer-Verlag.

Cranen, B. & Boves, L. (1985). Pressure measurements during speech production using miniature pressure transducers: Impact on models for speech production. *Journal of the Acoustical Society of America*, *77*, 1543-1551.

Draper, M.H., Ladefoged, P. and Whitteridge, D. (1960). Expiratory pressures during speech. *British Medical Journal*, *1*, 1837-1843.

Eady, S.J. & Cooper, W.E. (1986). Speech intonation and focus in matched statements and questions. *Journal of the Acoustical Society of America*, *80*, 402-415.

Erickson, D. & Atkinson, J. E. (1976). The function of strap muscles in speech. *Haskins Laboratories Status Report on Speech Research*, *SR-45/46*, 205-210.

Erickson, D., Baer, T. & Harris, K.S. (1983). The role of the strap muscles in pitch lowering. In D.M. Bless & J.H. Abbs (Eds.), *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, College-Hill Press, San Diego, 279-285.

Fourcin, A.J. (1974). Laryngographic examination of vocal fold vibration. In B. Wyke (Ed.), *Ventilatory and Phonatory Control Systems*, New York, 315-333.

Fowler, C.A. (1984). Current perspectives on language and speech production. In R. Daniloff (Ed.), *Recent Advances in Speech, Hearing and Language*, *Vol. 3*, College-Hill Press.

Fromkin, V. A. & Ohala, J. (1968). Laryngeal control and a model of Speech Production. *UCLA Working Papers in Phonetics*, *10*, 98-110.

Fry, D.B. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America*, *27*, 765-768.

Fujisaki, H. & Hirose, K. (1982). Modeling the dynamic characteristics of voice fundamental frequency with applications to analysis and synthesis of intonation. *Preprints of papers*, Working Group on Intonation, The XIIIth International Congress of Linguistics, Tokyo, 57-70.

Fujisaki, H., Hirose, K., & Ohta, K. (1979). Acoustic features of the fundamental frequency contours of declarative sentences in Japanese. Ann. Bull. RILP, 13, 163-173.

Garding, E., Fujimaura, O. & Hirose, H. (1970). Laryngeal control of Swedish word tone: A preliminary report of an EMG study. Annual Bulletin of the Research Institute of Logopedics and Phoniatrics, 4, Tokyo, 45-54.

Harris, K. S. (1981). Electromyography as a technique for laryngeal investigation. In C. L. Ludlow & M. O. Hart (Eds.), ASHA Reports: Proceedings of the Conference on the Assessment of Vocal Pathology, 11, 70-86.

Hirano, M. (1974). Morphological structure of the vocal cord as a vibrator and its variations. Folia Phoniatrica, 26, 89-94.

Hirano, M., Ohala, J. & Vennard, W. (1969). The function of laryngeal muscles in regulating fundamental frequency and intensity of phonation. Journal of Speech & Hearing Research, 12, 616-628.

Hirano, M., Vennard, W. & Ohala, J. (1970). Regulation of register, pitch, and intensity of voice. Folia Phoniatrica, 22, 1-20.

Hirose, H. (1971). Electromyography of the articulatory muscles: Current instrumentation and techniques. Haskins Laboratories Status Report on Speech Research, SR-25/26, 73-86.

Hixon, T. J., Klatt, D. H. & Mead, J. (1971). Influence of forced transglottal pressure changes on vocal fundamental frequency. Journal of the Acoustical Society of America, 49, 105.

Honda, K. (1983). Relationship between pitch control and vowel articulation. Haskins Laboratories Status Report on Speech Research, SR-73, 269-282.

Horii, Y. & Cooke, P.A. (1978). Some airflow, volume, and duration characteristics of oral reading. Journal of Speech & Hearing Research, 21, 470-481.

Ishizaka, H. & Matsudaira, M. (1972). Fluid mechanical considerations of vocal cord vibration. Monograph No. 8, Speech Communication Research Laboratory, Santa Barbara.

Isshiki, N. (1964). Regulatory mechanism of voice intensity variation. Jounal of Speech & Hearing Research, 7, 17-29.

Isshiki, N. (1965). Vocal intensity and air flow rate. Folia

Phoniatrica, 17, 92-104.

Kelso, J.A.S., Vatikios-Bateson, E., Saltzman, E.L. & Kay, B. (1985). A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling. Journal of the Acoustical Society of America, 77, 266-280.

Koyama, T., Harvey, J.E. & Ogura, J.H. (1971). Mechanics of voice production. II. Regulation of pitch. The Laryngoscope, 81, 47-65.

Koymama, T., Kawasaki, M. & Ogura, J.H. (1969). Mechanics of voice production. I. Regulation of vocal intensity. The Laryngoscope, 79, 337-354.

Kutik, E. J., Cooper, W. E. & Boyce, S. (1983). Declination of fundamental frequency in speakers' production of parenthetical and main clauses. Journal of the Acoustical Society of America, 73, 1723-1730.

Ladefoged, P. (1962). Sub-glottal activity during speech. Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki, 73-91.

Ladefoged, P. (1963). Some physiological parameters in speech. Language and Speech, 6, 109-119.

Larkey, L. S. (1983). Reiterant speech: An acoustic and perceptual evaluation. Journal of the Acoustical Society of America, 73, 1337-1345.

Leroy, L. (1984). The psychological reality of fundamental frequency declination. Antwerp Papers in Linguistics, 40, University of Antwerp, Wilrijk, Belgium.

Liberman, M.Y. (1975). The Intonational System of English, Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Liberman, M. & Pierrehumbert, J. (1982). Intonational invariance under changes in pitch range and length. Bell Laboratories Internal Technical Memorandum.

Liberman, M. & Streeter, L. A. (1978). Use of nonsense-syllable mimicry in the study of prosodic phenomena. Journal of the Acoustical Society of America, 63, 231-233.

Lieberman, P. (1967). Intonation, perception and language. Research Monograph No. 38. Cambridge: MIT Press.

Lieberman, P., Knudson, R. & Mead, J. (1969). Determination of the rate of change of fundamental frequency with respect to subglottal pressure during sustained phonation. Journal of the Acoustical Society of America, 45, 1537-1543.

Lieberman, P., Sawashima, M., Harris, K.S. & Gay, T. (1970). The articulatory implementation of the breath-group and prominence: Crico-thyroid muscular activity in intonation. Language, 46, 312-327.

Maeda, S. (1976). A Characterization of American English Intonation. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Mead, J., Bouhuys, A. & Proctor, D.F. (1968). Mechanisms of generating subglottic pressures. In A. Bouhuys (Ed.), Sound Production in Man, Annals of the New York Academy of Sciences, 177-181.

Monsen, R.B., Engebretson, A.M. & Vemula, N.R. (1978). Indirect assessment of the contribution of subglottal air pressure and vocal fold tension to changes of fundamental frequency in English. Journal of the Acoustical Society of America, 64, 65-80.

Ohala, J.J. (1972) How is pitch lowered? Journal of the Acoustical Society of America, 52, 124 (A).

Ohala, J.J. (1978) Production of tone. In V. A. Fromkin (Ed.), Tone: a linguistic survey. New York: Academic Press, 5-39.

Ohala, J., Hirano, M. & Vennard, W. (1968). An electromyographic study of laryngeal activity in speech and singing. Paper read at the 6th International Congress on Acoustics, Tokyo.

Ohala, J. & Hirose, H. (1970). The function of the sternohyoid muscle in speech. Annual Bulletin, 4, Research Institute of Logopedics and Phoniatrics. University of Tokyo, 41-44.

Pierrehumbert, J. (1979). The perception of fundamental frequency declination. Journal of the Acoustical Society of America, 66, 363-369.

Scherer, R.C., Druker, D. & Titze, I.R. (1987). Electroglottography and direct measurement of vocal fold contact area. Paper presented at the Fifth Conference on Vocal Fold Physiology, Tokyo, Japan.

Shattuck-Huffnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. MacNeilage (Ed.), The Production of Speech. New York: Springer-Verlag.

Shimizu, K. & Dantsuji, M. (1981). A study on the perception of fundamental frequency delination by Japanese speakers. Studia Phonologica XV, Institution for Phonetic Sciences, University of Tokyo, 18-30.

Shipp, T. (1975). Vertical laryngeal positions during continuous and discrete vocal frequency change. Journal of Speech & Hearing Research, 18, 707-718.

Shipp, T. & Izdebski, K. (1975). Vocal frequency and vertical larynx positioning by singers and nonsingers. Journal of the Acoustical Society of America, 58, 1104-1106.

Shipp, T. & McGlone, R.E. (1971). Laryngeal dynamics associated with voice frequency change. Journal of Speech & Hearing Research, 14, 761-768.

Simada & Hirose. (1970). The function of the laryngeal muscles in respect to the word accent distinction. Annual Report of the Research Institute of Logopedics and Phoniatrics, 4, 27-40.

Sonninen, A. (1968). The external frame function in the control of pitch in the human voice. Annals of the New York Academy of Sciences, Article 1, 68-90.

Sorenson, J.M. & Cooper, W.E. (1980). Syntactic coding of fundamental frequency in speech production. In R.A. Cole (Ed.), Perception and Production of Fluent Speech. Erlbaum, Hillsdale, N.J.

Sternberg, S., Wright, C., Knoll, R. & Monsell, S. (1980). Motor programs in rapid speech: additional evidence. In R.A. Cole (Ed.), Perception and Production of Fluent Speech. Erlbaum, Hillsdale, N.J.

Stetson, R.H. (1951). Motor Phonetics: A study of speech movements in action. North Holland, Amsterdam.

Stevens, K.N., Kalikow, D.N. & Willemain, T.R. (1975). A miniature accelerometer for detecting glottal waveforms and nasalization. Journal of Speech & Hearing Research, 18, 594-599.

't Hart, J. (1979). Explorations in automatic stylization of FO contours. IPO Annual Progress Report, 14, 61-65.

Thorsen, N. (1979a). Intonation contours in declarative sentences of varying length in ASC Danish. Ann. Rep. Inst. Phonet., 13, University of Copenhagen, 1-7.

Thorsen, N. (1979b). Lexical stress, emphasis for contrast, and sentence intonation in advanced standard Copenhagen Danish. Proceedings of the Ninth International Congress of Phonetic Sciences, Vol.2, 417-423.

Thorsen, N. (1979c). Lexical stress, emphasis for contrast, and sentence intonation in advanced standard Copenhagen Danish. Ann. Rep. Inst. Phonet., 13, University of Copenhagen, 59-85.

Thorsen, N.G. (1986). Sentence intonation in textual context - Supplementary data. Journal of the Acoustical Society of America, 80, 1041-1047.

Titze, I.R. (1980). Comments on the myoelastic-aerodynamic theory of phonation. Journal of Speech & Hearing Research, 23, 495-510.

Titze, I.R. & Durham, P.L. (1987). Passive mechanisms influencing fundamental frequency control. In T. Baer, C. Sasaki & K. Harris (Eds.), Laryngeal Function in Phonation and Respiration, Little, Brown & Co., Boston, 304-319.

Umeda, N. (1982). "F0 declination" is situation dependent. Journal of Phonetics, 10, 279-290.

van den Berg, J. (1958). Myoelastic-aerodynamic theory of voice production. Journal of Speech & Hearing Research, 1, 227-243.

Vanderslice, R. (1967). Larynx vs. Lungs: Cricothyrometer data reflecting some recent claims concerning intonation and "archetypality". UCLA Working Papers in Phonetics, 7, 69-79.

Williams, F. (1986). Reasoning with Statistics: How to read quantitative research. Holt, Rinehart & Winston, New York.

Yates, F.E. (1982). Outline of a physical theory of physiological systems. Canadian Journal of Physiology and Pharmacology, 60, 217-248.