77-4266

ERICKSON, Donna M., 1944-
A PHYSIOLOGICAL ANALYSIS OF THE TONES OF THAI.

The University of Connecticut, Ph.D., 1976
Language, linguistics

Xerox University Microfilms, Ann Arbor, Michigan 48106

A PHYSIOLOGICAL ANALYSIS OF THE TONES OF THAI

Donna M. Erickson, Ph.D.

The University of Connecticut, 1976

This study examines laryngeal muscle activity underlying the lexically contrastive pitch patterns of the distinctive tones of Central Thai.

The physiological interest of the study lies in the examination of the laryngeal control of fundamental frequency (F_0) of vocal fold vibration. Prior electromyographic (EMG) studies with speech report cricothyroid control of F_0 rises and falls and association of strap muscle activity with low F_0. This study of Thai tones, in which F_0 rises and falls are lexically determined examines more carefully the nature of strap muscle activity: whether the strap muscles are active or passive participants in F_0 lowering. The results of the investigation indicate that the cricothyroid controls both rises and falls in F_0; the strap muscles appear to actively lower the F_0 below a certain point in the speaker's mid voice range.

The linguistic interest of the study lies in the examination of a possibly invariant relationship between patterns of laryngeal muscle activity and tone features in the phonology. The tones of Thai have been described phonologically either as syllabic units (where the tonal contour is associated with the syllable as a whole), or as sequences of

segmental units (where tonal feature matrices are assigned to the vowels or vowel-like segments of the syllable). There is no phonological evidence to conclusively support either approach. The focus of the investigation is the analysis of the dynamic tones, those tones which show sharp rises or falls in F_0: whether these are to be treated as sequences of features of high and low tones or as features of unit tones with changing F_0 patterns. Acoustic data can be interpreted primarily to support a syllabic approach: the dynamic tones show distinct time-varying F_0 contours which encompass the entire syllable. The physiological data, on the other hand, can be interpreted primarily to support a segmental approach: the dynamic tones are analyzed as sequences of discrete occurrences of a reciprocal patterning of the cricothyroid and strap muscles. Herein is evidence for a one to one mapping of laryngeal muscle activity patterns onto the segmentally based tonal features [High] and [Low]. These findings are related to laryngeal features which describe the states of vocal fold tension underlying the pitch levels of the tones.

A final question examined in this study relates to the theory of tonogenesis which states that the low F_0 of voiced stops brought about low tones and the high F_0 of voiceless stops, high tones. The question examined is whether differential activity of the F_0 raising and lowering muscles is found for voiceless and voiced stops. The results show that initial F_0 differences between voiced and voiceless consonants support the tonogenesis theory but no laryngeal muscle activity differences correspond with the F_0 differences.

A PHYSIOLOGICAL ANALYSIS OF THE TONES OF THAI

Donna M. Erickson, B.A., M.A.

Ohio State University, 1966
University of Michigan, 1968

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at

The University of Connecticut

1976

APPROVAL PAGE

Doctor of Philosophy Dissertation

A PHYSIOLOGICAL ANALYSIS OF THE TONES OF THAI

Presented by

Donna M. Erickson, A.B., A.M.

Major Adviser Arthur S Abramson

Associate Adviser Thomas Gay

Associate Adviser Ignatius G Mattingly

The University of Connecticut

1976

ACKNOWLEDGMENTS

I take this opportunity to formally thank the following people and institutions which have provided me with the support and equipment necessary to carry out this study. Dr. Franklin S. Cooper, former president of Haskins Laboratories, was largely responsible for the creation of the Haskins EMG processing facilities. Dr. Alvin M. Liberman, current president and director of the Laboratories, provided the continuity of facilities necessary for the completion of the work. Dr. Katherine S. Harris, director of the EMG project, made possible the basic facilities which made this work feasible. The National Institute of Dental Research provides funding of the Haskins speech electromyographic project.[1] The Dental Research Center at the University of North Carolina[2]--specifically through the cooperation of George Allen--made the acoustic analysis system available. The staff of the Central Institute of English Language, Bangkok, Thailand--especially, Mrs. Mayuri Sukwiwat, director, and Mr. Jimmy Harris and Dr. Richard B. Noss, provided the facilities for the portion of the research which was performed in Thailand. The work in Thailand was supported by a grant from the National Science Foundation,[3] and the National Research Council of Thailand endorsed my year and one-half stay in the country.

[1] The major portion of this investigation was supported by PHS Research Grant No. DE01774 from the National Institute of Dental Research and in part by General Research Support Grant No. RR05596 from the General Research Support Branch of the National Institutes of Health.

[2] A portion of this investigation pertaining to the acoustic analysis of the study was supported by PHS Research Grant No. DE02668 from the National Institute of Dental Research and in part by General Research Support Grant No. RR5333 from the General Research Support Branch of the National Institutes of Health.

[3] NSF Dissertation Research Grant No. GS35200.

It is inevitable that many other people have necessarily contributed in various ways to a project which has had such a long gestation period. I am unhappy that I cannot give more full and adequate recognition to all the individuals who have contributed in so many ways. However, I realize that this acknowledgment is merely a formal recognition of feelings which I cannot capture on paper. I hope these people will enjoy the memory of their contributions to this work.

Don Smith (my husband) and the U. S. Army Special Forces Thailand conspired to interest me in the Thai language initially.

The original suggestion to investigate Thai tones from the perspective of laryngeal physiology came from Kathy Harris and Arthur Abramson. Kathy contributed to the ongoing work in many ways and was invaluable in the process of translating the data into English. Her encouragement maintained my interest through several episodes of personal despair.

Arthur Abramson, as my major advisor was involved in many aspects of the work. From helping with the original design and in making contact with the experimental subjects, to his suggestions on the final version of the thesis, he was present in one way or another throughout. Our time together in Thailand (1973-1974) was especially fruitful.

My other advisers, Ignatius Mattingly and Tom Gay, have contributed their own special skills--Ignatius his understanding of broader viewpoints and Tom his knowledge of EMG research. To all my advisers I owe an especial acknowledgment. Although they were unable to agree with each other on many issues, this was no handicap since it forced me to think independently and to insist on synthesis of conflicting approaches and viewpoints.

Dear people, whom I love so much and have helped me when I needed you, I thank you. I especially thank my gifted friend Sirmsree Sevatamorn who helped me with the language, and her family who showed me much of Thai culture and family values. I also thank Sirmsree for her valiant and courageous performance as subject for the EMG experiments, along with Phadhadej Dhamcharee, Chaisagna Taeratanachai, and Montri Jaiprasert who also underwent the same ordeal for the sake of science. May their larynges be forever blessed. Thanks go to Praiphon and Rachada Koomsup who gave help, hospitality, and insights into Thai phonology. I offer special thanks to Ms. Kalaya Chitrabongs of the Faculty of English at the University of Chiangmai, Thailand, for helping me collect data in Thailand, and to the staff and students at the English Language Center in Bangkok who provided assistance in various ways.

I thank each of you others who have helped during the arduous task of preparing the final manuscript--especially dear Agnes McKeon (may you never have to spend another Easter weekend drawing figures in a hot laboratory), Tina LaColla for your excellent typing in spite of my carelessness, Dorie Baker for your editing job (may your headaches vanish forever), and Donna Silverio for your good-humored and painstaking figure production. By no means least of all, I thank Alice Dadourian who at a critical point stated, "We'll get this done on time if I have to do all of the typing myself." Al Liberman also deserves thanks for his support in the final preparations of the work and for his faith in the necessity of "finishing."

I thank Jane Gaitenby for her proofreading of the final manuscript, Susan Gray for her editing of my sometimes idiosyncratic prose style, and Robin Kneubuhl for her everreadiness to type bits and pieces of the prefinal manuscript.

Thanks also is extended to Freddie Bell-Berti and Tom Baer for their helpful and patient suggestions and explanations relating to electromyographic research and laryngeal muscle functions. Jim Hirose and Seiji Niimi have gained my gratitude--Jim for his relatively painless and accurate insertions of the electrodes as well as encouragement and critical comments on early drafts; Seiji for his cheerful personality and expert explanations of laryngeal physiology.

The final thanks goes to my unofficial advisory committee, Jim Atkinson, Terry Halwes, and Gary Kuhn, who at various stages of the production encouraged me in my endeavor by offering concrete guidance with regard to both overall focus and specific details, and without whose help this thesis would not have materialized. I am also grateful to Mark Liberman and Jack Gandour for valuable insight concerning the linguistic aspects of the study.

There are many others to be thanked, including my obnoxious and loving husband, who absolutely insisted that I complete this thesis quickly. I also thank Jean Jessop Sutherland and Ray Walker for their help in developing a sense of my own worth and creativity. In addition, my parents, siblings, and mother-in-law are remembered for their distant but consistent emotional support in this endeavor.

TABLE OF CONTENTS

Page

List of Tables

List of Figures

CHAPTER I: INTRODUCTION

Like many other Southeast Asian languages, Thai is a tone language. In a tone language, as narrowly defined, the phonological specification of every word includes a distinctive pitch pattern for each syllable (Pike, 1948). That is, the tones are lexically contrastive. Such languages differ both in the nature of their pitch patterns and in the total number of tones. Tone languages do not differ from other languages in requiring the speaker's voice to rise and fall consonant with the rules of sentence intonation and emotional demands. Nevertheless, sentence intonation does not normally interfere with the tonal distinctions, although under certain conditions--for example, weak stress in Thai--certain tonal contrasts will collapse (Noss, 1964:28; Hiranburana, 1972).

The primary acoustic realization of tones is the fundamental frequency (F_0), which maintains a typical, relative contour for each tone with reference to the speaker's voice range (Abramson, 1962; Chiang, 1967; Howie, 1976). The F_0 contours are determined by variation in the rate of vocal fold vibration, which is itself regulated by the laryngeal muscles as well as by subglottal air pressure (Van den Berg, 1958).

Some interesting physiological and linguistic questions arise in connection with tone languages; in particular, what is the nature of the laryngeal mechanisms underlying the tonal contrasts, and what is the relation of the physiological data to representation of the tones in the phonology? This study examines these questions in connection with the tones of Central Thai. Specifically, the purposes of this study are the following:

1. To examine the laryngeal control of F_0 in Thai tones to see if definite patterns of laryngeal muscle activity can be identified that invariantly correspond to the various tone patterns previously described.

A tone language is especially interesting in this regard because the speaker must produce each syllable in a string with an intrinsic tonal specification as part of its makeup. In addition to the overall movements of F_0 for sentence intonation, F_0 must also vary from moment to moment in response to the phonetic instructions for the production of each morpheme. This is somewhat similar to the situation for lexical stress in English; but in Thai the phenomenon is more complex (five distinctive pitch contours are employed) and more pervasive (every syllable is involved). This investigation can provide insight into the nature of the laryngeal control of F_0 generally.

2. To determine whether such invariant patterns, if found, provide evidence relevant to phonological analysis of the tonal structure of Thai; specifically, to see whether F_0 contours in tones can be rationalized by a system of laryngeal features in an abstract phonological description, or whether tones must be described as specific unanalyzable F_0 contours (or laryngeal gestures) with no underlying structure.

3. To compare the acoustic and electromyographic (EMG) properties of tones with corresponding aspects of the production of stop consonants, with the intention of providing evidence relevant to a historical

description of the emergence of the tonal and consonantal systems of Central Thai. An old theory asserts that voiceless initial consonants gave rise to high tones, and voiced initial consonants to low tones (Maspéro, 1911), and that the tones then interacted with the consonants to change the consonantal inventory. The stop consonants of Thai have been shown to be differentiated by their initial F_0 characteristics (e.g., Erickson, 1975a). This study will seek corresponding patterns of laryngeal muscle activity since such differences in motor control patterns may be a necessary condition for the hypothesized mechanism of tonogenesis.

Pursuant to these questions, EMG and acoustic measurements were made on productions of four native speakers of Central Thai. The laryngeal muscles examined were those found to be operant in F_0 control, specifically the cricothyroid, the vocalis, and the strap muscles: the sternohyoid, sternothyroid, and thyrohyoid. The samples examined were the nine syllable types /baa, bii, buu, paa, pii, puu phaa, phii, phuu/ as spoken on the five tones, yielding 45 utterance tokens. F_0 contours were analyzed with the University of North Carolina Dental Research Center Phonetic Analysis System, an automatic pitch extraction system. Electromyographic data were analyzed with the Haskins Laboratories computerized EMG processing system. Details of these procedures are given later. At this point, more detailed background of the problem is given.

CHAPTER II: LINGUISTIC BACKGROUND

Phonologic, acoustic, and historical aspects of the tones are discussed in this chapter as background pertinent to the EMG study.

Part 1: Phonological Background

The phonological treatise discusses (A) the types of tones and types of phonological approaches (syllabic and segmental) to the analysis of tone languages in general, and leads to (B) the tonal system and phonological arguments for syllabic and segmental representation of the tones of Thai in particular. Also, (C) feature systems representative of both syllabic and segmental approaches are presented and (D) the implications of syllabically and segmentally based feature systems with regard to handling the temporal quality of dynamic tones are discussed in detail. In particular, the relevance of EMG investigation to the linguistic goal of deriving phonetic features is considered.

A. Phonological descriptions of tone languages. Generally speaking, languages referred to as tone languages are those in which contrastive "tonemes" differentiate lexical meanings. Two types of tones are usually found in these languages: "Level Tonemes" and "Gliding Tonemes" (Pike, 1948). The former are defined as tones "...in which, within the limits of perception, the pitch of the syllable does not rise or fall during its production." With a gliding toneme "...during the pronunciation of the syllable on which it occurs there is a perceptible rise or fall, or some combination of rise and fall, such as rising-falling or falling-rising" (Pike, 1948:5). The terms "level and contour" or "static and dynamic" have also been used to refer to these tone types (Abramson, 1962:9). The latter terms will be employed in this study. Acoustic analysis of these tone types reveals patterns of relatively level F_0 for static tones and relatively kinetic F_0 for dynamic tones (e.g., Abramson, 1962; Chiang, 1967; Howie, 1976). The degree to which (particularly in Thai) static tones have level F_0 and dynamic tones kinetic F_0, is discussed in Part 2 (Acoustic Background) of this chapter.

In general, there are two approaches to handling tones in the phonology: in one approach tones are associated with the syllable (i.e., the domain of the tone is the syllable) and the tonal inventory is the set of tonal patterns that differentiate words from one another (Pike, 1948). In the other approach, tones are associated with segments within the syllable (i.e., the domain of the tone is the segment) and the tonal inventory is the number of minimally distinctive tonal patterns permissible in a segment within the syllable. In this case, the dynamic tones are treated as sequences of static tones. The traditional approach for Southeast Asian languages has been to treat the tones as syllabic[1] units; in West African languages, however, the practice is to consider tones as segmental units. The reason given for the latter is that a segmental description can handle such phenomena as "downstep" (whereby high tones are lowered when they are preceded by low tones), as well as "tone copying" (wherein the immediately preceding tone is "copied onto an inherently toneless element") (Leben, 1973:139) which

[1]The term "syllabic" is used here in the same way as the term "suprasegmental," which appears in Leben (1973).

frequently occur in West African tone languages. An example of tone copying handled segmentally is given by Leben (1973:141): "When the immediately preceding element has what might be described as a contour tone, such as Mende $\frac{\text{mbu}}{\text{HL}}$ or $\frac{\text{mba}}{\text{LH}}$, the tone copied is not a falling or a rising tone, but rather the last level tone of the sequences HL or LH." Thus, the segmental approach is needed here to describe a tone copying phenomenon occurring with dynamic tones that could not be explained if the dynamic tones were syllabic units. Southeast Asian languages such as Thai or Burmese, on the other hand, have no such phenomena as tone copying or downstep, which would argue strongly for a segmental treatment of the tones, and these languages are generally treated with a syllabic description.

There are certain languages for which it is claimed that dynamic tones must be treated as syllabic units in the underlying structure and cannot be decomposed into sequences of level-tone segments. Such a case is the West African language Kru (Elimelech, 1974). This language is posited with the four tones--high, low, rising, and falling--and is subject to downdrift phenomena. It is argued that rising and falling tones must be represented as such in the phonology in order to account for tone sandhi with a sequence of rising tones; namely, that each subsequent elevation of a rising tone begins on a pitch identical to that ending the preceding tone (resulting in a successively higher and higher pitch in some instances). Were the rising tone derived from a sequence of low and high tones, an inaccurate description of tonal sequence would result. (See Elimelech, 1974, for details.)

Another language thought to require dynamic tones in its underlying structure in order to explain tone sandhi phenomena, is the

Tai[2] dialect of Lue spoken in Yunnan, China (Gandour, 1975). A general survey of the problems concerned with phonological representation of tones, in which these particular languages are discussed in greater detail, is presented by Fromkin (1974).

B. Description of tones of Central Thai. Central Thai has five lexically contrastive tones: mid, low, high, falling, and rising (Haas, 1956; Henderson, 1949; Abramson, 1962). The mid, high, and low tones are static or level; their F_0 trajectories are relatively flat. In contrast, the falling and rising tones are contour or dynamic tones; during these tones F_0 changes direction sharply (Abramson, 1962). A sixth tone, a high checked tone with some glottal constriction toward the end, has been described by Noss (1964). It would appear, however, that in a grammar of Thai that does not require an autonomous phonology, this tone should be treated as a predictable variant of the high tone, thus leaving a system of only five tones. This view is accepted here.

In some cases in Thai all five different tones may be contained on one syllable. An example is the syllable [na:], which may mean "field," "custard apple," "mother's younger sibling," "face," or "thick," depending on the tone of the syllable. Although the full system consists of five tones, there are contextual restrictions on the occurrence of some. For example, any syllable ending with a stop consonant normally takes only two different tones, though exceptions are found in a few loan words. Of course, even if there are no known restrictions to the occurrence of all five tones on a particular syllable, words with one or more of the theoretically possible tones may not occur in the lexicon.

[2]The term "Tai" refers to a group of genetically related languages, among which is the Central Thai dialect of this study.

There are basically two schools of thought concerning the representation of tones of Central Thai in the phonology and specifically, the dynamic tones. The traditional approach has been to treat the dynamic tones as unit phonemic tones (e.g., Hass, 1956; Henderson, 1949; Abramson, 1962; Noss, 1964; Warotamasikkhadit, 1967). A second approach (Trager, 1957; Leben, 1971; Gandour, 1974a) has been to treat the dynamic tones as sequences of two static tones: the falling tone is a high tone followed by a low tone; the rising tone, a low tone followed by a high tone.

In traditional analyses the tones are characterized by distinct pitch patterns that function to distinguish minimally one word from another, and are part of the phonological specification of every word in the language. (The example of [na:] has been given above.) The domain of the tone is the syllable.

Syllabic descriptions of tones tend to be couched in phonological terms that antedate generative phonology; an exception is the system proposed by Warotamasikkhadit (1967) in which he specifies three distinctive features [High], [Low], and [Falling] to describe the tones.[3] This system is discussed more thoroughly below. To a large extent, there are no strong phonological arguments for the syllabic approach of the type given by generative phonologists; however, support for such a description is manifest in the native (literate) use of the language and is provided for in writing. For instance, there is the five-way

[3]Even here, Warotamasikkhadit does not discuss the issue of syllabic versus segmental treatment of the tones per se; the possibility of a syllabically based feature system is made apparent in the course of his exposition.

lexical contrast in the language. Also, the Thai writing system specifies five tones that match the lexical contrasts. A further reason for viewing the five tones of Thai as syllable elements derives from acoustic evidence: the F_0 contour shows an everchanging pattern over time that is not segmented into discrete units. (This is discussed later in Part 1, Section C, of this chapter.)

Contrasting with this approach is the segmental analysis of the Thai tones in which the five lexically contrastive tones are further analyzed into three phonemic tones, which can be combined to yield the full complement of five tones. In this approach the number of tones is determined by the number of minimally distinctive F_0 patterns in a specific environment within the syllable, rather than as a function of the number of minimally contrasting words they delineate. In this sense then, tones can be viewed as segmental features, their domain being limited to a single segment within the syllable. The domain of the tone in this three-tone description of Thai is the "single" vowel, semivowel, or final nasal, rather than syllable or word, as with the traditional five-tone description.

Trager (1957), for reasons of economy of inventory within a structuralist framework, posits three pitch phonemes: high, mid, and low; Leben (1971) and Gandour (1974a), working with the premises of generative phonology, speak of three phonemic tones: high, mid, and low. In addition, Gandour proposes a distinctive feature system that describes these three tones in terms of two binary features [High] and [Mid]. Gandour's system is discussed in more detail in Part 1, Section C, of this chapter.

The strongest evidence for a segmental description rests in the phonotactic argument, which relates vowel length and distribution of tones in the language (Trager, 1957; Gandour, 1974a). Specifically, Thai has a phonemic vowel length distinction; moreover, the dynamic tones cannot occur on short vowels, unless the vowel is followed by a nasal consonant or a semivowel. (See Table 1, which is from Gandour, 1974a.)

TABLE 1: Distribution of lexical tones in different types of syllable structures in Thai.

Syllable structures		Lexical tones: Mid	Low	High	Falling	Rising
(1) CVV		X	X	X	X	X
(2) CV(V)C_f	C_f = mnŋwy	X	X	X	X	X
(3) CVVC_f	C_f = ptk	0	X	0	X	0
(4) CVC_f	C_f = ptk?	0	X	X	0	0

The correlation between distributions of vowel lengths and tone types can be expressed quite simply if the tones and vowels are both analyzed segmentally; the dynamic tones can be handled as sequences of high and low tones (such a sequence could only occur on a syllable with two vowellike segments); the level tones on the long vowels may be analyzed as a sequence of occurrences of the same tone. This would yield the following analysis:

High $\acute{V}$: → High-High $\acute{V}\acute{V}$

Low $\grave{V}$: → Low-Low $\grave{V}\grave{V}$

Mid $\bar{V}$: → Mid-Mid $\bar{V}\bar{V}$

Falling $\hat{V}$: → High-Low $\acute{V}\grave{V}$

Rising $\check{V}$: → Low-High $\grave{V}\acute{V}$

An assumption underlying this analysis is that vowel length is also a segmental feature. Some investigators question a segmental treatment of vowel length in Thai; for instance, Abramson (1974c) argues that the concept of double vowels should rest upon a phonetic basis, either in the form of evidence of rearticulation or perceptual processing of individual vowel segments; he finds evidence of neither. This point is discussed later in regard to the relation of the phonological representation to the speech signal (in Part 1, Section C, of this chapter).

There is, however, an independent rationale for transcribing vowel length with double vowels (e.g., /aa/), namely, to maintain parallelism with the vowel clusters /ia, ɨa, ua/. This practice is seen in several linguistic descriptions of Thai (Haas, 1956; Abramson, 1962; Noss, 1964). From a phonotactic point of view (e.g., Pike, 1948) this argument is convincing, unless one can justify treating the "clusters" as vowel phonemes of changing quality. This is analogous to the phonological question of whether to treat a diphthong as vowel + vowel, vowel + glide, or a single vowel phoneme of changing quality.

The argument for segmental analysis is somewhat circular. If, however, vowel length and tones are analyzed as segmental features, an interesting generalization can be made about vowel and tone occurrences: dynamic tones occur only with the double segments, that is, double vowels, including vowel clusters, or single vowels followed by either a nasal or semivowel (cf. Table 1). The absence of falling and rising tones is represented by unit contour features.

Abramson[4] argues, however, that the ability of a segmental analysis to capture this generalization is offset by a failure to account for

[4] Arthur Abramson, 1976: personal communication.

the nonoccurrence of the mid, high, and rising tones on double vowels ending with voiceless stop consonants.

Other arguments for segmental analysis presented in the literature are less convincing. Among these are arguments based on the occurrence of tonal collapse and vowel reduction in certain sandhi contrasts in running speech (Hiranburana, 1971, 1972; Leben, 1971). One such argument given by Leben is as follows: in running speech the long vowel in the first element of the compound $/\underset{\text{HL}}{\text{thi:}}\ \underset{\text{LH}}{\text{nai}}/$ is shortened and the tone is simplified to a mid tone, that is, $/\underset{\text{M}}{\text{thi}^{\cdot}}\ \underset{\text{LH}}{\text{nai}}/$. The single dot after the vowel of the first element in the running speech style indicates that some length is retained on vowels not followed by a consonant or glide in the same syllable, even though shortening applies. Leben (1971:124) suggests that one "can account for the connection between what appeared at first glance to be two separate operations by expressing long vowels as geminates, VV, and by expressing contour tones as sequences of level tones, with each level tone expressed as a feature on a vowel." That is, both vowel shortening and tone simplification can be expressed by a single rule: VV $\longrightarrow$ V.

Hiranburana (1971, 1972) also gives examples of tonal collapse in the language in support of segmental analysis of the tones; however, her observations differ to some extent from those presented by Leben (1971). For example, Hiranburana claims that dynamic tones reduce to high tone instead of the mid tone proposed by Leben. Both Leben and Hiranburana's arguments are challenged by Gandour's (1974a) acoustic analysis of the tones in running speech. Although Gandour does not look at the instances of tonal collapse specifically mentioned, he finds five tonal distinctions maintained in running speech and, particularly, that "neutralization of contour tones in fast casual speech on the unstressed initial syllable (containing a long vowel) of bisyllabic noun compounds does not occur" (Gandour, 1974a).[5] A similar finding of retention of tonal distinctions in running speech is reported in Abramson (1975a). Thus it would seem that arguments for a segmental approach based on tonal collapse phenomena, are at best inconclusive.

A different argument for segmental analysis is given by Gandour (1974a). He points to certain polysyllabic morphemes of Indic (Pali-Sanskrit) origin that show an interaction between tones and segments. In polysyllabic morphemes that commence with a sonorant segment and follow a short checked syllable carrying low tone, a syllable other than the first carries a low tone if it ends in a stop; otherwise, a rising tone. Gandour (1974a) claims that this generalization must occur at the lexical level of representation because high and low tones on short checked syllables are neutralized to mid tone in this position in everyday speech; for example:

sàʔwǎn 'heaven' sàʔmùt 'ocean'

But if the second syllable begins with a nonsonorant, these patterns of tone occurrences are not found, that is, another tone can occur on the second syllable, as for example, if (1) this second syllable begins with a nonsonorant (sàʔtaaŋ, 'Thai monetary unit') or (2) if the initial syllable carries other than a low tone (kháʔneen, 'grade', 'vote').

---

[5]Gandour does not argue against the generally accepted observation that tones collapse in certain unaccented grammatical morphemes such as particles or pronouns: such collapse, in fact, is seen in his own data [i.e., /khǎw/ (third person pronoun) appears to have a high tone in running speech].

According to Gandour (1974a:139) such occurrences "show that tone in Siamese is sensitive to surrounding segments and is thus evidence in support of segmental representation of tone in Siamese." Gandour's argument may be correct but should be accepted with caution, pending further studies of assignment of tone to loan words from nontonal languages.

In summary, a case for segmental description of the tones has been presented with regard to certain phonotactic arguments concerning (1) distinctive vowel length and distribution of dynamic tones on double vowels; and (2) the interaction of tones and segments in certain polysyllabic morphemes in Thai. Although there are no parallel arguments for a syllabic description of the tones in the literature, a case for it has been given here: grounds for a syllabic description exist in the traditional Thai grammars and writing system, as well as in the phenomenon of lexical meaning differentiated by tonal elements.

C. Phonological feature descriptions of Thai. With regard to generative phonological descriptions of the Thai tones, both syllabic and segmental approaches have been formulated in the literature. Warotamasikkhadit (1967) has described dynamic tones with syllabic features; Gandour (1974a), on the other hand, has described the dynamic tones with segmental features.

Below is the description of the five lexically contrastive tones given by Warotamasikkhadit. It includes the three features [High], [Mid], and [Fall].

Features/Lexically Contrastive Tones	Mid	Low	High	Falling	Rising
High	-	-	+		
Fall				+	-
Low	-	+	-		

Redundancy rules fill in the remaining feature specifications as follows:

Features/Lexically Contrastive Tones	Mid	Low	High	Falling	Rising
High	-	-	+	+	+
Fall	-	-	-	+	-
Low	-	+	-	+	+

The phonological feature [Fall] is context dependent: that is, the context [+High] and [+Low] enables the feature [-Fall] to specify that the features [Low] and [High] are ordered sequentially to account for the manifestation of a transition from low to high F_0 of the rising tone at the level of the speech event. The same context allows the feature [+Fall] to specify a different manifestation at the level of the speech event for the falling tone: a transition from high to low F_0. Thus, the feature [Fall] has no direct phonetic implications of its own, and in essence one needs two different definitions for the feature [+Fall].

Abramson[6] has suggested another approach to a syllabic view that eliminates the problem of ordering of the features [High] and [Low] involved with Warotamasikkhadit's feature [Fall]. It specifies a feature [Dynamic] that has a single definition, that is, "abrupt movement of the F_0 contour." The features [High] and [Low] indicate gradual movement in the directions named. The feature [+Dynamic] in combination with one of these features specifies in the systematic phonetic component abrupt movement in the direction named.

Features/Lexically Contrastive Tones	Mid	Low	High	Falling	Rising
Dynamic				+	+
High	-	-	+		+
Low	-	+	-	+	

[6] Arthur Abramson, 1976: personal communication.

Since Abramson's proposed feature system states essentially the same phenomenon as Warotamasikkhadit's, but more elegantly (it does not require two definitions for a single binary feature), his system will be referred to in a future discussion of syllabic treatments of the tones.

There is yet another generative account of the tones in the literature (Whitaker, 1969) in which are posited three tone features [High], [Rising], and [Falling]. This analysis is not considered in detail. The interested reader is referred to Gandour (1974a), whose critical assessment of the system is accepted here.

Contrasted with the syllabic description is Gandour's segmental description, accounting for five lexically contrastive tones with two phonological features [High] and [Mid].

Features/Lexically Contrastive Tones	Mid	Low	High	Falling	Rising
High	−	−	+	+ −	− +
Mid	+	−	−	− +	+ −

The feature [Mid] is used instead of [Low] to explain more efficiently the frequent occurrence of tones becoming [Mid] in certain environments and in connected speech (Gandour, 1974a:143). He has not shown how this system provides a generalization that prevents the mid tone from combining with the high and low tone in a dynamic tone. The relations of the features proposed by Abramson and Gandour to the acoustic F_0 data are discussed more fully in Part 2, Section C, of this chapter.

D. Theorectical implications with regard to segmental and syllabic phonological descriptions of the tones. Interesting implications for phonological theory arise from discussion of these two feature systems. With the proper phonetic implementation rules, both systems yield the appropriate tones; specifically, both yield dynamic tones with appropriate kinetic F_0 movements. The major difference between these feature representations is the point in a derivation at which the kinetic characteristics of the dymamic tones are introduced: at the surface output level, or in the deep structure phonology. Abramson's syllabic analysis posits dynamism (by use of the feature [Dynamic]) at the phonological level, whereas Gandour's segmental analysis allows dynamism to result from the processes that convert phonological features into a temporal speech signal.

Deciding between these two forms of description bears on the general problem of the relation of the speech event to the phonology.[7] Obviously, a relation exists between speech event and deep structure, but its nature is not apparent. For instance, onto which aspect of the speech event is the systematic phonetic level most readily mapped: articulatory gestures or acoustic signal? Lieberman (1970) suggests that certain features map more directly onto the acoustic signal, others onto the articulatory gestures, and yet others onto a complex relation of both.

How does one decide onto which aspects of the speech signal a particular figure maps most directly? The answer has important

[7]Although the meaning of the term "phonetics" is generally agreed upon by linguists, there is considerable potential for confusion owing to the similarity between the two terms "phonetics" and "systematic phonetics." Phonetics is concerned with describing the speech sounds (Ladefoged, 1975), and the study of phonetics includes study of speech events, such as acoustic signal, articulatory gestures, muscle activity, and neural commands. Phonology is concerned with describing abstract representations of the speech sounds, and the study of phonology includes a systematic phonetic level that describes the speech sounds in terms of matrices of features (Chomsky and Halle, 1968). The relationship between systematic phonetics (the lowest level in the phonology) and phonetics (the speech event) is the topic under discussion here.

phonological implications. If simple correspondence is found between categories of tone and specific F_0 contours, suggesting that tones map most directly onto the acoustic signal, it follows that dynamic tones must be represented at the systematic phonetic level, since the acoustic data show an everchanging F_0 over time that is not divided into discrete units. Viewing the five tones of Thai as syllabic elements is motivated in large part by such acoustic evidence. Examination of the F_0 contours of the tones in both citation forms and running speech induced Abramson (1975a:124) to state that "...the data lend no phonetic plausibility to arguments for the specification of the dynamic tones as temporal sequences of features that underlie two of the static tones." He concludes: "It seems psychologically far more reasonable to suppose that the speaker of Thai stores a suitable tonal shape as part of his internal representation of each monosyllabic lexical item."

His argument regarding tones might parallel his argument respecting vowel length (discussed in Section B, above), specifically, that signs of neuromuscular rearticulation are necessary to show that a level tone on a double vowel syllable is analyzable in terms of segments at the systematic phonetic level.

On the other hand, if an aspect of the articulatory control of F_0 is found to correspond simply to tone category, suggesting that tones map most directly onto the articulatory gestures (as, for example, a sequence of two discrete muscle gestures operant for the dynamic tones), a segmental treatment of the tones is plausible. If, however, the EMG signal shows a similarly time-varying and equally nondiscrete pattern as that of the F_0 contours, there sould be no impetus from such data to challenge a syllabic treatment of the tones as derived from the F_0 contours themselves.

Another means of examining the question of mapping tone features onto articulatory control of F_0 is in terms of the concept of "invariance." The notion of invariance has been a central concern of general theories of both speech production and perception. In the area of speech production, research has been directed toward the question of whether invariance exists at the level of motor commands, articulatory configuration (or state), or in the acoustic signal. These levels, arranged hierarchically, are presented below.

Systematic Phonetic Level
↓
Motor Commands
↓
Muscle Contraction
↓
Vocal Tract Configuration
↓
Acoustic Signal

A general view is that there is less variation in the speech chain as one works from the acoustic level up through to the systematic phonetic and phonological levels of representation. One model (Liberman et al., 1967) predicts that the motor commands will show greatest invariance with the phonological features. Thus, one motivation for early EMG work was to find the "invariant features" at the motor command level. However, in light of findings of ubiquitous variability in EMG data (MacNeilage and DeClerk, 1969), the assumption of invariant features in the patterns of muscle activity has, for the most part, been discarded (MacNeilage, 1970).

An alternative assumption most commonly accepted is that certain articulatory targets, or spatial targets, can be found that show isomorphism with the systematic phonetic level (MacNeilage, 1970). One instance of simple mapping of a phonetic feature onto an articulatory target occurs with the feature nasality: the velum is lowered (or relaxed) for [+nasal] and is raised for [-nasal]. Moreover, in this case there is a one-to-one correspondence with the articulatory gesture and muscle required to effect the gesture.

An example of less direct mapping of a phonetic feature onto an articulatory (target) configuration occurs with the feature [rounding] (Lieberman, 1970:310). That is, although the lips maneuver toward a vocal-tract shape that is longer and has a smaller orifice at the lips, the degree of rounding consequent to this feature is not invariant. Moreover, "Unlike the situation for '<u>nasality</u>,' where the same muscle always executes the relevant articulatory maneuver, different talkers appear to use different muscles to effect this articulatory maneuver."

To account for this, Lieberman introduced the concept of a feature as a "state function" at the articulatory level. That is, alternate patterns of muscular activity produce like articulatory "states" (as above, in lip rounding and unrounding). The conceptualization of abstract articulatory states described by Lieberman and abstract spatial targets toward which the articulations are aiming, as discussed by MacNeilage (1970) are quite similar. Neither would expect a generally simpler relation between phonology and muscle activity than between phonology and acoustic signal, since the pattern of muscle activities necessary to attain a given articulatory goal (state of target) may be quite complex.

With regard to tonal features, the primary articulatory goal is related to the state of tension of the vocal folds. Specifically, it is assumed that variations in tension of the vocal folds underlie the F_0 changes of tonal distinctions. As discussed in Chapter III, it is known that the vocal folds must be short, thick, and relatively slack to achieve low F_0; and long, thin, and taut to achieve high F_0. Thus, a corresponding hierarchical organization of laryngeal control can be outlined as follows:

Systematic Phonetic Level
↓
Motor Commands
↓
Muscle Contraction
↓
Vocal Fold State
↓
Acoustic Signal (F_0)

It may be that the correspondence between vocal fold state and tone feature is one of direct mapping of the phonological features [High] and [Low] onto vocal fold states of tension and laxity. This is the system proposed by Halle and Stevens (1971). They propose the features [Stiff] and [Slack] to describe the vocal fold states and assert that these are sufficient to describe three levels of tone distinctions that occur in languages:

Tone/Laryngeal Feature	Stiff	Slack
High	+	-
Mid	-	-
Low	-	+

Although indirect techniques such as electromyography, coronal X-ray photography, and high-speed motion pictures are used to measure other aspects of vocal fold function, the states of tension of the folds are not directly measurable. It is assumed, however, that the state of the folds may be inferred from measurements of the muscle activity presumed to contribute to these states. Therefore, rather than examining states of vocal fold tension directly, this study examined patterns of laryngeal muscle activity that bring about these states of vocal fold that determine the F_0 contours of the Thai tones.

The questions addressed in this study are (1) what is the nature of the EMG patterns underlying the five lexically contrastive tones, and (2) does the EMG evidence support either a segmental or a syllabic description of the tones, or both; specifically, do the patterns inherent in the data suggest a form similar to either, both, or neither of these phonological feature systems? (3) Do the patterns support a laryngeal state description of the type proposed by Halle and Stevens (1971)? A goal of this study, therefore, is to extract phonetic features from the EMG data.

Part 2: Acoustic Description of the Tones

In this part are discussed (A) previous acoustic analyses of the tones of Central Thai in citation form. This is followed by a discussion of (B) the static and dynamic contour distinction, which in turn is followed by a discussion of (C) the relation of the F_0 contours to phonological descriptions of the tones mentioned earlier. A final discussion entertains the question of (D) the relevance of acoustic signal data to phonological segmentation.

A. Review of the literature. The acoustic shapes of the tones in citation form have been described by Bradley (1909, 1916), Abramson (1962, 1975b), and Erickson (1974). Bradley's study is mentioned for historical purposes. It is not treated in detail here because of its limitations: only one sample from each tone type was analyzed, and the oscillographic analysis was performed with kymographic techniques, which were less reliable than modern methods of F_0 extraction.

The study by Abramson (1962) is based on spectrographic analysis of a large sample of the F_0 contours of Central Thai as spoken by a native speaker on words composed of syllables with various initial consonants. Briefly, the F_0 contours of the tones on long vowels as reported by Abramson can be described in the following manner (see Figure 2.1):[8] the mid tone is a relatively flat contour in the mid voice range that drops toward the end of the syllable. The high tone is a relatively flat contour that begins with a F_0 slightly higher than the mid tone and reaches an even higher F_0 before dropping slightly. The low tone is a relatively flat contour that begins with a F_0 slightly lower than

[8]Very similar results were obtained for a corpus of short vowels as well.

the mid tone and slopes downward to end on an even lower F_0. The falling and rising tones have relatively dynamic ranges of F_0 excursions compared to the other three tones. The falling tone begins with a F_0 slightly higher than the mid tone, rises to a high maximum, and then drops to a F_0 about as low as that of the low tone. The rising tone begins with a F_0 slightly lower than the mid tone, drops down to a low minimum, and then rises to a high F_0 comparable to that of the high tone.

Later studies by Abramson (1975a) examine F_0 contours of a single syllable type /khaa/ as spoken on each of the lexically contrastive tones (Figure 2.2). The contours in this study differ slightly from those of the earlier study in that in the later study the falling tone shows no initial rise. This may well be due to the effect of the initial voiceless aspirate: those consonant types are known to have higher fundamental frequencies than, for instance, voiced consonants (cf. Part 3 of this chapter). That the earlier study did not reflect this pattern for the falling tone might be a result of averaging: the earlier F_0 contours were averages of all syllable types. In addition, the high tone in this later study shows a dip in its trajectory that is not seen in the earlier study. No explanation is readily available for this; it may be a variant of this tone for some speakers under certain circumstances.

The study by Erickson (1974) reports acoustic descriptions of F_0 contours from the speakers in the present EMG study, and therefore will be reviewed in detail here. Averaged F_0 contours[9] for each of the speakers for each of the tones on the syllable /buu/ are displayed in

[9]See Chapter IV for a description of F_0 analysis and averaging techniques.

Figures 2.3 and 2.4. [These figures are redrawn from Erickson (1974) to show the five tones for each speaker on the same graph.[10]] In these figures the F_0 trace begins at the onset of phonation after release of the stop.

In general, the five tones are distinguished for each of the four speakers by five distinct F_0 patterns, as was seen in the studies by Abramson. The following characterizes the F_0 patterns in this study: for all speakers the low tone contrasts with the mid none in its more pronounced downward slope. Also, the low tone onsets with a lower F_0 than does the mid tone, across all speakers. Speaker MJ shows relatively little downward slope, as well as a somewhat high initial F_0 for the low tone, but as is seen in Figure 2.3 even this shallow slope serves to distinguish it from this subject's mid tone. It should be mentioned that under difficult listening conditions, these two tones are most likely to be confused (Abramson, 1974b). The high tone is clearly distinct from the mid and low tones, by virtue of its upward sloping F_0 contour. The falling tone shows similarity to the mid tone in that both show a final fall; however, for all speakers the final drop of the falling tone is considerably larger than that of the mid tone. The rising tone shows some similarity with the high tone in that the high tone may drop at the beginning, but this drop is slight. Speaker CT (Figure 2.3) shows no initial fall, his rising tone is clearly distinct from his high tones; the rising tone starts lower and shows a greater F_0 excursion.

---

[10]The averaged contours for speaker MJ displayed here are slightly different from those in Erickson (1974) due to the fact that the averages in the earlier study involved three repetitions, while these are from eight repetitions of the syllable.

The tonal contours in this study closely resemble those displayed in Abramson (1962: Figure 2.1), with two exceptions: (1) the acoustic description of the mid tone in Abramson (1962: Figure 2.1) shows a final F_0 drop of only 10 Hz, contrasted here with an average F_0 drop of about 30 Hz. (2) The descending slope of the low tone in these data is more gradual than that of Abramson, where it drops more abruptly and levels off more quickly. These differences may be due to artifacts in analysis techniques, or, more likely, the number and type of syllables elicited. The representative contour displayed in Abramson is taken from a single speaker on syllables ending with vowels, semivowels, and nasals; in this study the representative contour comes from four speakers on syllables ending in a long vowel.

B. Static and dynamic contour distinction. Abramson (1962) categorized the F_0 patterns of the five lexically contrastive tones into two groups: static and dynamic contours. The static contours--the mid, high, and low tones--are relatively flat; the dynamic contours--the falling and rising tones--have F_0 trajectories that change direction sharply.

Static contours

Inasmuch as all the tones show some movement, the terms "static" and "dynamic" are relative. Inspection of the trajectories in Abramson (1962, 1975b: Figures 2.1 and 2.2) and Erickson (1974: Figures 2.3-2.4) reveals that the "static" tones are not wholly static: the high tone tends to rise slowly, and the low tone to drop slowly; the latter effect is somewhat more noticeable. The contours displayed in Erickson (1974) also show initial perturbations in F_0 that are accountable in terms of F_0 characteristics of initial consonants that vary according to their voicing distinctions (cf. Erickson, 1975a; Gandour, 1974b). (An example of perturbation due to voicing of the initial consonant can be seen for the high and mid tones of speaker SS, Figure 2.3.) Another dynamic aspect of static tones is the tendency for the final F_0 to drop slightly, especially in the mid tone. This is an artifact of this tone's prepausal position and is thought to be an effect of decreased subglottal pressure (Abramson, 1976), which is known to occur frequently at the end of phonation (Lieberman, 1967), although no direct measurements of subglottal pressure have been made for Thai tones. Perceptual studies of Thai tones (Abramson, 1974b) having level F_0 trajectories without final falls indicate that F_0 levels can be identified reliably as static tones. Thus the significant characteristic of the static tones seems to be the relatively steady-state portion in mid syllable, that is, those portions that show little movement compared to the dynamic tones. One exception, however, is a high tone reported by Abramson (1975b), which shows virtually no steady portion (Figure 2.2).

Dynamic contours

The term dynamic contour is, again, relative. These contours tend to show more F_0 movement than their static counterparts. It might be said that there is a generic pattern for the dynamic tones: in the falling tone, a peak in F_0 is followed by a drop; for the rising tone, a drop in F_0 is followed by a peak. Such patterns are seen in Abramson (1962) and for three of the speakers (PT, SS, MJ) in Erickson (1974). Some variations, however, can be seen. For example, the falling tone reported by Abramson (1975b) (and displayed here in Figure 2.2) shows a tendency for F_0 to start and stay high, then drop abruptly downward. Even more variation is seen for speaker CT (Erickson, 1974) (here in

Figure 2.3), who shows a rising tone with a steady but slow rise without any initial dip. Nevertheless, it is apparent from comparison of his rising and static tones that the rising tone shows more movement than the others.

In conclusion, although it seems obvious upon examination of F_0 contours that they fall into the two categories of dynamic and static contours, it is difficult to present a simple description of the differences between the two types, applicable in all cases. Perceptual studies by Abramson are in progress to ascertain the critical acoustic characteristics of the two types of tones. Preliminary findings (Abramson, 1976) suggest that there is a perceptual basis for the distinction. The primary acoustic cue for the distinction is not always obvious from examination of F_0 contours from several speakers.

C. Relation of F_0 contours to phonological descriptions of the tones. The phonological feature descriptions proposed for Thai tones (cf. preceding section of this chapter) are based to a large degree on their F_0 characteristics. In this section, these features are scrutinized to determine whether they adequately describe the acoustic shapes of the tones as discussed above.

The features [Dynamic], [High], and [Low] proposed by Abramson[11] (cf. Part 1, Section C, of this chapter), clearly have a basis in the F_0 data. The features [+High] and [+Low] indicate gradual movements in the direction named; association of these features with [+Dynamic] results in an abrupt movement in the direction named. Thus, these features allow for a distinction between the characteristic types of F_0 movement associated with the static and dynamic tones. The tones that

[11]Arthur Abramson, 1976: personal communication.

are not marked for [Dynamic] show a gradual movement (i.e., static tones). Application of these features to the static tones results in the following tonal patterns: the high tone is a contour with gradual upward movement, the low tone, with gradual downward movement, and the mid tone, with virtually no movement. This is basically an accurate description of the mid tone in that the final drop that occurs in citation form is an intonational phenomenon before a pause. It would seem that the mid tone is the most nearly level of the static tones, since it is the only tone that "does not make extreme excursion into the high and low regions of the voice range" (Abramson, 1976).

An exception to this, however, is seen for speaker SS (Figure 2.3), who shows virtually no steady F_0 portion but a consistent downward slope (after the initial perturbatory rise due to the initial consonant) throughout the syllable. It is interesting to examine the aberrant high tone of Abramson (1975b) (here in Figure 2.2) in terms of this feature system. Even though this high tone shows a dip, at the same time it shows a gradual upward rise in F_0. The dip in the upward movement could be introduced by implementation rules that specify variations in terms of physiological and aerodynamic factors.

Similarly, application of the features to dynamic tones results in the following analysis: the falling tone starts high and then abruptly drops down; the rising tone starts low and then abruptly rises. The quality of "abruptness" is also relative: for instance, the rising tone of speaker CT (Figure 2.3) shows a less abrupt movement upward than the rising tones seen for other speakers.

In general, however, it would seem that Abramson's features are descriptively adequate for the general shapes of the F_0 contours of the five lexically contrastive tones of Central Thai. For the manifestation in the speech event, of course, specifications of actual F_0 excursions for the individual speaker would be needed, or, perhaps, specifications of gross muscular adjustments that yield a range of acceptable F_0 excursions with some of the variability ascribed to aerodynamic factors.

Gandour's two features [Mid] and [High] are more abstract relative to the F_0 characteristics of the tones than are Abramson's features. This is true on two accounts: first, Gandour chose the feature [Mid] instead of [Low] to account for a "phonological reality"; "acoustic reality" was a secondary consideration in his choice. More importantly, Gandour's segmentation of the dynamic tones on the phonological level is at best an abstraction of the F_0 contours, since no segmentation is seen on the acoustic level. The fact that the contours are continuous need not preclude an analysis of this kind. It is possible to segment the dynamic tones into high and low portions of the F_0 contours. This possibility is illustrated by a schematic display (Figure 2.5) of the dynamic tones of one of the speakers (PT) from Erickson (1974) (here 2.4). In this figure three points along the F_0 continuum--the initial F_0, the maximum/minimum F_0, and the offset F_0--were plotted and a line drawn connecting these points. Instances of dynamic tones where movement is less abrupt [i.e., speaker CT (Figure 2.3)] are more difficult to analyze in this way and more difficult to segment on the basis of the acoustic data.

D. Relevance of acoustic data to phonological segmentation. This discussion leads to considerations of the relevance of acoustic signal data to questions of phonological segmentation.

While no apparent traces of segmentation occur in the F_0 contours of the tones, this in itself is inconclusive. One might argue that the F_0 contours can, in principle, provide no evidence regarding the question of segmental versus syllabic analysis of the tones, since the F_0 contour cannot, in principle, vary discontinuously (e.g., in the manner of a step function) while voicing is maintained. The mass of the structures involved is such that changes in tension of the vocal folds or position of the cartilages must occur gradually. Whether or not the phonological structure involves discontinuities, the acoustic realization of these changes as F_0 contours must always be relatively smooth. Thus, logically, examination of the F_0 contours cannot provide direct evidence bearing on the phonological question at issue here. The problem of isolating tonal segments on the acoustic level is analogous to the well-known difficulty of isolating discrete units for segmental phonemes (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967), for example, the problem of isolating vocalic segments in diphthongs. The similarity is noted here and discussed in more detail in Chapter 5, Part 2.

The EMG data are more promising in this respect. Changes in muscle innervation can occur quite rapidly, and recordings of muscle activity might reflect the rapidity of such changes much more directly than can the F_0 contour. Further, the possibility arises when considering muscle activity, that certain discontinuities in the phonological representation of a syllable may be reflected as discrete changes involving successive activation of different muscles. Such changes are relatively invisible in the F_0 contour, since whatever factors contribute to determining F_0, whether changes in level or activation of a muscle, changes from one muscle to another, or changes in subglottal or supraglottal pressure, all combine to give a single contour from which the contributing influences cannot be determined except by very indirect means.

One question that will be addressed in this study, as earlier mentioned, is whether the EMG data can offer support for a segmental description of the tones, or whether one must conclude that there is no support for such a segmental description in the speech output data at either the acoustic or the articulatory level. This would not necessarily invalidate segmentation on the phonological level, but would suggest that no evidence of this type is available in the speech signal itself. If one wished to hold to a segmental analysis despite the lack of observable evidence in more peripheral speech events, one might propose that the segmental nature of the features is lost at a very early stage in the speech production process and is unavailable to present techniques of speech research.

Part 3: Historical Background

A further goal of the study is the interpretation of acoustic and physiological differences among stop consonants, to contribute to a historical account of the emergence of the tonal and consonantal systems of Central Thai. The stop consonants examined in this study are the bilabial phonemes /b,p,ph/. The phonetic characteristics of these phonemes are defined in terms of voice onset time (VOT) relative to the release of the occlusion: /ph/ is an aspirated voiceless stop in which voicing is delayed until somewhat after the release of the occlusion; /p/ is an unaspirated voiceless stop in which voicing onset is simultaneous with stop release; and /b/ is a voiced stop in which voicing commences prior to the release (Lisker and Abramson, 1964).

It is generally believed that the tone system of the Tai languages developed from the proto-Tai voicing distinction between initial consonants; namely, that high tones typically followed originally voiceless consonants, and low tones, voiced consonants (Maspéro, 1911; Burnay and Coedès, 1927; Li, 1966; Haudricourt, 1971, 1972; Matisoff, 1973; Brown, 1975). An example of this is diagramed below:

Proto-Tai (1)	Original tone distribution (2)	Modern tone distribution (3)
pha	phá	phá bá
ba	bà	phà bà

The sequence of events is presumed to have been as follows: first, phonemic contrasts of initial consonants; second, as a result of the tendency for F_0 to the higher for voiceless initial consonants, a tonal distinction of high and low tones evolved in which the high tone was associated with voiceless consonants and low tone with voiced; finally, as the consonants themselves changed with respect to voicing distinctions, the tonal distinctions became more independent of consonant type so that either high or low tones might occur with either voiced or voiceless consonants. For a more detailed description of this development of tone and consonant change, the reader is referred to Erickson (1975).

Experimental support of this theory derives from acoustic studies with both English stop consonants (e.g., Lehiste and Peterson, 1961; Mohr, 1971) and Thai stops (Gandour, 1974b; Erickson, 1975a), which have demonstrated that syllables having initial voiced stops show lower onset F_0 than those with voiceless initial stops; though the Thai studies disagree on whether voiceless aspirates or inaspirates have the higher F_0. Erickson (1975) found voiceless aspirated stops to have higher fundamental frequencies than do their inaspirate counterparts; Gandour (1974b) reported the opposite.

Both studies show less difference in F_0 onset values between voiceless aspirates and inaspirates than between F_0 onset for voiced and voiceless consonants. It is the latter difference that is relevant to the historical argument. The data support the historical explanation for tonogenesis in Thai, wherein differences originally present in F_0 in consonants became gradually more perceptible, until they were heard as pitch characteristics of the vowel. This characteristic pitch subsequently assumed phonemic import as a distinctive tone (Abramson, 1974a; Hombert, 1975). Some support for this explanation is found in perceptual studies of voicing characteristics of initial stop consonant influenced by F_0 done by Haggard, Ambler, and Callow (1970) for English, Fujimura (1971) for English and Japanese, and Abramson (1974a) for Thai. Appropriate F_0 shifts persuade listeners to modify their category assignments of consonants having ambiguous values of VOT. Possibly, during an epoch in the development of the language--when voicing categories for initial consonants were unstable--such pitch perturbations began to serve a role in distinguishing the voiced from the voiceless stop consonants. Perhaps they began to be exaggerated intentionally. As the VOT differences between the two categories stabilized, becoming capable of bearing the major perceptual load of the distinction, the F_0 differences were freed to become independently distinctive sound patterns.

It is known that both aerodynamic and laryngeal factors are involved in F_0 production (cf. Chapter III). Support of the tonogenesis explanation might be found partly in laryngeal physiology in the use of similar laryngeal mechanisms for production of both stops and tones, specifically, voiceless stops and high tones, and voiced stops and low tones. As stated in the discussion of laryngeal physiology (Chapter III), EMG studies have shown a correlation of cricothyroid activity with high F_0, and strap muscle activity with low F_0. It may be that voiceless consonants are accompanied by increased cricothyroid activity and decreased strap activity, and voiced consonants, decreased cricothyroid activity and increased strap muscle activity. If it were shown that active laryngeal control mechanisms contribute to the F_0 differences between the voiced and voiceless stops, the case for a role of such differences in tonogenesis would be stronger than if the F_0 differences among stops were found to be due solely to aerodynamic factors.

No EMG experiment has yet been done with Thai consonants; however, such studies are reported of laryngeal activity during production of English stops (Hirose and Gay, 1972) and Hindi stops (Dixit, 1975;

Kagaya and Hirose, 1975). These studies are relevant to this discussion in that both English voiceless aspirates and Hindi voiceless aspirates, inaspirates, and voiced inaspirates, are similar to corresponding Thai stops (Lisker and Abramson, 1964). Although simultaneous F_0 analysis wss done only in the study of Hindi stops by Kagaya and Hirose (1975), it is assumed that in all studies initial voiced stops show lower fundamental frequencies than voiceless ones. Results of these studies with regard to muscle activity for voiceless stops are confusing. Hirose and Gay (1972) report no increased cricothyroid activity, whereas Dixit (1975) and Kagaya and Hirose (1975) do. The latter two studies (of Hindi) disagree in certain details: Dixit finds increased cricothyroid activity for both aspirate and inaspirate voiceless stops, whereas Kagaya and Hirose find such activity for the inaspirates only. The latter study also reports a higher initial F_0 for the voiceless inaspirates than aspirates; the former study has no F_0 information.

In voiced stops both Hindi studies show decreased cricothyroid activity. Dixit's study also examined strap activity in connection with voiced stops but found no increase; rather, the data show increased strap activity for the voiceless stops. Strap activity has been observed with English voiced /b/ (Bell-Berti, 1973, 1975). Although English /b/ is phonetically not identical to any of the Thai stops (see Lisker and Abramson, 1964), this finding encourages examination of strap muscle activity for voiced consonants in Thai.

Although the studies cited above are suggestive, the results are inconsistent and do not speak directly to the topic of tonogenesis in Thai. What is needed is a direct investigation of the correlation between laryngeal muscle activity and F_0 differences associated with voiced and voiceless stops in Thai. The present study seemed an excellent opportunity for such an examination, and the appropriate comparisons were undertaken. The results of these examinations are presented in Chapter V, Part 3. Aerodynamic mechanisms may provide alternative evidence regarding the tonogenesis theory described above, and EMG studies of the laryngeal muscles that act to control glottal opening may be important in elucidating these mechanisms. But such mechanisms and the investigations appropriate to them were deemed outside the scope of the present study.

CHAPTER III: PHYSIOLOGICAL BACKGROUND: LARYNGEAL DETERMINANTS OF F_0

Many experimental studies have investigated the physiological factors controlling the fundamental frequency (F_0) of vibrations of the vocal folds. These have included electromyographic (EMG) studies of the laryngeal muscles (e.g., Sawashima, 1974) as well as aerodynamic studies of subglottal pressure (e.g., Van den Berg, 1957), and air flow (Klatt and Stevens, 1968). It is known that both laryngeal and respiratory functions effect F_0, but the exact role of each is unclear. The most widely accepted account of laryngeal vibratory behavior is Van den Berg's (1958) myoelastic aerodynamic theory. Briefly, it states that the vocal folds are set into vibration by pressure forces generated by the lungs during expiration. For phonation, the vocal folds are brought together by the adductor muscles of the larynx. Adduction causes an increase in subglottal pressure to be built up under the folds. When the subglottal pressure overcomes the glottal resistance, the vocal folds are blown apart. The decrease in pressure at the glottis while the folds are in the abducted position (Bernouilli effect) along with the elastic properties of the vocal folds, themselves, causes the vocal folds to return to the closed position. The entire cycle is thus repeated. The fundamental frequency corresponds to the rate at which the vocal folds complete one full vibratory cycle,which rate is determined primarily by the mass of the vocal folds, their tension, and the subglottal pressure.

Recent investigations of laryngeal and respiratory functions in F_0 control during speech, specifically, aerodynamic and electromyographic studies of English intonation (Vanderslice, 1967; Ohala, 1970; and Atkinson, 1973a) and Dutch intonation (Collier, 1975), suggest that laryngeal muscles have the primary role in the control of F_0, with aerodynamic factors a secondary. The laryngeal muscles primarily affecting F_0 and investigated in this study are the cricothyroid, vocalis, and strap muscles (sternohyoid, sternothyroid, and thyrohyoid). Fundamental frequency is a function of the complex interaction of length, mass, and tension of the vocal folds; parameters which are affected by these muscles. In the normal speech mode these muscles effect changes in F_0 primarily by altering tension of the folds. The cricothyroid changes the tension of the folds by modifying their length and mass; the vocalis muscle varies tension inherent in folds in the absence of change in their length or mass (Kotby and Haugen, 1970:207). The strap muscles are thought to reduce the tension by either shortening the folds or by increasing their mass, or by a combination of these strategies, as yet unspecified. A more detailed account of the manner in which these muscles affect vocal fold tension is given in the following review of laryngeal anatomy and physiology.

A summary of the anatomy of the larynx, including a description of the vocal folds, their cartilaginous attachments, and the laryngeal muscles which effect changes in F_0 (specifically those examined in this study), is presented below. It is followed by an account of the physiology of the larynx and the function of the laryngeal muscles in the control of F_0.

Anatomy of the larynx[1]

The larynx can be viewed as a tube with folds of soft tissue through which air passes from the lungs into the upper respiratory tract (see Figure 3.1). The internal laryngeal cavity is rimmed at the top by the aryepiglottic fold (the back and side rims of the epiglottis), and at the bottom by the uppermost tracheal ring which attaches to the cricoid cartilage. Immediately below the aryepiglottic fold are the ventricular, or false, vocal folds, and below these are the true vocal folds. The ventricular folds and the true vocal folds are separated by a cavity called the laryngeal ventricle. The area between the separated vocal folds is the glottis. The entire internal laryngeal cavity is lined with elastic membranes. The quadrangular membrane lines the larynx above the false vocal folds, and its upper edges constitute the aryepiglottic fold. The conus elasticus lines the larynx below the true vocal folds, and its upper edges constitute the vocal ligaments.

Recent work indicates that the vocal folds are made up of (1) a mucous membrane which consists of an epithelium and a superficial layer of mucous tissues; (2) an underlying ligament which consists of the elastic tissues of the conus elasticus membrane; and (3) the vocalis muscle (Hirano, 1974). The mucous membrane can vibrate by itself, or the entire fold including the vocal muscle, can vibrate. The mucous membrane and the elastic ligament are loosely connected, which allows the mucous membrane to vibrate without affecting the underlying ligament and the vocalis muscle. The fibers of the conus elasticus insert into

---

[1]For a more detailed description of the larynx, see e.g., Pernkopf (1963: Figures, 308, 324-332), Zemlin (1964), or Dickson and Maue (1970).

the vocalis muscle, which provides a tight coupling between the conus elasticus and the vocalis muscle, permitting the vocalis muscle to vibrate with the mucous membrane and the elastic ligament. In this sense the vocal folds can be regarded as a double-structured vibrator which consists of (1) a body made up of the vocalis muscle and the conus elasticus membrane, and (2) a cover made up of the mucous membrane (Hirano, 1974:90).

Cartilaginous attachments of the vocal folds

The vocal folds are paired and form a triangle in the horizontal plane, the apex of which is attached anteriorly to the thyroid cartilage, while the bases of the triangle are attached posteriorly to the vocal processes of the arytenoid cartilages (see Figure 3.2). The arytenoid cartilages are perched on the cricoid cartilage.

Thyroid cartilage. The thyroid cartilage is the largest laryngeal cartilage, comprising most of the anterior and lateral walls of the larynx (see Figure 3.3). The stability of this cartilage is due to its anchoring to the sternum and to the hyoid bone by the extrinsic muscles (Arnold, 1961:691). The thyroid is composed of two plates of cartilage fused in front and widely separated in back. The V-shaped notch where the two plates fuse anteriorly is the laryngeal prominence, or "Adam's apple." The vocal folds are attached (anteriorly) to this portion of the thyroid cartilage. In back, the borders of each thyroid lamina are extended superiorly and inferiorly as horns, known as the superior and inferior horns, respectively. The inferior horns articulate on either side with the cricoid cartilage.

Cricoid cartilage. The cricoid cartilage is like a signet ring with the signet part lying in back (known as the posterior plate), and the ring part lying in the front (known as the anterior arch). The cartilage forms the platform on which the arytenoid cartilages perch (see Figure 3.3). The cricoid cartilage sits on the uppermost tracheal ring.

Arytenoid cartilages. These cartilages are paired. Each member of the pair is broad at the base and diminishes in cross-sectional area as it rises toward an apex (see Figure 3.4). At the anterior corner of the base the vocal process projects, to which are attached the vocal folds. On the postero-lateral corner of the base is the muscular process, to which are attached the posterior cricoarytenoid, the lateral cricoarytenoid, and the thyroarytenoid muscles.

Cartilage articulations

Cricoarytenoid articulation. The arytenoids are attached to the cricoid plate in such a manner that closure is accompanied by an inferior and medial movement wherein the arytenoid cartilages approximate, and during glottal abduction the cartilages are pulled posteriorly and rotate outward, away from each other (Von Leden and Moore, 1961; Ardran and Kemp, 1966:644, 648).[2] (See Figure 3.5.)

Cricothyroid articulation. The ring of the cricoid cartilage is attached on each side to the inferior horn of the thyroid cartilage, permitting the cartilages to rotate around a common transverse axis (Arnold, 1961:689). (See Figure 3.6.) In theory, the thyroid cartilage can rotate forward and downward, or the cricoid can rotate upward and backward. However, since the thyroid cartilage is well anchored by the

[2] There is currently some discussion about this articulation concerning whether or not the movement is actually a rocking motion in the superior-dorsomedial-inferior-ventrolateral direction (Hirose, personal communication; Dickson, indirect personal communication).

strap muscles to the hyoid bone and the sternum, it is the cricoid cartilage which is believed to rotate in most instances (Faaborg-Andersen, 1957:13; Arnold, 1961:690; Ardran and Kemp, 1966:653).

The cricothyroid articulation also allows the cartilages to move either longitudinally or vertically without any rotating movement. In longitudinal movement either the thyroid cartilage moves forward or the cricoid cartilage moves backward. When there is vertical motion either the thyroid cartilage moves upward, or the cricoid cartilage moves downward. Dickson and Dickson (1972), however, question whether indeed the cartilages can articulate in the longitudinal plane.

Intrinsic laryngeal muscles

Vocalis muscle. The vocalis muscle is a bundle of fibers which comprise the bulk of the vocal fold proper.[3] The vocalis muscle is paired and attaches anteriorly to the thyroid plate and posteriorly to the vocal processes of the arytenoid cartilages (see Figure 3.7).

Cricothyroid muscle. The cricothyroid muscle, also paired, attaches in front to the lower edge of the thyroid cartilage and on the sides to the inferior horns of the thyroid cartilage. The front portion is the pars recta, and the side portion, the pars oblique (see Figure 3.8).

---

[3]There is a certain difficulty with nomenclature of which the reader should be aware. Strictly speaking, the vocalis muscle is part of the thyroarytenoid muscle. The thyroarytenoid muscle consists of two parts: the thyrovocalis and the thyromuscularis. The term "vocalis" in a strict sense refers to the thyrovocalis, but a tradition has developed in EMG literature wherein reference is made to the entire complex as the vocalis muscle. This practice is therefore continued in this study.

Extrinsic laryngeal muscles

Sternohyoid, thyrohyoid, and sternothyroid (see Figure 3.9). The sternohyoid attaches to the sternum and to the hyoid bone near the midline, superficially. (The hyoid bone is a horse-shoe shaped bone which anchors the tongue superiorly and the thyroid cartilage inferiorly.) The thyrohyoid attaches to the major horn of the hyoid bone and to the thyroid plate on the oblique line away from the surface. The sternohyoid lies deepest and attaches to the sternum and the thyroid plate on the oblique line. All of these muscles are paired.

Physiology

Vocalis muscle. The vocalis muscle functions to increase F_0 by isometrically increasing tension of the vocal folds. The role of the vocalis in F_0 elevation is confirmed by a number of electromyographic experiments with singing and sustained phonation (Faaborg-Andersen, 1957; Hirano et al., 1969, 1970; Gay et al., 1972). Hirano et al. (1969) and Gårding, Fujimura, and Hirose (1970) found this to be true for speech also.

Cricothyroid muscle. The primary function of the cricothyroid muscle in speech is to control F_0, especially rises in F_0 (e.g., Arnold, 1961). This has been confirmed by numerous EMG experiments with both speech and singing (Katsuki, 1950; Faaborg-Andersen, 1957, 1965; Arnold, 1961; Hirano et al., 1969, 1970; Sawashima et al., 1969; Gårding et al., 1970; Simada and Hirose, 1971; Gay et al., 1972; Atkinson, 1973a; Collier, 1975). The cricothyroid effects rises in F_0 primarily by contraction of the pars recta, causing the anterior part of the cricoid cartilage to approach the thyroid cartilage (Faaborg-Andersen, 1957:13; Arnold, 1961; 691). This action tilts the cricoid plate (the posterior lamina) anteriorly and superiorly, thereby increasing the distance between the thyroid insertion of the folds in the front and their insertion on the vocal processes of the arytenoid cartilages in back. (See Figure 3.10 as compared with Figure 3.11a.) In this fashion, the folds are elongated and tensed and there is an increase in F_0. The space between the thyroid and cricoid cartilages in front serves as an index of tension of the vocal folds (Ferrein, 1741: 426-7; Sonninen, 1956:58; Faaborg-Andersen, 1957; Ardran and Kemp, 1966). Reduction of the cricothyroid space indicates that the folds have been elongated, with a resulting increase in F_0. Conversely, enlargment of the space indicates that the folds have been shortened, with a concomitant reduction in F_0 (see Figure 3.11). Decreases in cricothyroid activity have been observed in connection with low F_0 by Collier (1975). In these cases, low F_0 has been brought about by relaxation of the cricothyroid muscle.

Interarytenoids, posterior cricoarytenoid and lateral cricoarytenoid. The primary function of the interarytenoids and the lateral cricoarytenoid is glottal closure, whereas that of the posterior cricoarytenoid is widening of the glottal aperture. Activity of these muscles during rises in F_0 has been reported in various electromyographic studies but their role in regulating F_0 has not been found to be as consistent as that of the vocalis or cricothyroid muscles (Van den Berg and Tan, 1959; Hirano et al., 1969; Ohala, 1970; Gay et al., 1972).

Extrinsic laryngeal muscles

The extrinsic muscles (also known as the infrahyoid or strap muscles) may affect the vibratory properties of the vocal folds, although this role is certainly secondary to that of the cricothyroid and vocalis muscles in control of F_0 and is not well understood. Perhaps the best study on the topic is that by Sonninen (1956) which examined the anatomy and physiology of the sternothyroid muscle in a number of subjects and recorded its electric potential while the subjects were singing. His study suggests that the sternothyroid has a dual effect: it can function to assist in elongation and tension of the folds at high F_0, and to shorten and relax them at low F_0. Sonninen explains this anomaly in terms of: (1) the permissible movement in the articulation of the cricoid and thyroid cartilages, and (2) the ability of the folds to be lengthened or shortened as the distance between the cricoid and thyroid attachments changes. As mentioned above, the narrowing or enlarging of the anterior space between the cricoid and thyroid cartilages serves as an index of tension on the folds. The main observed effect of sternothyroid contraction is rotation of the thyroid cartilage anteriorly and slightly inferiorly (Sonninen, 1956:25). The effect of this at high F_0 when contraction of the cricothyroid elevates the cricoid arch toward the thyroid cartilage, is to increase the distance between the cricoid and thyroid (anterior and posterior) attachments of the vocal folds, thereby increasing their length and tension. The approximation of the cricoid and the thyroid cartilages anteriorly suggests that this strategy has been employed to tense and lengthen the folds (Figure 3.12).

At low F_0 when the cricothyroid muscle is inactive, and the system is more loosely coupled, contraction of the sternothyroid can have a secondary effect of tilting the cricoid cartilage anteroinferiorly at the site of the cricothyroid articulation (Sonninen, 1956:74). If the cricoid cartilage deviates from the horizontal to a greater extent than the thyroid cartilage, then the distance between the cricoid and thyroid attachments of the vocal folds will become smaller, contributing to their shortening and relaxation. This is indicated by an enlarged cricothyroid space (Figure 3.11c). Head position also influences the effect of sternothyroid contraction. For instance, when the head is angled backwards, contraction of the sternothyroid shortens the folds and lowers F_0, whereas with a forward positioning, as tends to be the posture of singers at high F_0, contraction of the sternothyroid results in increased tension of the vocal folds (Sonninen, 1956:51). The account of the effect of extrinsic musculature on tension of the vocal fold, control of the distance between the cricoid and thyroid attachments, is known as the "Rahmen Funktion" (Schilling, 1937), or frame function theory (Sonninen, 1956).

Although Sonninen discusses only the sternothyroid, various electromyographic experiments of pitch control in singing suggest that two other muscles--the sternohyoid and thyrohyoid--are also probably capable of affecting F_0 in two distinct ways. Thus, Sonninen (1956), Faaborg-Andersen and Sonninen (1960), Hirano et al. (1967), and Ohala and Hirose (1970) report sternohyoid activity at both high and low F_0; Faaborg-Andersen and Sonninen (1960) report thyrohyoid activity as well. Other electromyographic studies with speech report activity of these muscles at low F_0 only (Ohala, 1970; Simada and Hirose, 1971; Atkinson, 1973a; Collier, 1975). The extrinsic muscles also aid production of certain supraglottal speech gestures--such as jaw opening (as for [a]) or tongue retraction (in release of [k]). Since the extrinsic muscles control the position and movement of the hyoid bone, it is logical that gestures which require fixation or depression of the hyoid bone involve some activity of the extrinsic muscles. Such activity has been observed in a number of electromyographic experiments, e.g., Faaborg-Andersen and Vennard (1964), Ohala and Hirose (1970), Simada and Hirose (1970), Harris (1974), and Collier (1975).

Although taken as a whole EMG studies of extrinsic muscles yield complex results, studies specifically concerned with speech give somewhat simpler results. No study of strap muscle activity in speech has found strap activity for high F_0. The difference between results from studies of speech and of singing may be an effect of the fact that normal pitch levels in speech are confined to the lower portion of the speaker's total pitch range.

A question which has puzzled experimenters in electromyographic research on laryngeal control of F_0 in speech is "How is F_0 lowering accomplished: by passive relaxation of the cricothyroid alone, or by active involvement of the extrinsic muscles?" (Ohala, 1972). The above-cited results suggest that the answer could be the latter--the strap muscles are active during low F_0 in speech. One of the goals of this study is to determine if such activity of the strap muscles is necessary to account for the observed changes in F_0, or if cricothyroid relaxation is sufficient. Examination of Thai in this regard is especially interesting since each word has distinctive rises and falls in F_0 associated with the tonal contrasts.

CHAPTER IV: EXPERIMENTAL METHODS AND PROCEDURE

Electromyography is the recording of bioelectric potentials from motor units within a muscle. The motor unit is the functional unit of a muscle and consists of (1) a nerve cell body, (2) one nerve fiber, and (3) a number (10-600) of muscle fibers. A single motor unit can fire at rates up to about fifty times per second. To make smooth, prolonged muscle contractions, many motor units fire repeatedly in volleys, combining to effect a constant contraction. A muscle contraction is accompanied by an electrical potential detectable by paired electrodes placed in or near the muscle. The magnitude of this potential is directly proportional to the rate of firing of a single motor unit and the number of motor units actually firing, and also to the strength of the contraction.

One limitation inherent in electromyography is the inability to detect absolute amplitude of muscle contraction. This is due to both the random and asynchronous pattern of motor unit depolarization. The asynchronous firing of single motor units results in an interference pattern which results directly from the phase differential of the signals. The random nature of the firing, the fact that for a given contraction only certain motor units fire, and fire at different rates, results in the signal strength at the electrode varying inversely with the distance between electrodes and active motor units.

For a more detailed account of the principles and techniques of electromyography, the reader is referred to Buchthal (1957), Galambos

(1962), Faaborg-Andersen (1964), Cooper (1965), and Basmajian (1974).

Materials

The subjects, three men and one woman, were native speakers of the Thai dialect of Bangkok and its environs, which forms the basis of the official standard language. Three of them were students at the University of Connecticut, and one was a student at the University of Hartford.

The nine nonsense syllables /baa, bii, buu, paa, pii, puu, phaa, phii, phuu/ served as samples for obtaining the acoustic and physiological data. Each syllable was spoken on each of the five tones, making a total of 45 utterance types, in the carrier phrase (əə _____), ("Yes, it is a _____"). The carrier phrase was employed to minimize contamination from nonspeech activity in the EMG recording. Although the presence of the carrier phrase may have made the productions a little less like citation forms, in this study it was deemed of primary importance to study short utterances without the complicating effects of varying contexts. With the resulting data as a base, one might then go on to examine these syllable types in longer utterances.

Methods

Each utterance type was repeated 16 times to balance for utterance-to-utterance variation, and to average out random noise in the electrical activity of the contracting muscle. The laryngeal muscles examined, as mentioned earlier, included the cricothyroid, the vocalis, the sternothyroid, the sternohyoid, and the thyrohyoid. There was simultaneous recording of the acoustic signal for F_0 analysis.

The experiment was conducted utilizing the facilities and techniques in practice at Haskins Laboratories. Hooked-wire electrodes were fashioned from paired platinum-iridium wires bent or slightly hooked at the ends to prevent displacement from the muscle fiber. The insertions were done by Hajime Hirose, M.D. Techniques for placing electrodes into these muscles are described by Hirose (1971).

EMG data processing

The electromyographic signals were recorded and processed by the Haskins Laboratories EMG recording and data processing system (Port, 1971; Kewley-Port, 1973). In addition to muscle potentials, a calibration signal of known amplitude (300 microvolts) was recorded periodically throughout the experimental session for subsequent conversion of the EMG amplitude into microvolts. The following steps were followed to derive the averaged signal used in the displays in the discussion of the results. First, a visual display of the raw EMG signals, the audio signal, and the octal time code track was obtained from an oscillographic (Visicorder) output. Each individual token (repetition) of a given utterance type was identified, and a lineup point for each token was specified. The computer (Honeywell DDP-224) then rectified and integrated (filtered) the EMG signals to produce a smooth envelope of the muscle contraction pattern. Next, an editing program was used which allowed the processed tokens to be displayed individually on a display scope for the purpose of deleting (1) obvious errors in the lineup points, (2) inactive channels of muscle activity, (3) channels which showed electrode displacement had occurred during the experiment, and (4) tokens which showed obvious aberrant spikes from, for example, short circuiting at the electrode. The editing program also enabled the experimenter to visually ascertain whether or not the tokens of a single utterance-type showed a consistent pattern of activity. Finally, the averaged EMG signals were derived from a hard copy printout of the display. [For a detailed description of the processing system used in this study the reader is referred to Bell-Berti (1973:26-28).]

F_0 analysis

The F_0 contours of the tones were derived from speech data stored on audio tape and analyzed by the Phonetic Analysis System (PAS)[1] at the Dental Research Center of the University of North Carolina. The technique calculates F_0 from interpeak intervals. The output of the PAS is an ink-jet display (Mingograph), in which vertical line length is directly proportional to the F_0 period. The sampling interval is fixed (ca. 5 msec) and so occurs independently of the actual F_0 period. The overall F_0 contour is then derived by tracing a horizontal line connecting the end points of all the vertical lines produced by the device (vis. display sample, Figure 4.1). Average F_0 contours were derived in the following manner. Initially, each contour drawn on the ink-jet display (i.e., the horizontal line connecting the end points of the vertical lines) was copied onto tracing paper.[2] From each family of curves, determined by utterance-type, an averaged contour was determined by eye. Because scatter was minimal the representative contours were easily chosen (Figure 4.2).

Representative F_0 contours were determined for each utterance type from eight tokens arbitrarily selected from the total set of tokens of that type produced by each speaker. Although 16 repetitions of each utterance type were available, it was found that eight repetitions gave as good an average F_0 contour as 16. No attempt was made to

[1]Fonema Company of Sweden, designed by G. Fant and J. Liljencrants.

[2]Eight contours for three of the speakers, and 16 for the other speaker.

combine the representative contours of the different subjects since individual differences in range of F_0 and in the forms of various F_0 contours were considerable. For similar reasons no attempt was made to combine representative contours from different syllable types. Of course, contours were not combined from different tone types. The primary concern of the F_0 analysis was to allow comparison of F_0 contours with EMG records. The unit of comparison was the utterance type for a given subject. (The term "utterance type" refers to a unique consonant-vowel-tone combination of which there are 45 per speaker, as explained on page 50). Averaged F_0 contours from the eight tokens were paired with the averaged EMG contours from all of the available tokens for a given speaker-utterance type combination.

CHAPTER V: ANALYSIS AND DISCUSSION OF ELECTROMYOGRAPHIC DATA

The results of the electromyographic experiment on Thai tones are discussed in three parts. Initially, the role of the laryngeal muscles in F_0 production and the implications thereof for a theory of laryngeal control of F_0 are examined. The second part discusses the EMG patterns for each of the tones and the implications for a linguistic feature description of the tones in the phonology. Finally, the laryngeal activity associated with consonant production is interpreted in connection with the historical account of the emergence of tones and consonants in Central Thai.

Part 1: Role of Laryngeal Muscles in F_0 Production

In this part are discussed (A) laryngeal activity associated with the F_0 patterns, (B) correlation of EMG with F_0, (C) timing relation of EMG to F_0, and (D) reciprocity of cricothyroid with strap muscle activity.

Satisfactory EMG data were obtained for each of the speakers with certain exceptions, as outlined in Table 2. Starred (*) muscles in the Table are those with usable signals (100 μV or above) which figure in the amplitude and timing analyses discussed in Sections B and C. Muscles with weak levels of activity (50 μV or less) will be discussed only in Section A in the examination of the EMG/F_0 display figures, since they reflect tendencies seen for muscles with usable signals; they will not be subjected to amplitude and timing analyses, as will those with higher amplitudes.

TABLE 2: Recording acceptability of muscles by speaker.

Speaker/ muscle:	Cricothyroid	Vocalis	Sternohyoid	Sternothyroid	Thyrohyoid
Speaker CT	weak	contam.[a]	weak	*	weak
Speaker SS	contam.	weak	*	*	weak
Speaker PT	*	weak	weak	weak	*
Speaker MJ	*	*	*	weak	*

[a]contaminated

A. Description of laryngeal activity associated with F_0 patterns. The following discussion focuses on the five F_0 patterns of the syllable /buu/, as representative examples of association of muscle activity with F_0. For all analyses lineup was at the moment of the stop release, represented as the vertical line on the displays. The vertical scale factors were set in the computer program to maximize the vertical range in the plots without overload. For this reason, amplitude will vary for different speakers for the same muscle. This is mentioned because the differing voltage values can make visual comparison of amplitudes across speakers misleading.

Vocalis activity

Figures 5.1-5.3 show vocalis activity for speakers MJ, PT, and SS. The vocalis shows slight fluctuations with respect to F_0. This is most apparent for speaker MJ (Figure 5.1), probably because amplitude of his vocalis activity is higher than that of the other speakers. The other speakers show fairly flat activity throughout all the tones, regardless of F_0 change, except that speakers PT and, to some extent, SS, show increases of activity at the end of the syllable with the high tone (and to some extent, with the rising tone as well). This increased vocalis activity is thought to be an effect of the glottal constriction observed in the audio signal at the end of the syllables for especially the high tone for these speakers. Vocalis activity has also been noted to occur with glottalized Swedish consonants and Danish stød (Gårding et al., 1970; Hirose, 1974).

Cricothyroid activity

Figures 5.4-5.6 show cricothyroid activity for speakers MJ, PT, and CT. In general, the cricothyroid increases in activity with high F_0 and decreases with low F_0. The same pattern was seen for the vocalis but the change in cricothyroid activity is clearly greater than that seen for the vocalis. Each of the speakers shows essentially the same pattern. As can be seen in examination of cricothyroid activity for speaker MJ (Figure 5.4), the cricothyroid shows increased activity for the high F_0 associated with the high tone and also the high F_0 portions of the falling and rising tones. It shows low activity for the low F_0 of the low tone and the low portions of the dynamic tones, as well as the mid F_0 range of the mid tone. Also, the cricothyroid tends to show decreases in activity for the low F_0 portions of these tones. Speakers PT and CT (Figures 5.5 and 5.6) show the same pattern as seen for MJ, but the level of cricothyroid activity for the mid F_0 range of the mid tone is higher than that for the low F_0 of the low tone. This was not seen for speaker MJ who showed low level activity for F_0 patterns for both low and mid tones with activity for the mid tone being only slightly higher. It is noted that the F_0 levels, however, for the mid and low tones for speaker MJ are minimally distinctive. A rise in cricothyroid activity can be seen for speaker PT after the initial fall during the low F_0 of the low tone; however, this is thought to be a consequence of the /uu/ vowel, since for /aa/and /ii/, the cricothyroid remains low throughout the low tone. Also, a decrease in cricothyroid activity at the end of the high tone is noted to occur for this speaker and is associated with the increased vocalis activity for final glottal constriction. Gårding et al. (1970) also noted decreased cricothyroid activity to occur with vocalis activity associated with glottal constriction of the glottal stop in Swedish.

Strap muscle activity

Figures 5.7–5.10 show strap muscle activity for all four speakers. The strap muscles on the whole tend to show peaks of activity during low F_0 with relatively little or no activity during high F_0; each speaker shows the same configuration. Typical patterns of strap muscle are exemplified by speaker MJ (Figure 5.7). Increases in strap muscle activity occur with the final F_0 fall of the mid tone, with the downward sloping F_0 of the low tone, with the final fall of the falling tone, and with the initial fall of the rising tone. There are no peaks of activity for the high F_0 of either the high tone or of the final portion of the rising tone. Moreover, all three strap muscles manifest the same activity pattern.

Speaker PT (Figure 5.8) shows this same pattern, except that an increase in strap activity occurs at the end of the high tone which correlates with the vocalis activity associated with glottal constriction seen for this speaker on this tone.

Speaker SS (Figure 5.9) shows the same basic pattern of strap muscle activity with low F_0 but with certain variations. There is no increase of strap muscle activity at the end of the mid tone, but neither is there the sharp final fall of F_0 as seen for the other speakers for this tone. Moreover, peaks of strap muscle activity can be seen before the release of the stop /b/ (high activity with initial /p/ also occurs) which correlate with initial perturbations of F_0 for this speaker. The effect of strap muscle activity in connection with initial consonants is discussed more fully in Chapter V, Part 3. In addition, high strap muscle activity for the vowel /aa/ occurs across all tone environments irrespective of F_0 characteristics. Strap activity with /aa/ has been noted in earlier EMG experiments and is thought to be associated with jaw opening (cf. Chapter III).

Speaker CT (Figure 5.10) has reasonably high amplitude for the sternothyroid muscle only, but the typical pattern of increased activity during low F_0 can be seen here also. In general, although there are interspeaker differences, the tendency is for strap muscle activity to occur in conjunction with low F_0. The relation of these muscles to F_0 is examined more thoroughly in the following correlation analysis.

B. Correlation of EMG with F_0 (Method of Correlation). Analysis of the correlation between muscle activity and F_0 might ideally involve continuous cross-correlation of the two signals. Such analysis could not be done at the time of this study, since a coordinated F_0/EMG computerized analysis program was not yet available. An approximate measure of correlation based on a smaller sample size is possible. EMG activity should be correlated with F_0 values somewhat later to account for the mechanical latency of muscle activity. The latencies used in this study are based on contraction time[1] values reported in the literature for other species (e.g., Sawashima, 1974), corrected for the larger

[1] The term "contraction time" specifically refers to the period during which work is done by the muscle and can be assessed only by examination of excised muscles. The term "latency" as used here refers to the time lag between observed muscle activity and observed acoustic event.

size of the strap muscles in man. The values agree with those of mean response time[2] determined recently by Atkinson (1973a). In his electromyographic study of the role of laryngeal muscles in regulating F_0 pattern of English intonation, Atkinson measured the latency between observed muscle activity and observed F_0 by shifting the EMG data in incremental steps relative to the F_0 data and calculating the correlation coefficient for each successive shift. The "shift" (incremental step) which showed the highest correlation was determined to be the latency value. There was a range within the incremental steps for which was seen the same high correlation: the values used in this study fall within this range. Values for thyrohyoid were not reported by Atkinson (1973a) but a smaller value was selected for this muscle since it is smaller than the other strap muscles. Table 3 shows the values used in this study compared with those reported by Atkinson and by investigations in other species.

TABLE 3: Latencies used in this study compared with values reported by other studies.

	Erickson	Atkinson	Other studies	
Muscle	latencies (msec)	M.R.T. (msec)	C.T. (msec)	species
Vocalis	15	15	14-21	cat, dog
Cricothyroid	50	40	30-44	cat, dog
Sternohyoid	100	120	50	cat
Sternothyroid	100	70	no data	
Thyrohyoid	50	no data	52	dog

It is assumed that muscle latencies are consistent across different speakers and constant throughout the course of the syllable. Since latency is primarily a function of muscle size, the latter assumption

[2]The term "mean response time" is used in the same way as "latency."

is reasonable, though it may be that a relaxing gesture is more rapid than a tensing gesture. (This will be explored further in a study in progress by M. Liberman and Erickson.) Across speakers, there may be differences due to variance in larynx size, especially across sex, but data are insufficient to permit adjustment for this.

Since the analysis is limited by the number of samples that can be handled, a question exists regarding the number of samples necessary to give reasonable pictures. Clearly, in the case of a static F_0 trajectory one sample is sufficient. After removal of the F_0 perturbations from initial consonants and syllable final effects, the static tones show a midsyllable region of relatively static F_0. A sample can be taken anywhere in the midsyllable region. A point 150 msec into the syllable was selected as lying within midsyllable region for all static F_0 syllables of all speakers.

In the case of a dynamic (time varying F_0 trajectory) more samples are needed to include all regions of interest, i.e., positive and negative peaks as a minimum. There are two possible approaches to sampling the dynamic F_0 trajectories; one is to select a fixed F_0 point, i.e., positive and negative peaks; the other is to carefully select a fixed point within each of the regions surrounding the positive and negative peaks. If the first approach is used, a discontinuous (i.e., bimodal) plot results. (This is a consequence of selecting two points which have relatively the same high and low F_0 values; i.e., if there is a positive correlation between muscle and F_0, such an analysis will result in grouping high F_0 with high muscle activity levels and low F_0 with low muscle activity values.) To fill in the plot more samples would be necessary. Using the second approach, a more complete scatter plot results for the same sample size. (This is because the two points will show a wider range of F_0 values, and this a greater scattering of points.) The second approach was selected; two points were chosen along the F_0 contour, one at 100 and the other at 300 msec into the syllable. These points were found to lie within the region of positive and negative peaks for all dynamic F_0 contours and all speakers.

There are 27 syllable types for the static tones, and 18 for the dynamic tones, a total of 45 syllable types for each speaker. All data are plotted in the form of scatter graphs. As mentioned above, the syllable types which show relatively weak levels of activity (50 μV or less) will be excluded from the scatter plots.

Vocalis/F_0

Scatter plots of vocalis activity versus F_o for speaker MJ for static and dynamic tones are displayed in Figure 5.11. (The other speakers had weak levels of activity and are not analyzed in this fashion.) A strong correlation appears for static but not for dynamic tones. The dynamic tones show some correlation between F_0 and vocalis activity in the F_0 range 125 to 180 Hz, however, even in this range, the correlation appears stronger for static than for dynamic tones. This distinction needs further investigation.

Other muscles examined in this study do not show differences with respect to the dynamic/static F_0 distinctions; therefore, data for the two categories of F_0 patterns are combined in the F_0/EMG scatter plots presented for the other muscles.

Cricothyroid/F_0

The results show systematic increase in cricothyroid activity with increased F_0. This is seen in the scatter plots for speakers MJ and PT (Figure 5.12). Here the graphs show a positive correlation with cricothyroid activity and F_0. Superficial examination of the two graphs suggests that speaker MJ shows two groupings whereas speaker PT shows a single grouping. This may reflect a difference in articulatory strategy between these speakers. This possibility is discussed at length on page 66 and 67.

Strap muscles/F_0

The results show an increase of strap muscle activity with decreased F_0 for each of the strap muscles. Thyrohyoid versus F_0 is displayed in Figure 5.13 (speakers PT and MJ); sternohyoid versus F_0 in Figure 5.14 (speakers SS and MJ); and sternothyroid versus F_0 in Figure 5.15 (speakers SS and CT). The plots show a consistent trend across muscles and speakers: more strap activity is evident with low F_0 than with high F_0, and there is greater variability at the high F_0 values. Within the low F_0 range, there appear to be variations in EMG and F_0 correlation: most strap muscles show negative correlation with F_0 in this low range, but one muscle (the thyrohyoid of speaker PT, Figure 5.13) shows a more extreme pattern with the thyrohyoid acting in an all-or-none fashion. This difference is discussed in greater detail below. Also, peaks of activity are seen with high F_0 for speakers SS and CT (Figures 5.14 and 5.15). These occur with the vowel /aa/ on high and falling tones and are associated with jaw opening for this vowel (cf. Chapter III).

Summary of EMG/F_0 correlations

To summarize, F_0 correlates positively with both the vocalis and cricothyroid, but more consistently with the latter. This finding is congruent with other EMG studies of speech which report that the cricothyroid plays a more consistent role in controlling F_0 in speech than does the vocalis (e.g., Sawashima et al., 1969). The strap muscles show negative correlation with F_0; again, this is in agreement with other EMG findings for speech (e.g., Atkinson, 1973a). The relation between strap muscle activity and low F_0 is discussed more fully below.

C. Timing relationship between muscles and F_0. There is strong evidence that the cricothyroid and strap muscles are involved in F_0 changes. The question addressed herein is whether the muscles initiate these changes. To answer this question, an examination of temporal relationships between muscle event and F_0 was carried out. The rationale is that if the muscle event commences after or simultaneously with the F_0 event, the muscle cannot be said to initiate the change in F_0. If, on the other hand, the muscle event preceeds the F_0 event, then it is possible that the muscle initiates the change in F_0. It must be mentioned that the aim of this part of the study is not to assess muscle latencies, since that examination would require examination of timing of tokens for individual F_0 and EMG data. Rather, the aim is to assess the timing of muscle event relative to F_0 in order to state whether a particular muscle initiates the F_0 change.[3] Discussion of timing relations will be based on those utterance types which show clear onsets and

[3]A more precise analysis of timing relations as well as a method for assessment of muscle latencies will be possible when the autocorrelation programs mentioned earlier become available.

peaks of both F_0 and muscle activity, specifically the rising and falling tones. The low and high tones are not examined because the precursor "carrier phrase" which is on the mid tone shows coarticulatory contextual effects with the following tones, and onset of the rise or fall in the tones is difficult to locate.

It has been suggested that the cricothyroid may initiate both rises and falls in F_0, and that strap muscles can initiate falls in F_0. This section examines temporal relations between F_0 rise and onset of cricothyroid activity and F_0 fall and increase of strap activity as well as decrease of cricothyroid activity. The timing relation between vocalis and F_0 rises is not examined since there are no clear positive or negative peaks in the EMG display for this muscle. Also, since all the speakers show the same basic pattern of timing relationships, all the data for each of the muscles are pooled throughout this section. The number of observations included in each of the comparisons in this section is given in the corresponding figure.

For rises in F_0, timing measurements were made for the cricothyroid at the point where it begins to increase in activity relative to onset of F_0 rise in the rising tone (see Figure 5.16). The results, as shown in Figure 5.17 indicate that an increase in F_0 is always preceeded by an increase in cricothyroid activity.

For falls in F_0, timing measurements were made for both the cricothyroid and strap muscles at that point where the cricothyroid begins to decrease and strap muscle activity begins to increase, both relative to the time at which the F_0 begins to fall in the falling tone (see Figure 5.18).

Erickson and Atkinson (1975) used similar measurements for the sternohyoid muscle on individual tokens of the falling tone from a single Thai speaker and of high falling intonation contours from a single English speaker. They found that the drop in F_0 was always preceded by a decrease in cricothyroid activity, and the sternohyoid activity always increased after the beginning of the F_0 fall. The data suggested that F_0 falls in speech are initiated passively, by relaxation of the cricothyroid, and that the sternohyoid is not used to actively initiate F_0 falls in speech, although sternohyoid activity is present during low F_0. The timing measurements of the present study extend to the sternothyroid and thyrohyoid muscles as well as the sternohyoid. The results confirm the findings of Erickson and Atkinson (1975) for the falling tone. The results shown in Figure 5.19 indicate that a decrease in F_0 is always preceded by a decrease in cricothyroid activity, but that the strap muscles begin to increase in activity <u>after</u> the onset of the F_0 fall. In this figure it appears that the strap muscles do not initiate F_0 falls. It must be remembered that only one particular type of contour has been examined for F_0 falls. The strap muscles may show a different relation to F_0 falls for other tones or other frequency ranges. In order to investigate this, the mid tone, which has a final fall beginning in the mid voice range, was selected for further observation of timing relationship between strap activity and F_0 fall.[4]

Figure 5.20 plots the timing of strap activity relative to the onset of the F_0 fall for the mid tone. In this case the strap activity

---

[4]This was somewhat difficult since the onset of the F_0 fall in such cases is difficult to locate with as much certainty as it was for the falls from high to low F_0 as in the falling tone.

appears to increase before the fall in F_0. These results would indicate that the strap muscles may initiate falls in F_0 when the F_0 drops from a mid F_0 value. From this, one might speculate that strap activity is initiated only whenever the F_0 falls below a certain threshold value and that the strap muscles are active in achieving F_0 values below this threshold. To explore the hypothesis that onset of strap activity is conditioned by a threshold F_0 level, the falling tone was reexamined; this time the reference point was during the fall, specifically that point at which the falling F_0 reached the same absolute F_0 value that occurred on the mid tone for the speaker (see Figure 5.21). Timing measurements with respect to the F_0 fall from this point of the falling tone were made and the results displayed in Figure 5.22. It appears that strap activity begins prior to the mid point of the fall in F_0: after the beginning of the F_0 fall, but before it has progressed past the level of the mid tone.

These findings lend support to a "modal shift" theory of F_0-lowering (see Atkinson, 1973a, b and unpublished). The data suggest that the speaking range can be divided into high, mid, and low voice, and that an F_0 drop from the high to the mid range might be accomplished by relaxation of the cricothyroid, whereas a drop from mid to low range involves increased strap activity as well.

One could speculate about possible trade-off strategies as the F_0 drops below a certain mid point in the voice range. These speculations are based on cricothyroid and thyrohyoid data from speakers MJ and PT discussed in Sections A and B. The first phase of F_0-lowering seems to be the same for both speakers: to drop from high to low F_0 the cricothyroid, greatly tensed at high F_0, relaxes. The thyrohyoid is relaxed during this phase. The second phase would occur when F_0 reached a certain mid value and the strap muscles went into operation. During this phase different speakers use different strategies to further lower F_0. One strategy appears to be to use an increase in thyrohyoid activity in an all-or-none mode to enter the low range, with graded activity of the cricothyroid used to finely adjust F_0 within the low range, as in the high range. Speaker PT, in the present experiment, appears to use this strategy. F_0 for this speaker corresponds to the level of cricothyroid activity throughout the range of F_0 observed in this experiment, while thyrohyoid activity appears only in the low range. In this range F_0 appears uncorrelated with gradations in thyrohyoid activity.

Thus, speaker PT appears to be a good example of what is meant by a "modal shift." The thyrohyoid muscle appears to be used in much the same way as the gearshift in an automobile, while the cricothyroid muscle is used in a manner analogous to the accelerator. This is clearly different from the simple statement that the thyrohyoid is active at low F_0. The latter pattern, which is perhaps analogous to a car with an accelerator and a brake but not gearshift, is shown by speaker MJ.

Speaker MJ's strategy is to use strap muscle activity to enter the low range, but also to use graded activity of these muscles for fine tuning within this range. In the low range, F_0 is relatively uncorrelated with cricothyroid activity (which appears to be held at a minimal level), though it is strongly negatively correlated with thyrohyoid activity in the low F_0 range. (Compare the data on these speakers for cricothyroid activity in Figure 5.12 and for thyrohyoid activity in Figure 5.13.) Since these speakers are the only two for whom good cricothyroid records were obtained it is impossible to generalize these strategies from this data base.

As well, it is possible to speculate about a threshold level of F_0 at which the strap muscles begin to actively lower F_0. With this in mind, the scatter plots of EMG/F_0 correlation for all tones (Figures 5.13-5.15) were examined to see if there may be an F_0 value below which the strap muscles become extremely active. In the table below are displayed the F_0 values for each speaker and strap muscle obtained by examining the scatter plots. Mid point values within the voice range for each speaker are also shown. It would seem one can speak of a certain threshold F_0 value at which the strap muscles become active, which value, moreover, is approximately midrange

TABLE 4: Threshold F_0 values in Hz.

Speakers/Muscles	Thyrohyoid	Sternohyoid	Sternothyroid	Mid point of voice range
Speaker PT	135	---	---	152
Speaker MJ	160	160	---	157
Speaker SS	---	225	225	252
Speaker CT	---	---	145	153

In general no simple answer exists to the question of active versus passive laryngeal control of F_0 lowering in speech. It appears that the strap muscles only contribute actively to lowering F_0 when the pitch is to drop below a threshold level. Once within the region of strap muscle activity the pattern remains complex, since different speakers may use different strategies for lowering F_0 further within this low range. This would seem a promising area for further research.

D. <u>Reciprocity of cricothyroid and strap muscles</u>. It is apparent from the above discussion of timing relations between both cricothyroid activity and F_0 rises, and cricothyroid and strap muscle activity and F_0 falls, that there is reciprocity between these muscles during F_0 changes. This relation is displayed in Figures 5.23 and 5.24 for the two speakers MJ and PT. For F_0 rises, there is an initial decrement in strap muscle activity, followed by an increase of cricothyroid activity (Figure 5.23). For F_0 falls, on the other hand, there is first a decrease in cricothyroid activity followed by an enhanced strap activity (Figure 5.24). A physiological explanation may rationalize the observed reciprocity of the cricothyroid and the strap muscles. It is fairly well understood that contraction of the cricothyroid can exert tension on the vocal folds (see Chapter III); it has been suggested that contraction of the strap muscles can bring about lax vocal folds, although it is not clear how this is done. Two theories speak to this. On the one hand, Ohala (1972) has proposed a "vertical tension" theory in which the strap muscles (particularly the sternohyoid) are thought to facilitate a decrease in vertical tension on the vocal folds; this results in flaccid folds vibrating at low frequencies. Specifically, he proposes that the sternohyoid causes the larynx to be lowered, and that this lessens the vertical tension on the folds. (He does not mention the strap muscles other than the sternohyoid.) The other theory which seeks to account for strap activity during low F_0 is the "frame function" theory, proposed by Schilling (1937) and later elaborated by Sonninen (1956). This theory suggests that the strap muscles, specifically the sternothyroid, can relax antero-posterior tension on the folds by decreasing the distance between their cricoid and thyroid attachments, thus thickening the folds (cf. Chapter III).

Conceivably a combination of effects may result: contraction of the strap muscles may simultaneously reduce the vertical tension as well as modify antero-posterior tension of the folds. Nevertheless, there remain many unanswered questions as to the nature of the laryngeal adjustments produced by strap muscle activity with regard to low F_0. It is apparent that the question is complex and cannot be answered by examining only a select few of the laryngeal muscles. A comprehensive study which examines laryngeal strategies in conjunction with various laryngeal forces during low F_0 is now in progress.

Part 2: Linguistic Analysis

In this section are addressed the linguistic questions introduced in Chapter II: (1) What is the nature of the EMG patterns underlying the five lexically contrastive tones; (2) does the EMG evidence support either a segmental or a syllabic description of the tones or both; specifically, do the patterns inherent in the data suggest a form similar to either, both, or neither of these phonological feature systems; and (3) do the patterns support a laryngeal state description of the type proposed by Halle and Stevens (1971)?

The initial section (A) of this part relates to the first question: it examines the data to see specifically whether there are distinct patterns of EMG associated with each of the tones. Section B considers a phonetic feature description in light of the data and the relation of these to the phonological feature descriptions discussed in Chapter II; Section C examines abstract patterns of muscle activity in terms of articulatory features describing states of the larynx characteristic of tone production. A final section (D) discusses the notion of invariance introduced in Chapter II.

A. Laryngeal activity associated with the tones. Noting distinct F_0 contours for each of the tones, as well as correlation of activity of the laryngeal muscles with F_0, distinct patterns of muscle activity are expected for each of the tones. Figures 5.25 and 5.26 show laryngeal activity for the cricothyroid and strap muscles for the five tones on the syllable /buu/ for the two speakers, MJ and PT, who had good cricothyroid data. The vocalis showed steady low level activity for the most part in this experiment, and is not included in this discussion. For ease of display only one strap muscle is shown in these figures because it was shown earlier (Figures 5.10-5.13) that all the strap muscles show essentially the same pattern of activity. Since the thyrohyoid shows high amplitude activity for both speakers, this muscle was selected as the representative strap muscle.

The five tones can be generally characterized by patterns of cricothyroid and strap muscle activity which are typical for both speakers. The mid tone has relatively steady cricothyroid activity, with a final peak of strap muscle activity corresponding to the fall in F_0 at the end of the syllable. A slight dip in cricothyroid activity occurs with the rise in strap activity.

The low tone generally is characterized by a peak of strap muscle activity which coincides with the F_0 drop of the tone. The cricothyroid shows a dip at this point. The level of cricothyroid activity for the low tone is generally somewhat lower than that for the mid tone for speaker PT; for speaker MJ, on the other hand, the cricothyroid level is about the same for the two tones.

The high tone is characterized by increased cricothyroid and decreased strap muscle activity. One speaker (PT) shows a final peak of thyrohyoid activity, which occurs with increased vocalis activity and may contribute to the final glottal constriction for this speaker.

For the falling tone, the cricothyroid shows an increase in activity shortly before the rise in F_0, which is followed by a peak of strap muscle activity coinciding with the fall in F_0. The rising tone shows the inverse: a peak of strap muscle activity coincident with the initial F_0 drop succeeded by elevated cricothyroid activity shortly prior to the F_0 rise. During the initial drop in F_0 there is also a decrement in cricothyroid activity.

In summary, the cricothyroid and strap muscles together present a unique and consistent configuration of activity for each of the five tones.

One of the original purposes of the study was to contribute to the understanding of the laryngeal mechanism for F_0 lowering. As discussed in Chapter III, the issue is whether F_0 lowering (as in the low part of the falling tone in Thai) is accomplished entirely passively, by relaxing the cricothyroid, or actively, by tensing the strap muscles. An early hypothesis, discussed in Erickson (1975b) was that cricothyroid activity alone was sufficient to account for F_0 lowering, at least in the range of pitches observed in these experiments. Thus it was believed that the cricothyroid muscle alone can effectively differentiate all of the tone types in Thai.

These conclusions were based on examination of data from speaker PT. Later analysis of the data from speaker MJ showed overlap between the distributions of peak activity levels for the mid and low tones. On reanalysis of the data for speaker PT, similar overlap in cricothyroid activity level distributions was found for these tone types. This overlap is present even in cases where the F_0 levels measured for the same tokens do not overlap. Figure 5.27 illustrates this finding. For speaker PT cricothyroid levels are correlated with F_0 levels, even in the lower range of F_0 containing the mid and low tones but by the stronger criterion of nonoverlap of distributions it becomes clear that some other muscle must be contributing to the distinction between the mid and low tones.

Further examination of the corresponding distributions of strap muscle activity for the mid and low tones showed no overlap for either speaker. Figures 5.28 and 5.29 show distributions of cricothyroid and strap muscle activity for the static tones plotted against each other for the two speakers. Characteristic levels of EMG activity were determined in the following way: for the static tones, a peak or level of activity in the midsyllable region was selected, specifically, at 100 msec after release of stop and onset of phonation. The sample size for the static tones is 18 utterance types per speaker (cf. Chapter IV). There were nine syllable types per tone.

From these figures it is apparent that while there is some overlap in the distributions of cricothyroid activity for the low and mid tones for both speakers, there is no overlap in the distributions of strap muscles activity. Thus it appears likely that the strap muscles are involved in actively lowering F_0 during the production of the low tone, and that this activity is an important factor contributing to the distinctiveness of the F_0 patterns for the mid and low tones.

These considerations are pursued further in the next section.

B. Discussion of phonetic features. This section examines how the characteristic patterns in the EMG data relate to the phonological features proposed for the Thai tones. Specifically, the question addressed here is whether the EMG data support either a segmental description of the type proposed by Gandour or a syllabic description of the type proposed by Abramson. Also, examined here is the question of whether or not the patterns support a laryngeal state description of the type proposed by Halle and Stevens (1971).

As defined in this study, the syllabic approach treats tones as wholistic units, the domain of which is the syllable itself. The segmental approach treats tones as segmental units, and in the case of

Gandour's description, as matrices of two features, which can combine to make the five lexically contrastive tones. The domain of the segmental tone is the vowel or vowellike element of the syllable (cf. Chapter II).

Support for a segmental description of the tones is seen in the successive positive and negative peaks of both cricothyroid and strap muscle activity underlying the dynamic tones. For example, the falling tone shows cricothyroid activity with first a positive peak followed by a negative peak and strap muscle activity with first a negative peak followed by a positive peak. The pattern of change in activity levels for the two muscles corresponds exactly to the pattern of feature assignments which characterize the dynamic tones in a segmental description. This is elaborated upon shortly. In addition, the static tones show peaks of muscle activity which correspond to patterns of their phonological feature assignments: the high tone shows a positive peak of cricothyroid activity and a negative peak of strap muscle activity; the low tone shows the opposite pattern; and the mid tone, for one speaker (PT) shows neither positive nor negative peaks of either muscle during the initial portion of the utterance. For the other speaker (MJ), the cricothyroid shows low activity for both the mid and low tones.

The peak of strap muscle activity toward the end of the mid tone would appear to be related to the final fall which is characteristic of this tone in prepausal position. Since this tone typically shows final fall only in prepausal position, it has generally been assumed that the fall was an effect of decreasing subglottal air pressure at the end of a breath-group. This finding of strap muscle activity in connection with this final fall suggests that the laryngeal muscles participate in such nonphonemic F_0 changes as well as in F_0 changes underlying phonemic distinctions. Of course, it may be that this prepausal fall is phonologically relevant after all, but this will not be known until an adequate analysis of Thai sentence intonation is available. It is unlikely that this final fall is an intrinsic aspect of the mid tone in Thai, since Abramson (1974b) has shown that synthetic syllables with identifiable mid tones can be produced with no final fall at all.

In order to further assess the relation of these muscles to the phonological features of tone, the characteristic EMG activity for each of the tones was determined, and the values for cricothyroid and strap muscle activity were plotted against each other.

Static tones

The static tones are displayed in Figures 5.28 and 5.29. Tokens of three static tones fall into three groups according to cricothyroid and strap muscle activity. The high tone is characterized by high cricothyroid activity and low strap muscle activity; the mid tone, by low cricothyroid activity and low strap muscle activity; and the low tone, by low cricothyroid and high strap muscle activity. The terms "high" and "low" activity are relative. The data can be reformulated in the following tables, using + and - to indicate relative levels of high and low activity.

Tones/Muscles	Cricothyroid	Strap Muscles
High	+	-
Mid	-	-
Low	-	+

The similarity of the occurrence of peaks of cricothyroid and strap muscles on tones described by the phonological features [High] and [Low] is striking. This analysis would seem to provide phonetic justification for the two features [High] and [Low] used by Abramson to describe the static tones. It would also agree with Gandour's analysis if his feature [Mid] were changed to [Low]. In fact, he has, partly on the basis of examining these data, indicated that [Low] might be a more appropriate feature.[5]

Dynamic tones

Addition of the dynamic tones to the schema developed for the static tones yields the following results (Figures 5.30 and 5.31). The points in the graph are the positive and negative peaks of muscle activity. There are 27 utterance types per speaker and nine syllable types per tone. The high and low portions of the dynamic tones group with the high and low tones. The descriptive power of the feature system developed for the static tones is not lost when the high and low portions of the dynamic tones are incorporated into the system. Addition of the dynamic tones does not weaken the ability of the cricothyroid and strap muscles to distinguish the categories of tones; the boundaries between the tones are not less clear when the dynamic tones are included.

The data have certain implications about the relation of laryngeal features to phonological features. Specifically, the display of static and dynamic tones together seems to support Gandour's segmental analysis of the tones rather well, given the change in features [Mid]

[5] Jackson Gandour, 1976: personal communication

to [Low]: the dynamic tones can be described in terms of features of the sequence of high and low tones. The same cannot be said for Abramson's feature [Dynamic], inasmuch as this feature has no simple correlate in muscle activity. The definition "abrupt movement in F_0" indicates a consequence of two successive muscle gestures. For instance, in the case of the lexical falling tone the feature [Dynamic] would indicate an initially high cricothyroid activity and low strap activity which changes then to low cricothyroid activity and high strap muscle activity. The fact that the feature [Dynamic] does not have a simple correlate in muscle activity does not disenfranchise the syllabic analysis; but the specific methods employed in the present representation of the data provide no support for an analysis which includes the feature [Dynamic]. Any similar analysis postulating a covert temporal directional feature such as [Dynamic] would fare poorly in light of these data.

Are there instances however in which one is compelled to indicate the feature [Dynamic] in the phonology, and if so, what might be the indications of such on a phonetic level? For instance, it has been claimed that Khru (Elimelich, 1974) is a language for which a dynamic tone must be represented on the phonological level; are there perhaps any indications of such in the EMG data? This question would apply to examination of the Thai dynamic tones as well. One might assume that the feature [Dynamic], in that it carries covert temporal atributes in its definition, is reflected by differences in the nature of the motor commands, which in turn are reflected in differences in the nature of the EMG data. Thus, the feature [Dynamic] would be realized by a different motor command structure than the feature sequence [High] [Low].

One test of this assumption might be to compare the nature of the muscle activity in running speech in (1) the context of a falling tone, and (2) the context of high tone followed by a low tone. Presumably, although not necessarily, if a sequential command constrained by syllable boundary is implicit in the feature [Dynamic], it would be evident from the patterns of muscle activity seen for the two situations. Such an experiment is now in the planning stages.

C. Articulatory features. Formulation of phonetic features in terms of muscle activity per se is not necessarily desirable, since the effect of muscle activity on the acoustical signal may be variable due to postural and other artifacts, as well as interspeaker differences in muscle usage. For example, it is known that the strap muscles behave differently for different postures of the head. Contraction of the strap muscles when the head is in normal position for speech is associated with low F_0, as seen in this study. Sonninen (1956), however, has shown that when the head is bent backward, contraction of the strap muscles can be associated with tense vocal folds. Further, as for interspeaker differences in muscle usage, as pointed out earlier, one speaker (PT) uses the cricothyroid muscle to contribute to differentiation of the mid from the low tone, whereas the other speaker (MJ) does not use the cricothyroid for this function.

One of the goals of physiological research in linguistics is to better ascertain articulatory phonetic features. This laryngeal EMG study of Thai tones was intended to assess the states of the vocal folds involved in production of the F_0 patterns of the tones by providing indirect evidence from the levels of laryngeal muscle activity bearing directly on the determination of these states.

Taking the present state of knowledge of laryngeal physiology, together with the results from correlation of muscle activity with F_0 discussed earlier, and the data displayed in the above figures, it is possible to infer a state, that is, amount of tension, of the vocal folds for each of the tones. It could be inferred, for instance, that for high tone the combination of high cricothyroid activity and low strap muscle activity would produce tense vocal folds; for low tone, the combination of low cricothyroid activity and high strap activity would produce lax vocal folds; and for the mid tone, the combination of both low (or relatively low) cricothyroid activity and low strap muscle activity would produce "normal" folds. The result of these laryngeal states is the appropriate high, low, or mid F_0 characteristics of the tones.

Thus, the following set of articulatory features is suggested by the EMG data for the three phonemic tones in a segmental analysis.

Tones/Features	Tense	Lax
High	+	-
Mid	-	-
Low	-	+

The featuees [Stiff] and [Slack] could be used equally well; however, the features [Tense] and [Lax] are used here to avoid certain implications about consonantal features which are made by the features proposed by Halle and Stevens (1971).

One of the merits of this feature description is that it can account for four tone levels. Although Thai has only three levels of tone, four levels, at the maximum are thought to be available in tone languages (Pike, 1948:5). It has also been suggested that in nontone languages, such as English, there are four levels of contrastive pitch available (Liberman, 1975). Specifically, Liberman, in his acoustic analysis of English intonational patterns, found the four levels, high, high-mid, low-mid, and low. In terms of the laryngeal features proposed above, the high-mid level would be indicated by [+Tense] [+Lax]; the low-mid level, by [-Tense] [+Lax]. Preliminary results in an EMG investigation being carried out by M. Liberman and D. Erickson suggest confirmation of these features: specifically, for the high-mid level there is both increased cricothyroid and strap muscle activity. It may be that the strap muscles act as brakes on the cricothyroid tensing mechanism, although it is not understood from the physiological view point how this might happen.

D. On the notion of invariance. The question introduced in Chapter II, "Onto which aspect of the speech event, acoustic, or articulatory, do the tonal features map more easily?" would seem to be largely one of focus. That is, both the F_0 contours and the EMG activity are phonetic aspects of the tones and both relate to possible phonological descriptions of the tones: a syllabic description of the tones permits a simple mapping of systematic phonetic features onto acoustic events; a segmental description permits a simple mapping of systematic phonetic features onto articulatory events.

With regard to the question of invariance, also raised in Chapter II, the figures (Figures 5.30 and 5.31) show invariance of the phonological features with the muscle activity levels.[6] That is,

[6] The problem of invariance is somewhat different here, where both the phonological analysis and the analysis of the speech event are at issue, than in the more usual case where investigators pursue the acoustic, auditory, articulatory, or electromyographic correlates of firmly established phonological categories.

specifically, the phonological feature [High] has the invariant muscle correlates, high cricothyroid activity and low strap muscle activity. This is true whether the feature [High] applies to the high tone, or the high portions of the dynamic tones. The phonological feature [Low], on the other hand, has the invariant muscle correlates, low cricothyroid activity and high strap muscle activity. Again, this is true whether the feature [Low] applies to the low tone or the low portions of the dynamic tones. The invariant relationship between muscle activity and phonological features is further illustrated in these displays by the nonoverlapping distributions of muscle activity for the phonological features [High] and [Low]. Also, the fact that the mid tone is definable in terms of the distribution of muscle activity is further support for the phonological features [High] and [Low], as well as support for the claim that the muscle activity shows invariance with the phonological features.

It may be asked why invariance has been found here on a level where investigators of other phonetic distinctions have found ubiquitous variability. It seems appropriate to try to answer this question. In terms of the source-filter theory of speech production (e.g., Fant, 1960), one is looking at the principle source of excitation of the vocal tract in this study. The traditional segmental phonemes, and the vocal tract resonances which determine the acoustic representatives of these phonemes, are influenced severally by the muscles of the tongue, the lips, the velum, the jaw, and the larynx, among others. The only anatomical structure which directly influences the fundamental frequency contour, on the other hand, is the vocal folds, and these vocal folds are controlled, in ways which affect F_0, by only a few muscles. It is not, then, surprising to find that the relationship between the phonological features and the laryngeal muscle activity is more direct than that between the features of the segmental phonemes, and considerably more complicated than combinations of muscular activities which create them.

It seems that ubiquitous variability is not as ubiquitous as one supposed. There is an invariant relationship between the tones, on the one hand, and the muscle activity, on the other. One cannot be so confident in a description of the relationship between the tones and the laryngeal states, since knowledge of the laryngeal states is an indirect inference from the muscle activity. One cannot even be sure that the pattern of results would not change in important ways if one examined tones in other languages, or other speakers, or other syllables, or running speech. All one can be sure of is that for these tones, in the Thai language, with these speakers, on the occasions when the utterances were recorded and analyzed by these particular methods, a completely invariant relationship between the Thai tones and the muscle activity of the cricothyroid and strap muscles of the speakers was found. There were no exceptions in the data to this invariant pattern.

Part 3: Historical Notes

The tonogenesis theory discussed in Chapter II, Part 3, argues that low tones emerged from initial voiced consonants, and high tones from initial voiceless consonants. Thus, one might expect to find differences in F_0 and EMG activity for the stop consonants. Indeed, experimental studies have shown acoustic and electromyographic differences associated with the variations in voicing of the consonants (cf. Chapter II).

In light of these considerations, the F_0 and EMG activity for the stop consonants in this experiment were examined in the following manner.[7] The initial F_0 was measured at the onset of phonation after the release of the stop. The EMG peak activity was measured at or slightly before the onset of phonation after the stop release. In certain cases where no peaks were apparent but level activity occurred in the region of the lineup point, measurements were made at the lineup point. In order to examine the differences in F_0 due to differences in initial consonants, the initial F_0 was normalized across vowel and tone types for each set of similar vowel/tone combination and the deviation from the average F_0 for each of these sets was plotted to show the relative distribution of onset F_0 for the different classes of initial stop consonants. The comparisons of muscle activity were prepared similarly.

[7]The F_0 measurements here differ from those of Erickson (1975a) in that the former study looked at initial F_0 only on the syllable types /buu, puu, phuu/ on the mid tone, whereas in this study measurements were made for 45 utterance types (composed of the five tones, and three vowels and three consonants, thus allowing 15 utterance types per stop type).

The results of the examination of initial F_0 of the consonants are displayed in Figure 5.32. Voiced /b/ consistently has a lower initial F_0 than the voiceless stops /p,ph/, and the voiceless aspirate tends to have higher F_0 than the voiceless inaspirate. As concerns the latter point, some inconsistency is evident among the speakers. For instance, speakers MJ and SS show clearly higher mean F_0 for /ph/ than for /p/ whereas speakers CT and PT show about equally high F_0 for /p/ and /ph/. Regarding the tonogenesis hypothesis, the comparison of most import is between the voiced stops on the one hand and the voiceless stops on the other. In terms of this comparison the F_0 data for each of the speakers are consistent with the hypothesis. Thus, this study replicates the findings of Gandour (1974b) and Erickson (1975a).

Given the tendency for voiced stops to have lower F_0 than their voiceless cognates, muscle activity at or near the stop release was examined to detect any correlations with F_0. Based on EMG experiments with English and Hindi stops, it is expected that increased cricothyroid activity for voiceless stops and increased strap muscle activity for voiced stops might be seen with Thai stops (cf. Chapter II). This, of course, would be consonant with the tonogenesis theory in that cricothyroid activity would be associated with both voiceless stops and high tones, and strap muscle activity with both voiced stops and low tones.

Results of the EMG investigation of strap muscle activity in initial stops are displayed in Figures 5.33 and 5.34. A great deal of inconsistency both across and within speakers is evident here. Only one speaker (SS) (Figure 5.33) shows a clear difference in the distributions of muscle activity for different stop types (for the sternohyoid muscle), producing both /b/ and /ph/ with approximately equal low levels of activity, and /p/ showing the highest level of activity. The difference between levels of activity for /b/ and /p/ is clearly counter to the prediction of the tonogenesis hypothesis. Other speakers show no clear differences across consonant categories, and what slight differences are present are in the wrong direction, from the point of view of the hypothesis.

As for cricothyroid activity associated with stop consonants, Figure 5.35 displays these results. No consistent differences in muscle activity corresponding to the different consonants are apparent, although there is a tendency for /ph/ and /b/ to show greater activity than /p/. That /b/ shows a similar pattern of activity with /ph/ in this study is not explicable at this time. Again, the small differences found are, if anything, contrary to the predictions of the tonogenesis hypothesis.

In conclusion, then, there seems to be a tendency for voiced stops to have initial low F_0 and voiceless stops to have initial high F_0. Voiceless aspirates evince a higher F_0 than the voiceless inaspirates, although this finding is not consistent across speakers. The F_0 values for these consonants are compatible with the tonogenssis theory of low tones emerging from voiced consonants and high tones from voiceless consonants. There is, however, no strong correspondence of laryngeal muscle activity with the observed differences in initial F_0 of the consonants. It would seem that these results, rather then shedding any light, increase the confusion already present in the literature with regard to the question of laryngeal activity during stop production. In conclusion it appears that the present EMG data do not offer support for the tonogenesis theory. The possibility of aerodynamic factors may be mentioned to account for differences in F_0 of the consonants, though no data in this study bear on this.

Part 4: Summary of Results

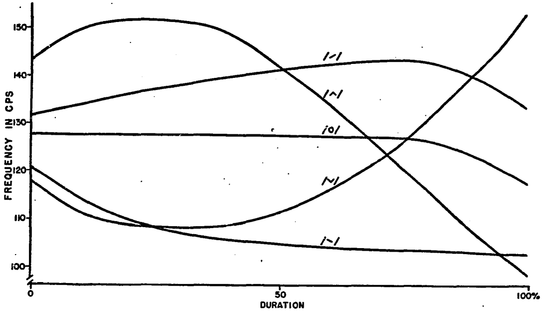
The results of this electromyographic study of Thai tones are discussed in three parts: the first part discusses muscle activity and F_0; the second, muscle activity and tones; and the third, muscle activity and consonants.

The first part is subdivided into three sections: correlation of muscle activity with F_0; timing of muscle activity relative to F_0; and reciprocity between the cricothyroid and strap muscles. Muscle activity correlates with F_0 positively for the cricothyroid, and negatively for strap muscles. Timing of the muscles relative to F_0 shows cricothyroid activity preceding F_0 rises with the cricothyroid suppressed before F_0 falls. This suggests that the cricothyroid can initiate both rises and falls in F_0. The strap muscles show a more complicated relation to F_0 falls. It seems that there is a threshold F_0 value (near midrange) at which the strap muscles become active for F_0 falls. A fall in F_0 from high to low would seem to involve first, relaxation of the cricothyroid, then, increase of strap muscle activity when the threshold F_0 is reached. The data show a reciprocity of the cricothyroid and strap muscles, which is explained in terms of the effect of these muscles upon the vocal folds. The cricothyroid tenses the folds, the strap muscles relax the folds. How the strap muscles might operate in relaxation of the folds is discussed.

The linguistic analysis initially considers the EMG activity for the tones, and subsequently interprets the phonetic data in light of phonological feature theory. In summary, a unique and consistent configuration of laryngeal muscle activity was found for each of the five lexically contrastive tones in Thai. The muscles involved are the cricothyroid and strap muscles. The data support a segmental description in that the dynamic tones can be analyzed as sequences of discrete occurrences of a reciprocal patterning of the cricothyroid and strap muscles. Specifically, then, support for a modified version of Gandour's feature descriptions is seen in the EMG data. This is an extremely interesting finding, especially since the phonological evidence for a segmental description of Thai tones seems not to be conclusive.

Furthermore, it is seen that the Halle and Stevens (1971) laryngeal features have a basis in laryngeal muscle data: relative increases or decreases in cricothyroid activity correspond with the tense state of the larynx (i.e., [±Stiff]), and relative increases or decreases in strap muscle activity correspond with the lax state of the larynx (i.e., [±Slack]). Although Thai has only three tone levels, the possibility of these features handling a maximum of four levels is briefly discussed.

The final section, historical notes, briefly discusses the tonogenesis theory of voiced and voiceless consonants. Examination of F_0 of initial consonants shows that voiced stops tend to have lower F_0 than voiceless ones, but there is no correspondence of these distinctions with the EMG data. The possibility of aerodynamic factors causing these differences in initial F_0 of consonants is mentioned speculatively.



TONES ON DOUBLE VOWELS

Figure 2.1

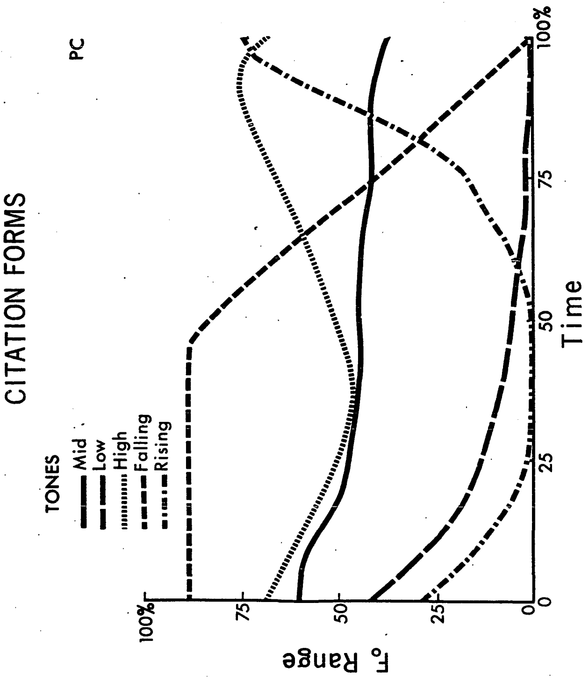

CITATION FORMS
PC
TONES
Mid
Low
High
Falling
Rising
100%
75
50
25
0
F0 Range
0
25
50
75
100%
Time

Figure 2.2

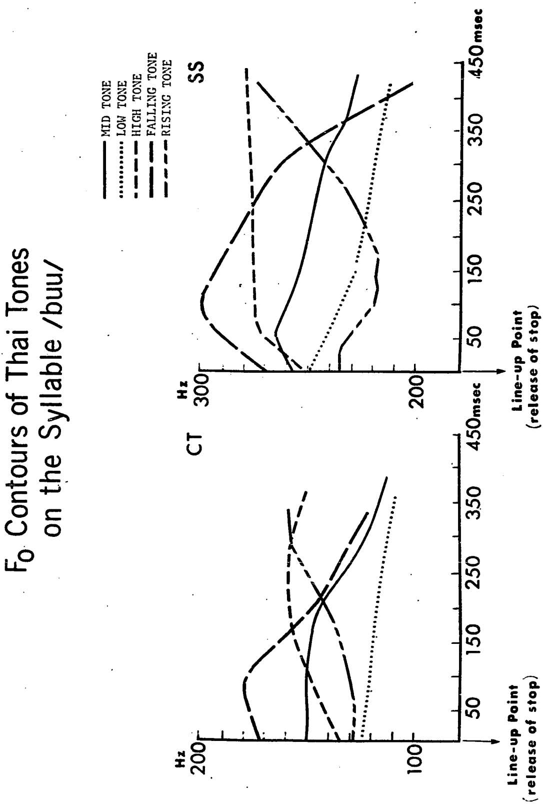

F0 Contours of Thai Tones
on the Syllable /buu/
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
CT
SS
Hz
200
100
Hz
300
200
50
150
250
350
450msec
Line-up Point
(release of stop)

Figure 2.3

F_0 Contours of Thai Tones on the Syllable /buu/

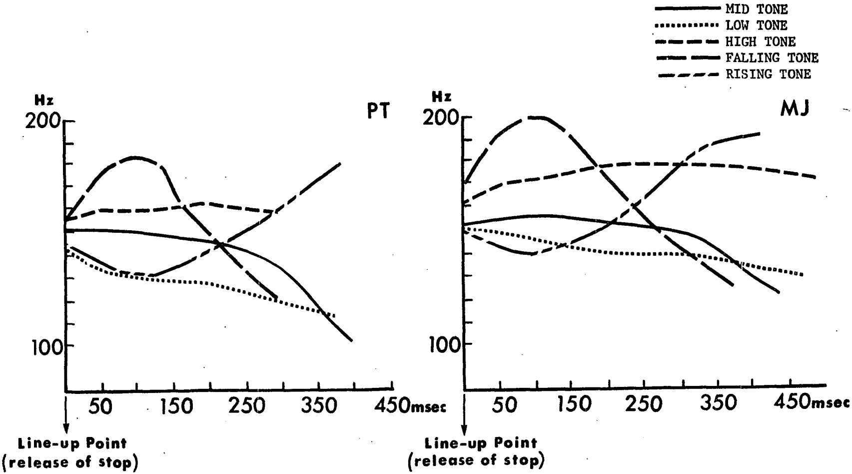


Figure 2.4

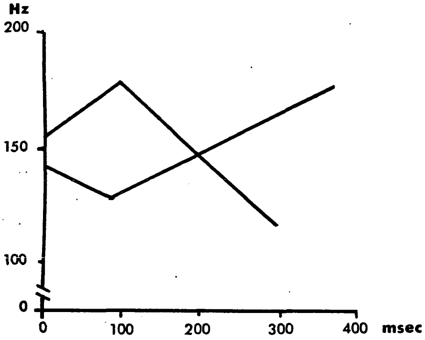
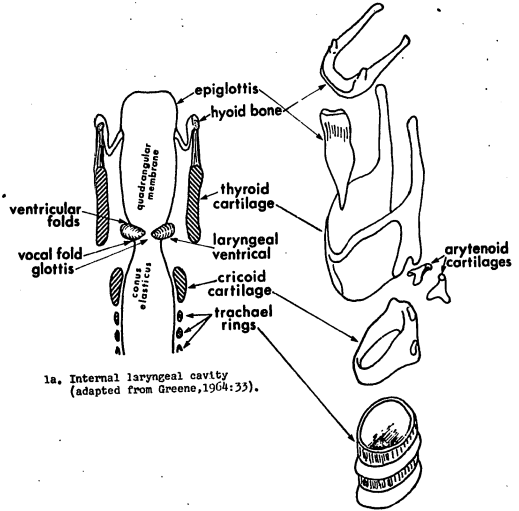


Figure 2.5: Schematic drawing of dynamic tones for speaker PT on the syllable /buu/.



1a. Internal laryngeal cavity
(adapted from Greene,1964:33).

1b. Hard structures of the larynx
(adapted from Shearer, 1968:41).

Figure 3.1: The larynx.

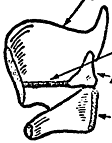


Figure 3.2 Cartilaginous Attachments of the Vocal Folds

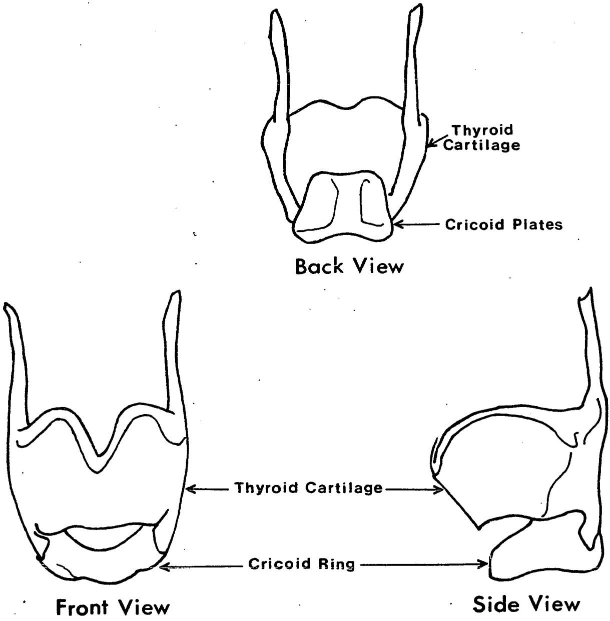


Figure 3.3: Cricoid and thyroid cartilages.

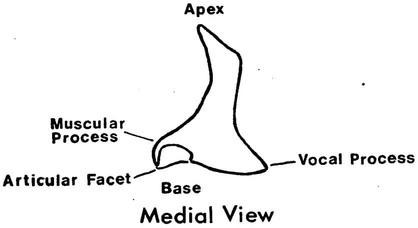


Figure 3.4: Arytenoid cartilage.

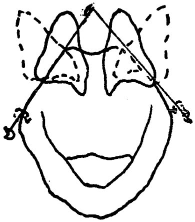


Figure 3.5: Rotating movement of cricoarytenoid articulation (adapted from Sonesson, 1970:53).

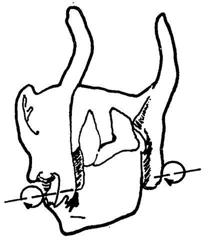


Figure 3.6: Rotation around circothyroid articulation (adapted from Zemlin, 1964:121).

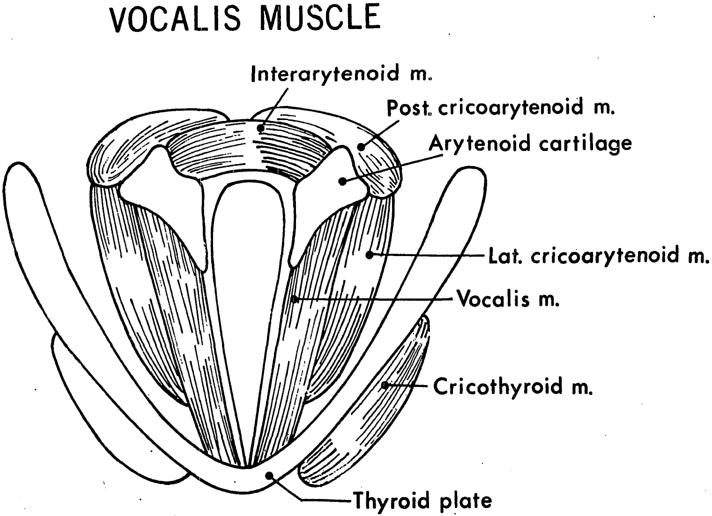


Figure 3.7: Vocalis muscle (adapted from Hirano and Ohala, 1969:370).

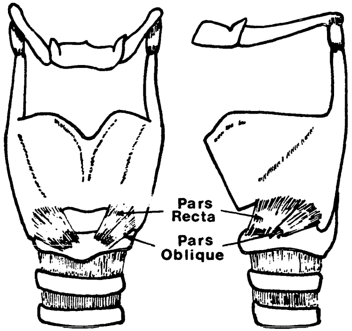


Figure 3.8: Cricothyroid muscle.

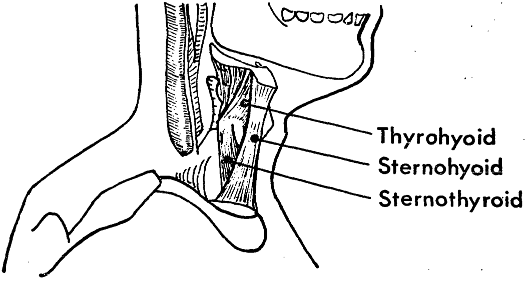


Figure 3.9: Strap muscles (adapted from Friedman, 1950: Figure 22).

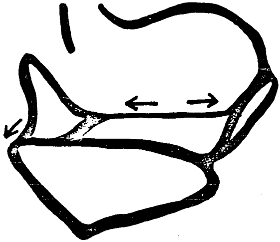
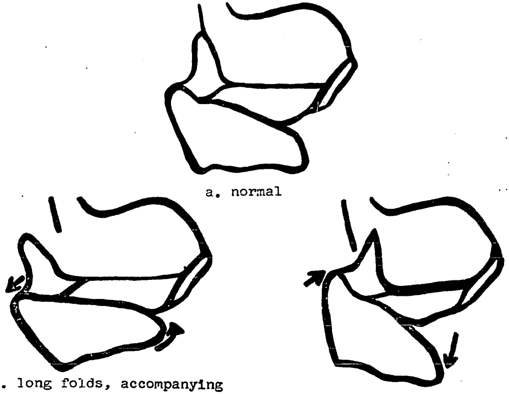


Figure 3.10: Lengthening of vocal folds due to cricothyroid contraction (adapted from Pernkopf, 1963:335).



a. normal

b. long folds, accompanying small crico-thyroid space.

c. short folds accompanying enlarged crico-thyroid space.

Figure 3.11: The crico-thyroid space as an index of vocal fold length and tension (adapted from Pernkopf,1963:335).

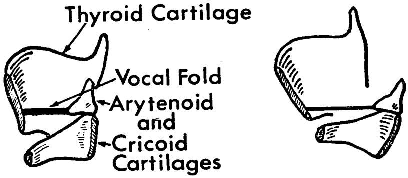


Figure 3.12: Schema indicating the production of a longer and thinner vocal fold upon tilting the thyroid and cricoid cartilages on each other (adapted from Hollinshead, 1954:436).

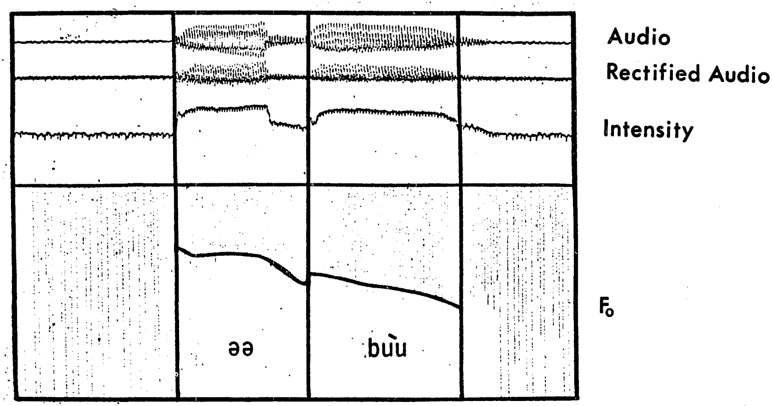


Figure 4.1: Sample of display of Phonetic Analysis System output.

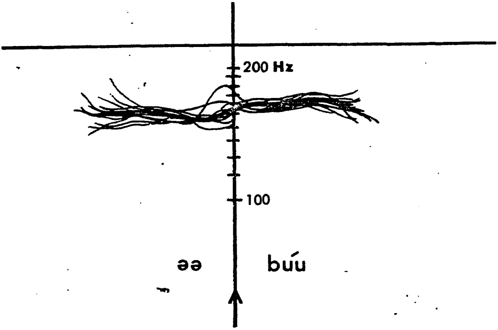


Figure 4.2: Sample of method of selecting an average F_0 contour from a family of curves.

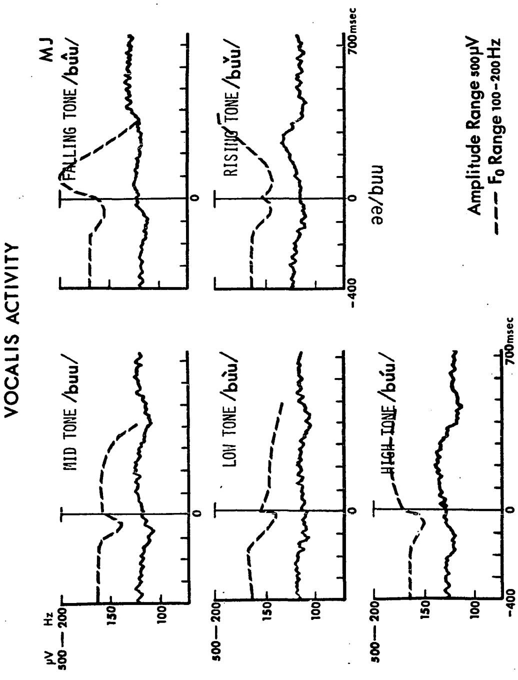

VOCALIS ACTIVITY
MJ
MID TONE /buu/
LOW TONE /bùu/
HIGH TONE /búu/
FALLING TONE /bûu/
RISING TONE /bǔu/
µV Hz
500 — 200
150
100
-400
0
700msec
nng/ee
Amplitude Range 500µV
F0 Range 100-200Hz

Figure 5.1

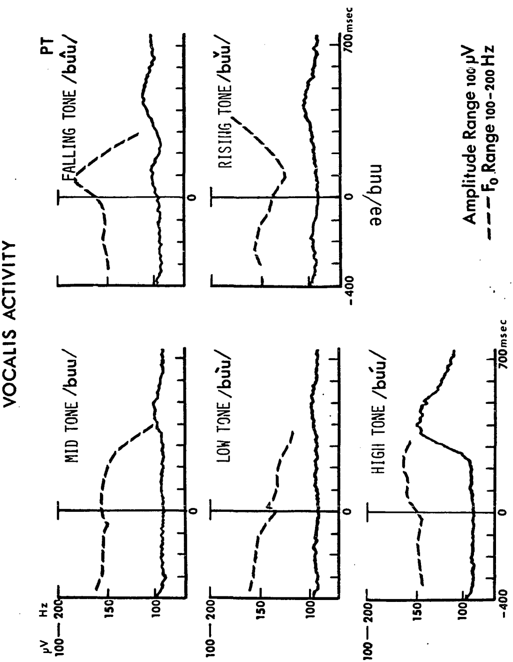

VOCALIS ACTIVITY
PT
MID TONE /buu/
LOW TONE /bùu/
HIGH TONE /búu/
FALLING TONE /bûu/
RISING TONE /bǔu/
μV Hz
100—200
150
100
-400
0
700 msec
Amplitude Range 100 μV
F0 Range 100-200 Hz

Figure 5.2

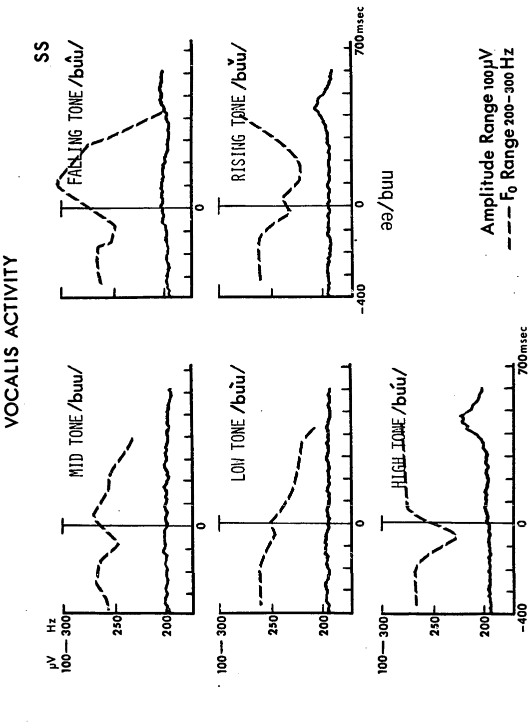
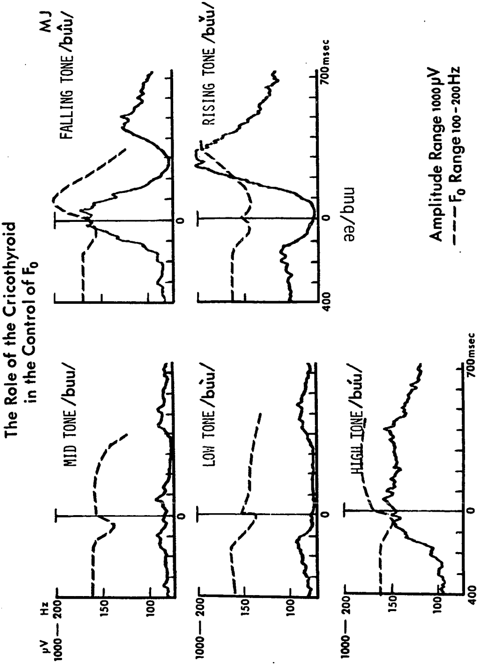

VOCALIS ACTIVITY
SS
MID TONE /buu/
LOW TONE /bùu/
HIGH TONE /búu/
FALLING TONE /bûu/
RISING TONE /bǔu/
µV Hz
100—300
250
200
-400
0
700msec
Amplitude Range 100µV
--- F_0 Range 200-300 Hz

Figure 5.3



The Role of the Cricothyroid in the Control of F_0
MJ
MID TONE /buu/
LOW TONE /bùu/
HIGH TONE /búu/
FALLING TONE /bûu/
RISING TONE /bǔu/
µV Hz
1000 — 200
150
100
400
0
700msec
Amplitude Range 1000µV
--- F_0 Range 100-200Hz

Figure 5.4

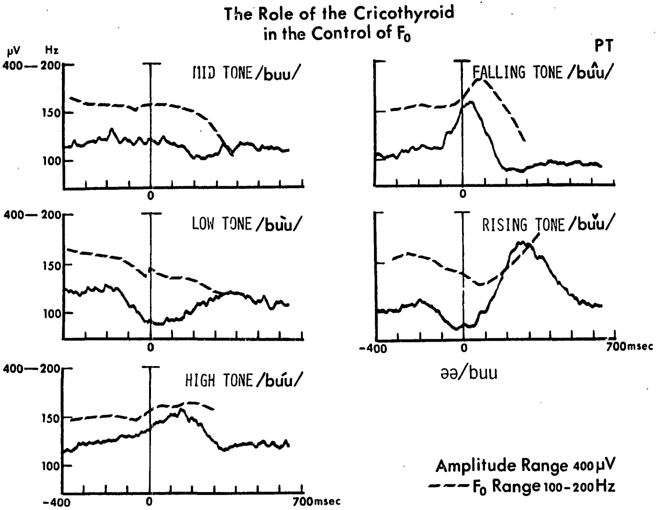

The Role of the Cricothyroid
in the Control of F_0
PT
μV Hz
400—200
150
100
0
-400
700msec
MID TONE /buu/
LOW TONE /bùu/
HIGH TONE /búu/
FALLING TONE /bûu/
RISING TONE /bǔu/
əə/buu
Amplitude Range 400μV
F_0 Range 100-200Hz

Figure 5.5

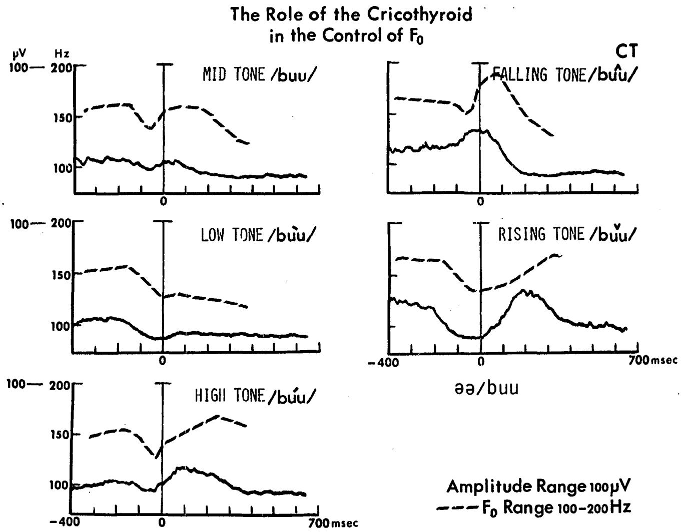

The Role of the Cricothyroid
in the Control of F_0
CT
μV
Hz
100— 200
150
100
0
MID TONE /buu/
FALLING TONE/bûu/
LOW TONE /bùu/
RISING TONE /bǔu/
HIGH TONE /búu/
-400
700 msec
əə/buu
Amplitude Range 100μV
F_0 Range 100-200Hz

Figure 5.6

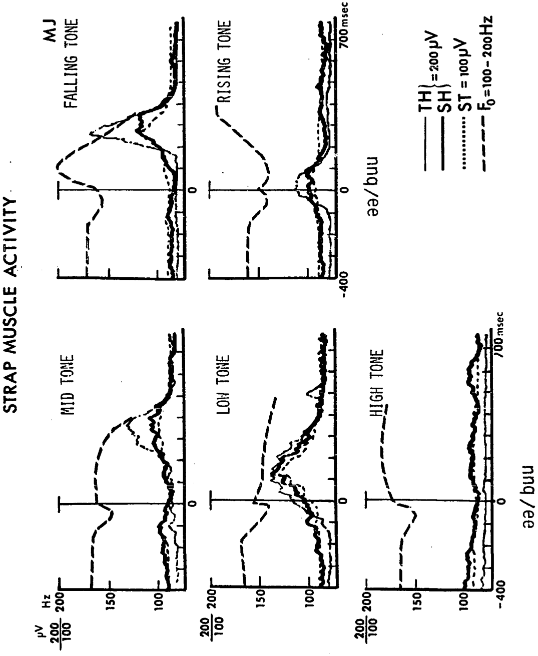

STRAP MUSCLE ACTIVITY
MJ
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
TH
SH
=200µV
ST = 100µV
F0=100-200Hz
-400
0
700 msec

Figure 5.7

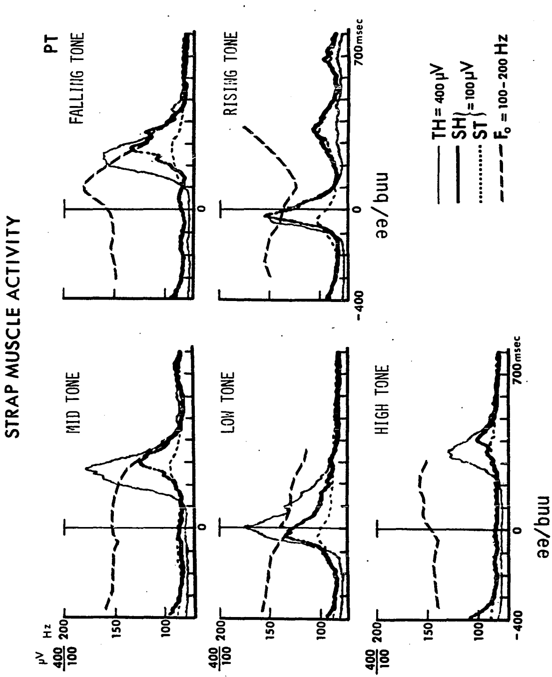

STRAP MUSCLE ACTIVITY
PT
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
TH = 400µV
SH
ST = 100µV
F_o = 100-200 Hz
700msec
-400
0
µV
Hz
400
100
200
150
100

Figure 5.8

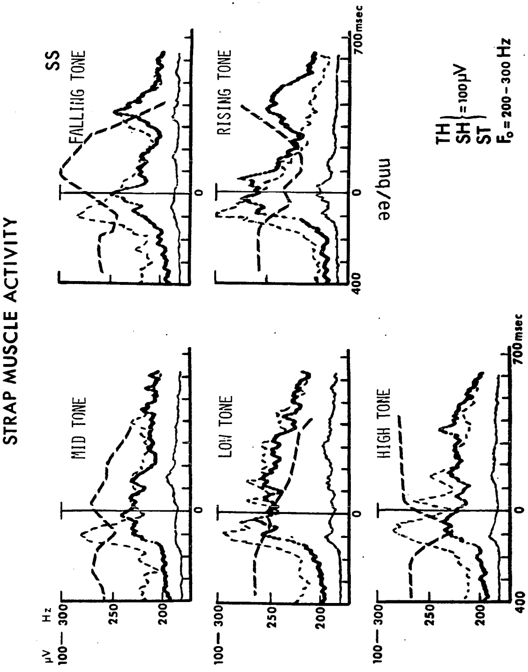

STRAP MUSCLE ACTIVITY
SS
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
μV Hz
100— 300
250
200
400
0
700msec
TH
SH =100μV
ST
F_o=200-300 Hz

Figure 5.9

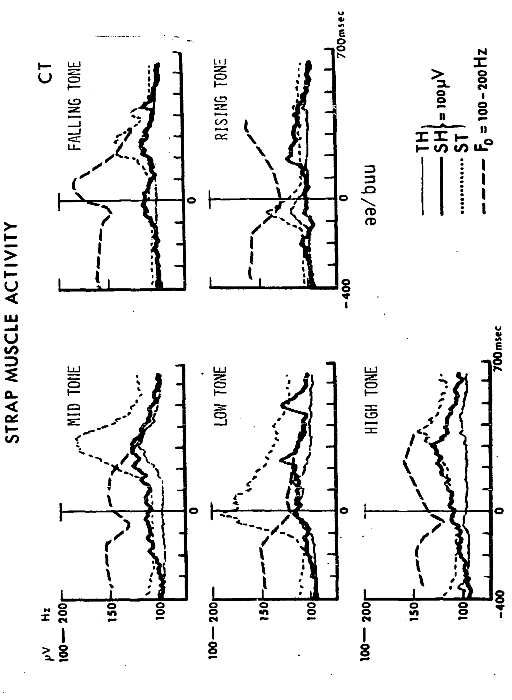

STRAP MUSCLE ACTIVITY
CT
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
µV Hz
100 — 200
150
100
-400
0
700msec
TH
SH = 100µV
ST
F0 = 100-200Hz

Figure 5.10

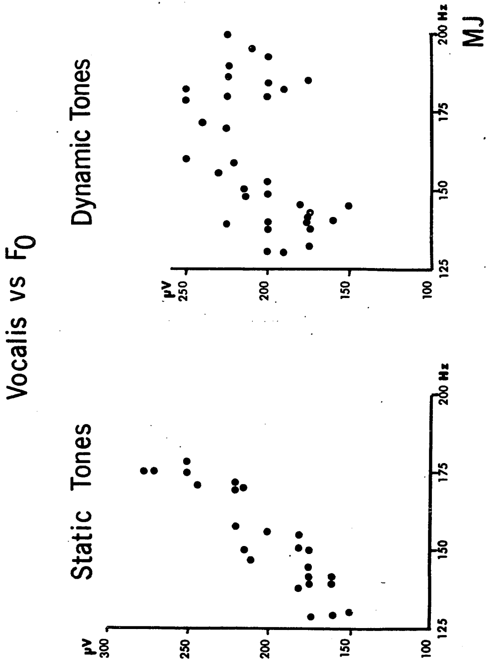
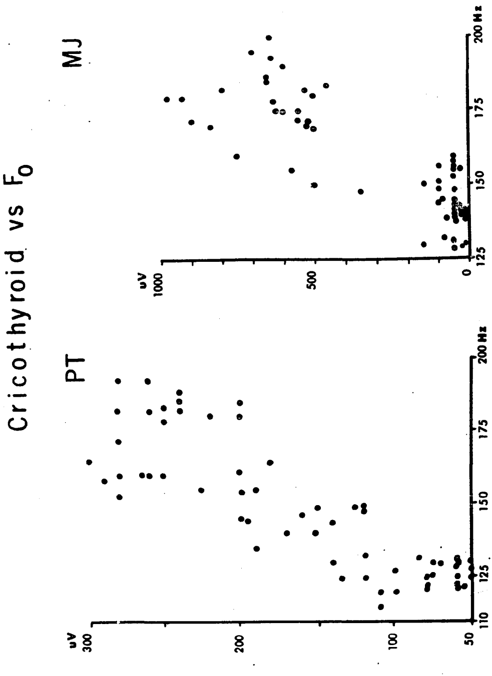

Vocalis vs F0
Static Tones
Dynamic Tones
μV
300
250
200
150
100
125
150
175
200 Hz
MJ

Figure 5.11



Cricothyroid vs F0
PT
MJ
uV
300
200
100
50
110
125
150
175
200 Hz
1000
500
0

Figure 5.12

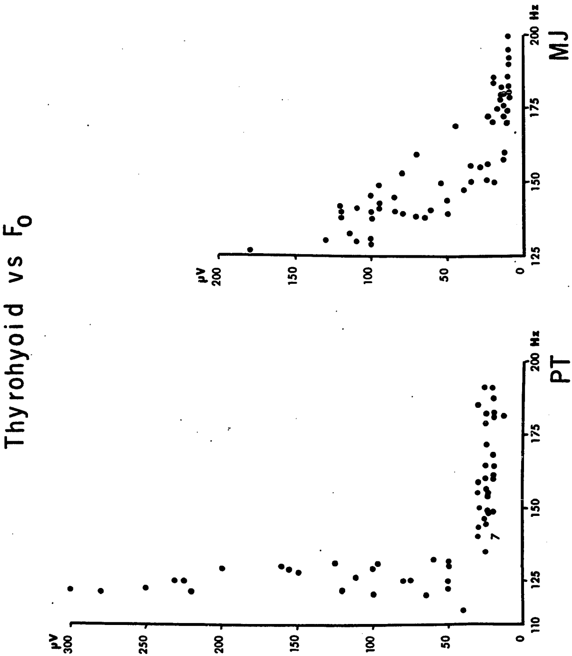

Thyrohyoid vs F_o
PT
MJ
μV
300
250
200
150
100
50
0
110
125
150
175
200 Hz
7

Figure 5.13

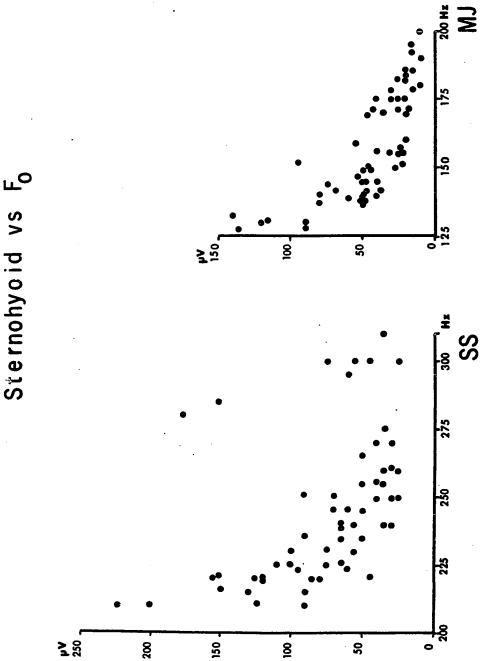

Sternohyoid vs F_0
SS
µV
250
200
150
100
50
0
200
225
250
275
300
Hz
MJ
µV
150
100
50
0
125
150
175
200 Hz

Figure 5.14

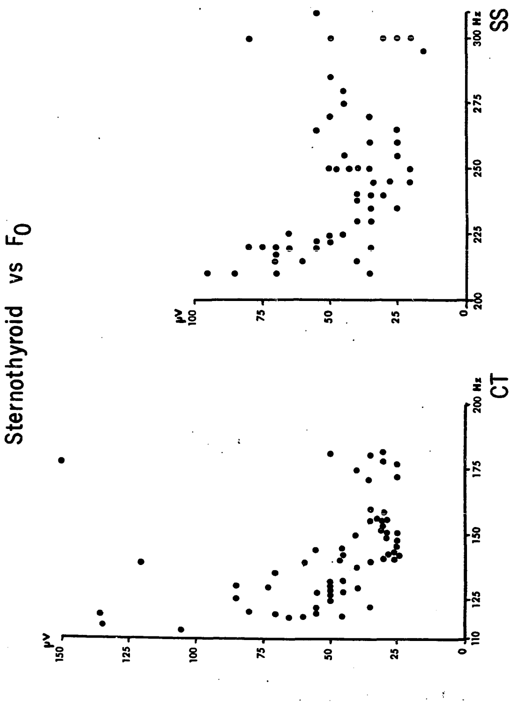
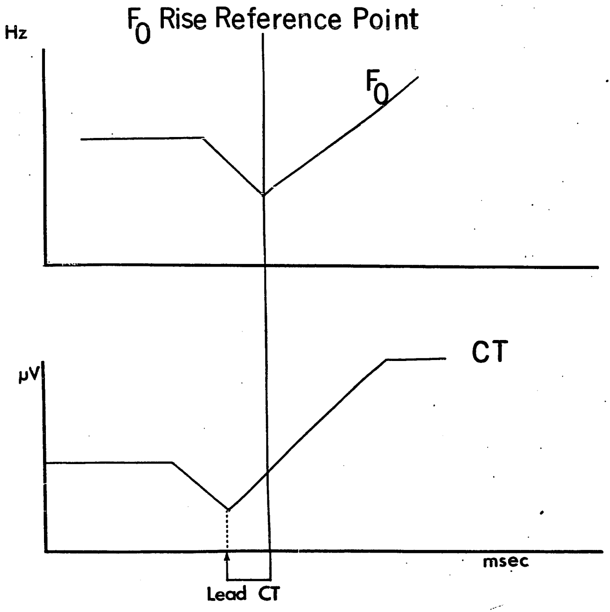

Sternothyroid vs F_0
μV
150
125
100
75
50
25
0
110
125
150
175
200 Hz
CT
μV
100
75
50
25
0
200
225
250
275
300 Hz
SS

Figure 5.15



Schematic presentation of Cricothyroid activity in relation to F_0 Rise

Figure 5.16

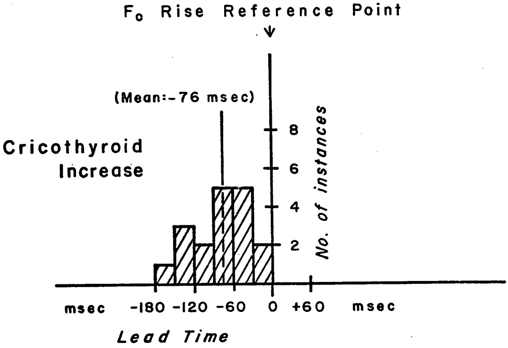
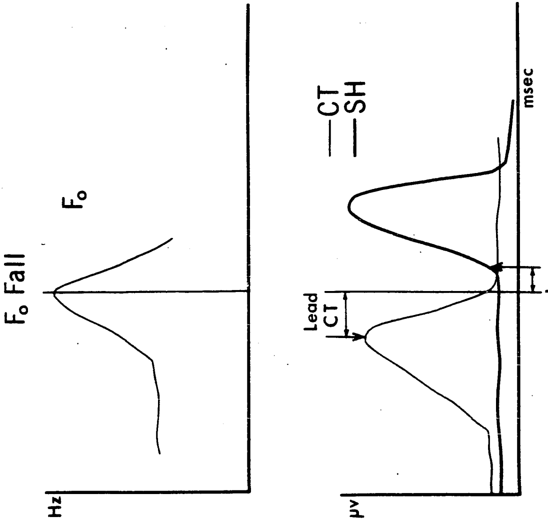


Figure 5.17: Timing of cricothyroid activity relative to F_0 rise (Rising Tone).

Data Base: Cricothyroid, 2 speakers, 9 syllables/speaker.



Schematic presentation of Cricothyroid and Sternohyoid activity in relation to F_o Fall.

Figure 5.18

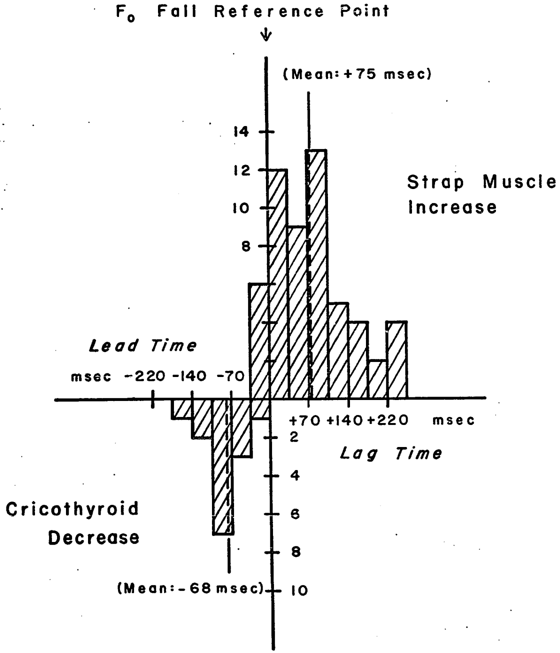


Figure 5.19: Timing of cricothyroid and strap muscle activity relative to F_0 fall (Falling Tone).

Data Base: Cricothyroid, 2 speakers, 5-9 syllables/speaker.
Strap muscles, 4 speakers, 1-2 muscles/speaker, 5-9 syllables/speaker.

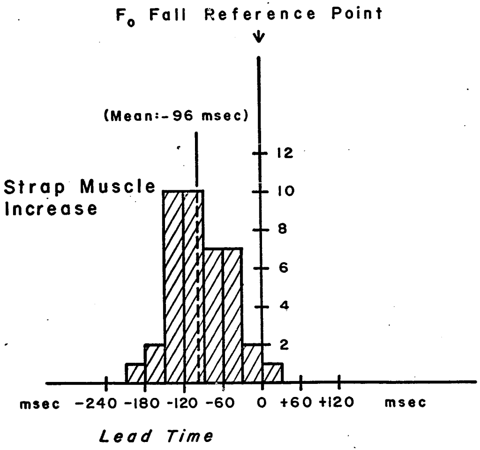
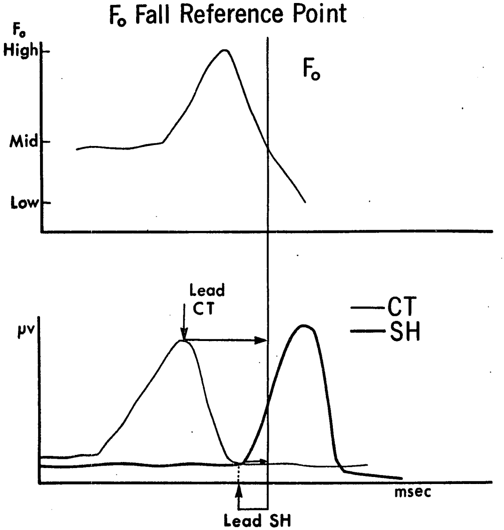


Figure 5.20: Timing of strap muscle activity relative to F_0 fall, mid voice range (Mid Tone).

Data Base: 3 speakers, 1-2 muscles/speaker, 6-9 syllables/speaker.

Figure 5.21



Schematic presentation of Cricothyroid and Sternohyoid activity in relation to F_o Fall from mid point in the Fall.

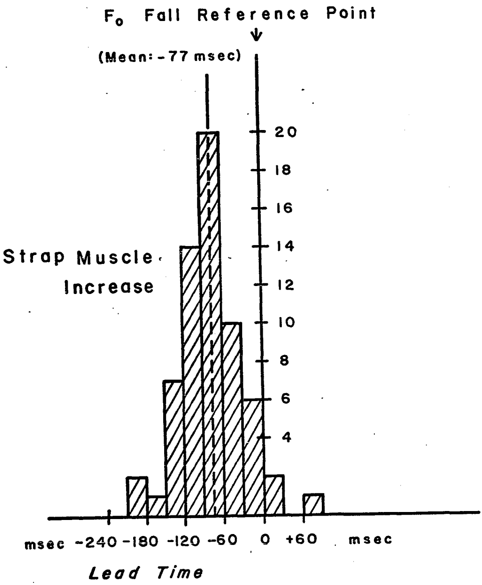


Figure 5.22: Timing of strap muscle activity relative to F_0 fall, mid voice range, (Falling Tone).

Data Base: 4 speakers, 1-2 muscles/speaker, 9 syllables/speaker.

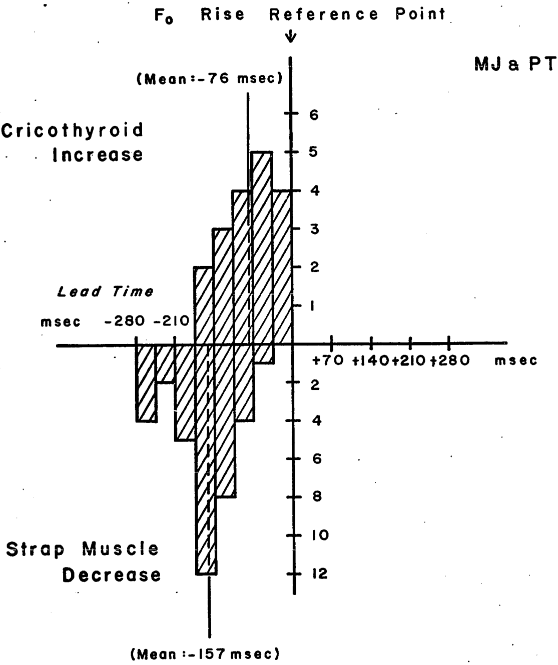


Figure 5.23: Reciprocity of muscles during F_0 rise (Rising Tone).

Data Base: Cricothyroid, 2 speakers, 9 syllables/speaker. Strap muscles, 2 speakers, 2 muscles/speaker, 9 syllables/speaker.

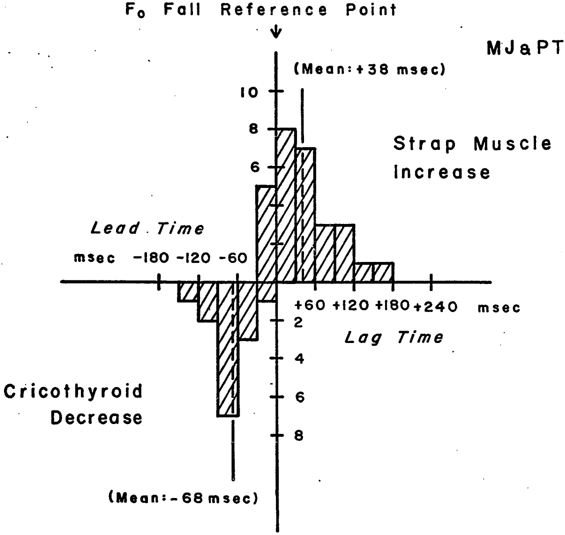


Figure 5.24: Reciprocity of muscles during F_0 fall (Falling Tone).

Data Base: Cricothyroid, 2 speakers, 5-9 syllables/speaker. Strap muscles, 2 speakers, 2 muscles/speaker, 5-9 syllables/speaker.

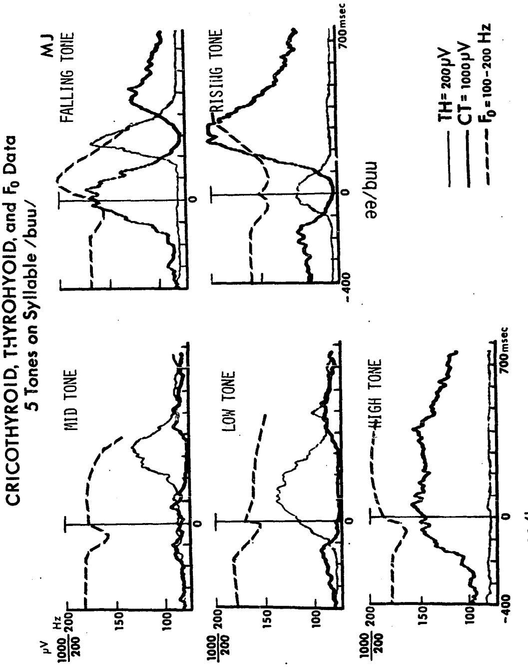

CRICOTHYROID, THYROHYOID, and F₀ Data
5 Tones on Syllable /buu/
MJ
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
/buu/
/ee/
/buu/
/ee/
μV
Hz
1000
200
200
150
100
-400
0
700msec
TH= 200μV
CT= 1000μV
F0=100-200 Hz

Figure 5.25

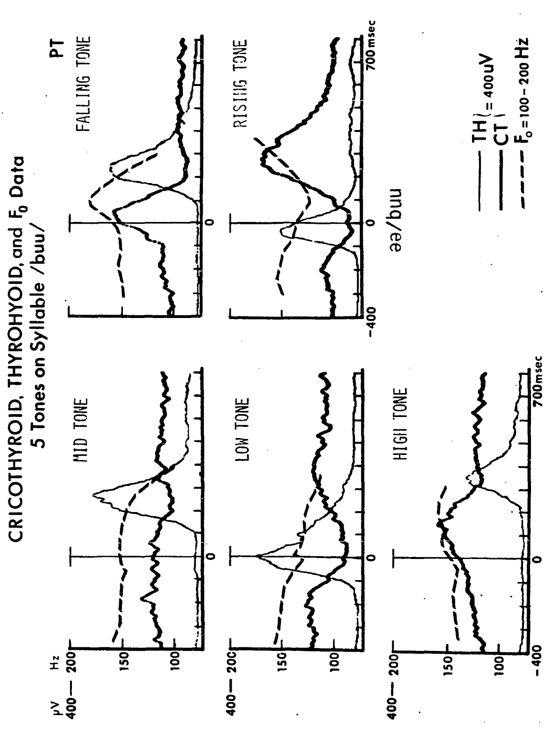

CRICOTHYROID, THYROHYOID, and F_0 Data
5 Tones on Syllable /buu/
PT
MID TONE
LOW TONE
HIGH TONE
FALLING TONE
RISING TONE
μV
Hz
400— 200
150
100
-400
0
700 msec
buu/ee
TH
CT
= 400uV
F_o=100-200 Hz

Figure 5.26

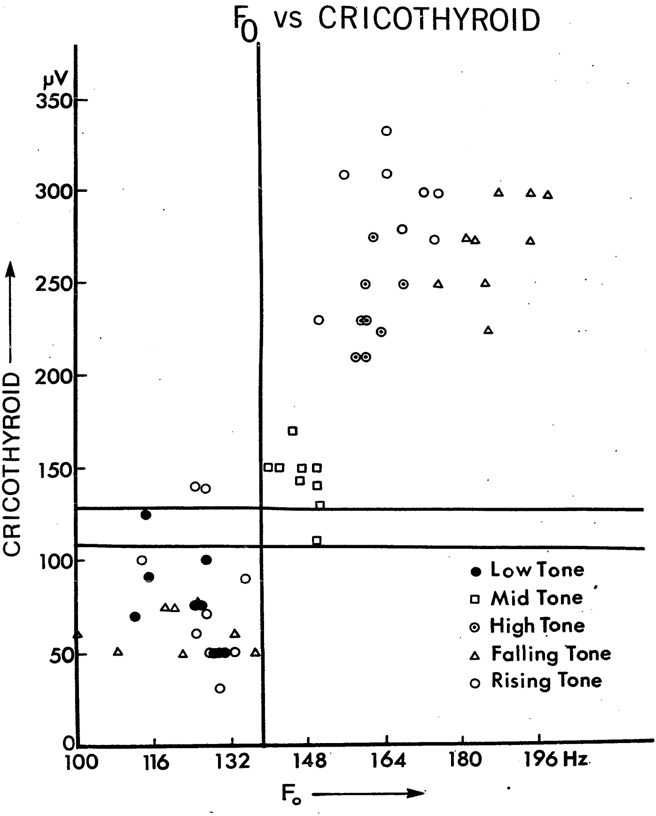


Figure 5.27. (Adapted from Erickson,1975b)

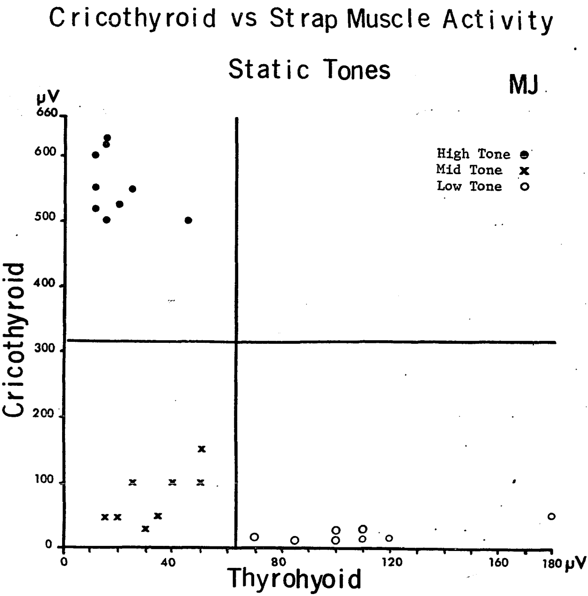

Cricothyroid vs Strap Muscle Activity
Static Tones
MJ
μV
660
600
500
400
300
200
100
0
Cricothyroid
High Tone
Mid Tone
Low Tone
0
40
80
120
160
180 μV
Thyrohyoid

Figure 5.28

Cricothyroid vs Strap Muscle Activity
Static Tones

PT

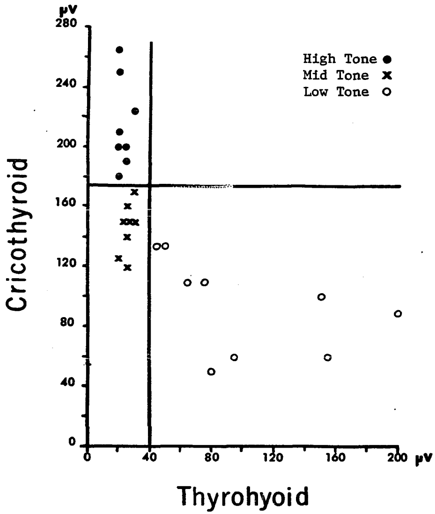


Figure 5.29

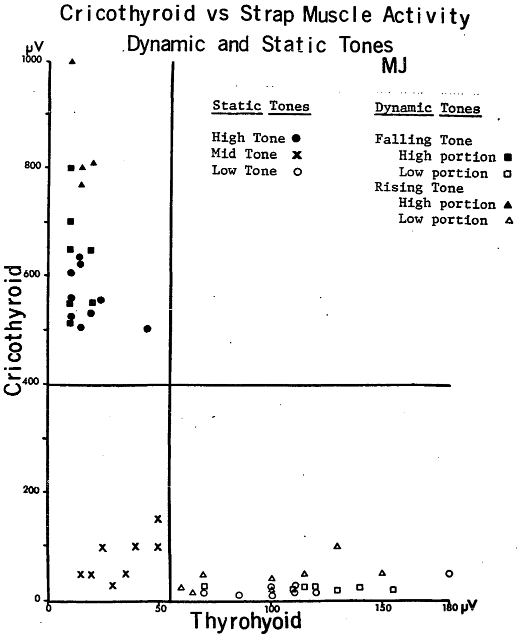

Cricothyroid vs Strap Muscle Activity
Dynamic and Static Tones
MJ
Static Tones
High Tone ●
Mid Tone x
Low Tone o
Dynamic Tones
Falling Tone
High portion ■
Low portion □
Rising Tone
High portion ▲
Low portion △
µV
1000
800
600
400
200
0
Cricothyroid
0
50
100
150
180 µV
Thyrohyoid

Figure 5.30

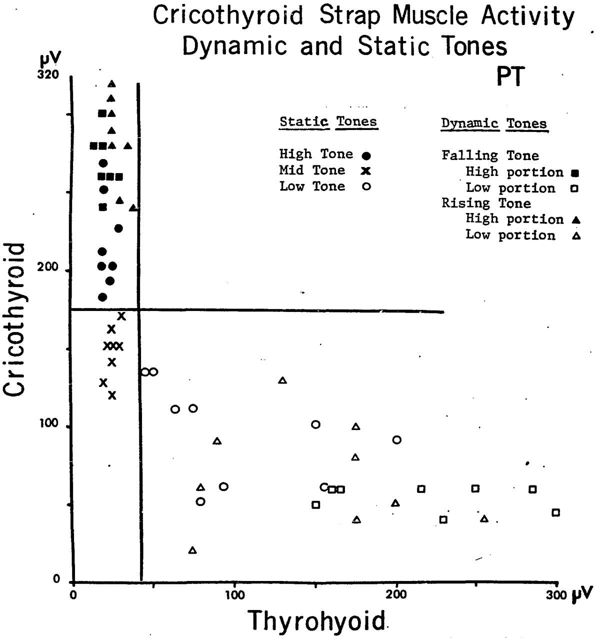

Cricothyroid Strap Muscle Activity
Dynamic and Static Tones
PT
µV
320
200
100
0
Cricothyroid
0
100
200
300 µV
Thyrohyoid
Static Tones
High Tone ●
Mid Tone x
Low Tone o
Dynamic Tones
Falling Tone
High portion ■
Low portion □
Rising Tone
High portion ▲
Low portion △

Figure 5.31

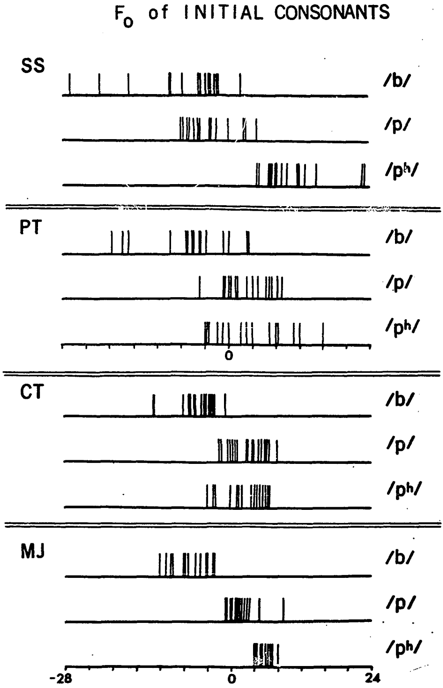

F_0 of INITIAL CONSONANTS
SS
/b/
/p/
/pʰ/
PT
/b/
/p/
/pʰ/
0
CT
/b/
/p/
/pʰ/
MJ
/b/
/p/
/pʰ/
-28
0
24

Figure 5.32

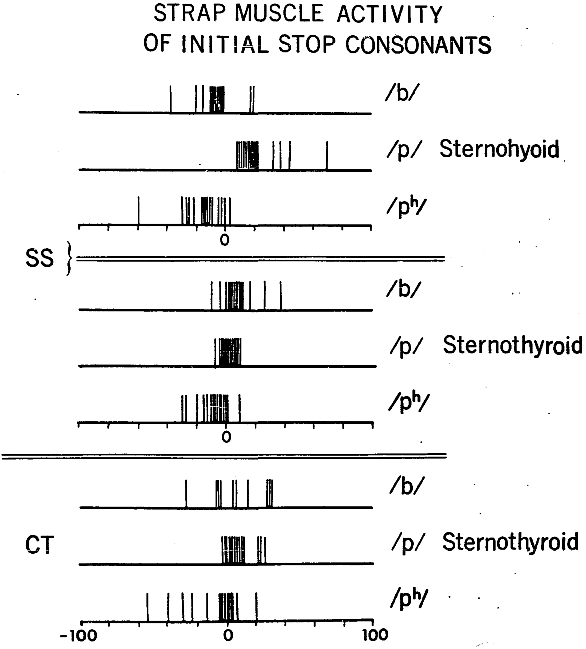

STRAP MUSCLE ACTIVITY
OF INITIAL STOP CONSONANTS
/b/
/p/ Sternohyoid
/pʰ/
0
SS
/b/
/p/ Sternothyroid
/pʰ/
0
/b/
CT
/p/ Sternothyroid
/pʰ/
-100
0
100

Figure 5.33

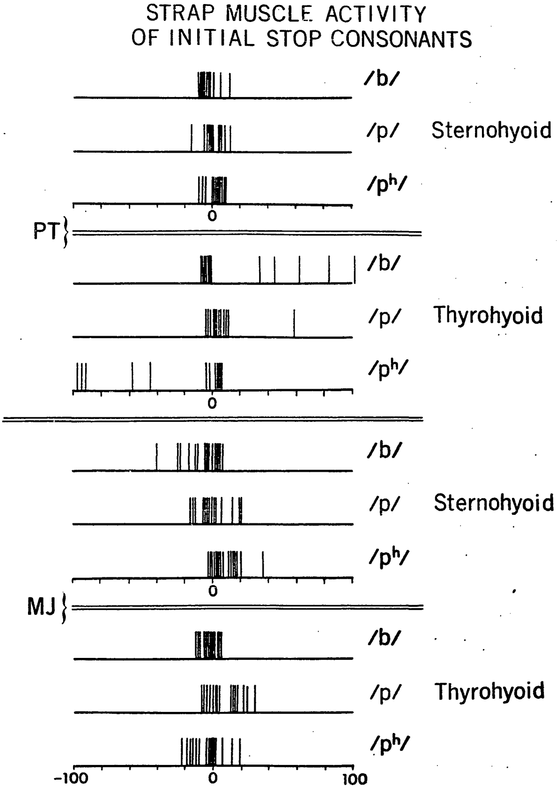

STRAP MUSCLE ACTIVITY
OF INITIAL STOP CONSONANTS
/b/
/p/
Sternohyoid
/pʰ/
0
PT
/b/
/p/
Thyrohyoid
/pʰ/
0
/b/
/p/
Sternohyoid
/pʰ/
0
MJ
/b/
/p/
Thyrohyoid
/pʰ/
-100
0
100

Figure 5.34

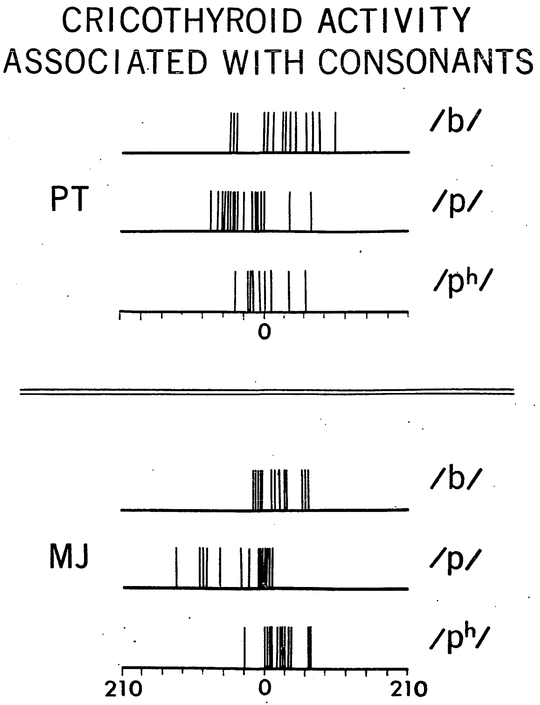

CRICOTHYROID ACTIVITY
ASSOCIATED WITH CONSONANTS
PT
/b/
/p/
/pʰ/
0
MJ
/b/
/p/
/pʰ/
210
0
210

Figure 5.35

REFERENCES

Abramson, A. S. (1962) The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments. (Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics), Publication 20.

Abramson, A. S. (1974a) Pitch in the perception of voicing states in Thai: Diachronic implications. Paper presented at the annual meeting of the Linguistic Society of America, New York, December. [Also in Haskins Laboratories Status Report on Speech Research SR-41 (1975), 165-174.]

Abramson, A. S. (1974b) The tones of Central Thai: Some perceptual experiments. Haskins Laboratories Status Report on Speech Research SR-39/40, 91-104. [Also in Studies in Tai Linguistics in Honor of William J. Gedney, ed. by J. G. Harris and J. R. Chamberlain. (Bangkok: Central Institute of English, 1975), pp. 1-16.]

Abramson, A. S. (1974c) Experimental phonetics in phonology: Vowel duration in Thai. Pasaa, Notes and News about Language Teaching and Linguistics in Thailand 4, 71-90.

Abramson, A. S. (1975a) The coarticulation of tones: An acoustic study of Thai. Haskins Laboratories Status Report on Speech Research SR-44, 119-126.

Abramson, A. S. (1975b) Thai tones as a reference system. Haskins Laboratories Status Report on Speech Research SR-44, 127-136.

Abramson, A. S. (1976) Static and dynamic acoustic cues in distinctive tones. J. Acoust. Soc. Am., Suppl. 59, S42(A).

Ardran, G. M. and F. H. Kemp. (1966) The mechanism of the larynx I: Movements of the arytenoid and cricoid cartilages. Brit. J. Radiol. 39, 641-654.

Arnold, G. E. (1961) Physiology and pathology of the cricothyroid muscle. Laryngoscope 71, 687-753.

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

Atkinson, J. E. (1973b) Physiological factors controlling F_0: Results of a correlation analysis. J. Acoust. Soc. Am. 54, 319(A).

Atkinson, J. E. (Unpublished) A correlation analysis of the physiological factors controlling voice pitch.

Basmajian, J. V. (1974) Muscles Alive. Their Functions Revealed by Electromyography, 3rd edition. (Baltimore: The Williams and Wilkins Co.).

Bell-Berti, F. (1973) The velopharyngeal mechanism: An electromyographic study. Unpublished Ph.D. dissertation, City University of New York. (Issued as Supplement to Haskins Laboratories Status Report on Speech Research.)

Bell-Berti, F. (1975) Control of pharyngeal cavity size for English voiced and voiceless stops. J. Acoust. Soc. Am. 57, 456-461.

Bradley, C. B. (1909) Graphic analysis of the Siamese tones. In Transactions and Proceedings of the American Philological Association 40, xcv-xcvii.

Bradley, C. B. (1916) On plotting the inflections of the voice (University of California Publications in Anthropology, Archaeology, and Ethnology) 12, 195-218.

Brown, M. J. (1975) The great tone split: Did it work in two opposite ways? In Studies in Tai Linguistics in Honor of William J. Gedney, ed. by J. G. Harris and J. R. Chamberlain. (Bangkok: Central Institute of English), pp. 33-48.

Buchthal, F. (1957) An Introduction to Electromyography. (Copenhagen: Gyldendal).

Burnay, J. and G. Coedès. (1927) Note sur les tons et les initiales du vieux siamois à l'époque de Sukhodaya. J. Siam Soc. 21, 103-117.

Chiang, H. T. (1967) Amoy-Chinese tones. Phonetica 17, 100-115.

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

Collier, R. (1975) Physiological correlates of intonation patterns. J. Acoust. Soc. Am. 58, 249-255.

Cooper, F. S. (1955) Research techniques and instrumentation: EMG. ASHA Reports No. 1, 153-168.

Dickson, D. R. and W. M. Maue. (1970) Human Vocal Anatomy. (Springfield, Ill.: Charles C Thomas).

Dickson, D. R. and W. M. Dickson. (1972) Functional anatomy of the human larynx. Trans. Pa. Acad. Opthalmol. Otolaryngol. Spring, 29-38.

Dixit, R. P. (1975) Neuromuscular aspects of laryngeal control: With special reference to Hindi. Unpublished Ph.D. dissertation, University of Texas at Austin.

Elimelech, B. (1974) On the reality of underlying contour tones. Working Papers in Linguistics (University of California at Los Angeles) 27, 74-83.

Erickson, D. (1974) Fundamental frequency contours of the tones of Standard Thai. Pasaa, Notes and News about Language Teaching and Linguistics in Thailand 4, 1-25.

Erickson, D. (1975a) Phonetic implications for an historical account of tonogenesis in Thai. In Studies in Tai Linguistics in Honor of William J. Gedney, ed. by J. G. Harris and J. R. Chamberlain. (Bangkok: Central Institute of English), pp. 100-111.

Erickson, D. (1975b) A laryngeal description of Thai tones. Paper presented at the annual meeting of the Linguistic Society of America, San Francisco, December.

Erickson, D. and J. A. Atkinson. (1975) The function of strap muscles in speech. J. Acoust. Soc. Am., Suppl. 58, S12(A). [Also to appear in Haskins Laboratories Status Report on Speech Research SR-45/46 (1976).]

Faaborg-Andersen, K. (1957) Electromyographic investigation of intrinsic laryngeal muscles in human. Acta Physiol. Scand., Suppl. 41, 140.

Faaborg-Andersen, K. (1964) Electromyography of the laryngeal muscles in man. In Research Potentials in Voice Physiology, ed. by D. W. Brewer. (New York: State University of New York), pp. 105-129.

Faaborg-Andersen, K. (1965) Electromyography of Laryngeal Muscles in Humans: Techniques and Results. (Basel, Switzerland: S. Karger).

Faaborg-Andersen, K. and A. Sonninen. (1960) The function of the extrinsic laryngeal muscles at different pitches. Acta Otolaryngol. 51, 89.

Faaborg-Andersen, K. and W. Vennard. (1964) Electromyography of extrinsic laryngeal muscles during phonation of different vowels. Ann. Otol., Rhinol., and Laryngol. 73, 248ff.

Fant, G. (1960) Acoustic Theory of Speech Production. (The Hague: Mouton).

Ferrein, A. (1741) De la formation de la voix de l'homme. Histoire de l'Academie Royale des Sciences de Paris 51, 409-432.

Friedman, S. M. (1950) Visual Anatomy I. Head and Neck. (New York: Harper & Row).

Fromkin, V. A. (1974) On the phonological representation of tone. Working Papers in Linguistics (University of California in Los Angeles) 27, 1-17.

Fujimura, O. (1971) Remarks on stop consonants: Synthesis experiments and acoustic cues. In Form and Substance, ed. by L. L. Hammerich, R. Jakobson, and E. Zwirner. (Copenhagen: Akademisk Forlag), pp. 221-232.

Galambos, R. (1962) Nerves and Muscles. (Garden City, N. Y.: Doubleday).

Gandour, J. (1974a) On the representation of tone in Siamese. Working Papers in Phonetics (University of California in Los Angeles) 27, 118-145.

Gandour, J. (1974b) Consonant types and tone in Siamese. J. Phonetics 2, 337-350.

Gandour, J. (1975) Evidence from Lue for contour tone features. Pasaa, Notes and News about Language Teaching and Linguistics in Thailand 5, 39-52.

Gårding, E. (1970) Word tones and larynx muscles. Working Papers (Phonetics Laboratory, Lund University) 3, 20-45.

Gårding, E., O. Fujimura, and H. Hirose. (1970) Laryngeal control of Swedish word tones: A preliminary report on an EMG study (Research Institute of Logopedics and Phoniatrics, University of Tokyo Annual Bulletin) 4, 45-54.

Gay, T. G., M. Strome, H. Hirose, and M. Sawashima. (1972) Electromyography of intrinsic laryngeal muscles during phonation. Ann. Otol., Rhinol., Laryngol. 81, 401-410.

Greene, M. C. L. (1957) The Voice and Its Disorders. (New York: Macmillian).

Haas, M. R. (1956) The Thai System of Writing. (Washington, D.C.: American Council of Learned Societies).

Haggard, M., S. Ambler, and M. Callow. (1970) Pitch as a voicing cue. J. Acoust. Soc. Am. 47, 613-617.

Halle, M. and K. Stevens. (1971) A note on laryngeal features. Quarterly Progress Report (Research Laboratory of Electronics, MIT) QPR-101, 198-213.

Harris, K. S. (1974) Physiological aspects of articulatory behavior In Current Trends in Linguistics, XII, ed. by T. Sebeok et al. (The Hague: Mouton), pp. 2281-2302.

Haudricourt, A. G. (1971) L'apparition des registres des langues à tons ponctuels. In Proceedings of the Seventh International Congress of Phonetic Sciences, ed. by A. Rigault and R. Charbonneau. (The Hague: Mouton), pp. 895-896.

Haudricourt, A. G. (1972) Two-way and three-way splitting of tonal systems in some far Eastern languages. Translated by Christopher Court. In Tai Phonetics and Phonology, ed. by J. G. Harris and R. B. Noss. (Bangkok: Central Institute of English Languages), pp. 58-86. Originally, 1961, Bipartition et tripartition dans les systèmes de tons. Bull Soc. Linguistique Paris, pp. 163-180

Henderson, E. J. A. (1949) Prosodies in Siamese: A study in synthesis. Asia Major (New Series) 1, 189-215. Reprinted in Prosodic Analysis, ed. by F. R. Palmer. (London: Oxford University Press, 1970), pp. 27-53.

Hiranburana, S. (1971) The role of accent in Thai grammar. Unpublished doctoral dissertation, University of London.

Hiranburana, S. (1972) Changes in the pitch contours of unaccented syllables in spoken Thai. In Tai Phonetics and Phonology, ed. by J. G. Harris and R. B. Noss. (Bangkok: Central Institute of English), pp. 23-27.

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

Hirano, M., Y. Koike, and H. von Leden. (1967) The sternohyoid muscle during phonation: Electromyographic studies. Acta Otolaryngol. 64, 500-507.

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

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

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

Hirose, H. (1974) Functional differentiation of the intrinsic laryngeal muscles. Paper presented at Haskins Laboratories Workshop on Speech Production, 17-18 January, pp. 7-8.

Hirose, H. and T. Gay (1972) The activity of the intrinsic laryngeal muscle in voicing control. Phonetica 25, 140-164.

Hollinshead, W. H. (1954) Anatomy for Surgeons: Vol. I. The Head and Neck. (New York: Harper & Row).

Hombert, J-M. (1975) Towards a theory of tonogenesis: An empirical, physiologically and perceptually-based account of the development of tonal contrasts in languages. Unpublished Ph.D. dissertation, University of California at Berkeley.

Howie, J. M. (1976) Acoustical Studies of Mandarin Vowels and Tones. (Cambridge, England: Cambridge University Press).

Kagaya, R. and H. Hirose. (1975) Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants (Research Institute of of Logopedics and Phoniatrics, University of Tokyo Annual Bulletin) 9, 27-46.

Katusuki, Y. (1950) The function of the phonatory msucles. Japan J. Physiol. 1, 29-35.

Kewley-Port, D. (1973) Computer processing of EMG signals at Haskins Laboratories. Haskins Laboratories Status Report on Speech Research SR-33, 173-183.

Klatt, D. H. and K. N. Stevens. (1968) Studies of articulatory activity and air flow during speech. In Sound Production in Man, ed. by A. Bouhuys. (New York: Annals of the New York Academy of Sciences) 155, 42-55.

Kotby, M. and L. Haugen. (1970) The mechanics of laryngeal function. Acta-Oto-Laryngologica 70, 209-210.

Ladefoged, P. (1975) A Course in Phonetics. (New York: Harcourt Brace Jovanovich).

Leben, W. (1971) On the segmental nature of tone in Thai. Quarterly Progress Report (Research Laboratory of Electronics, MIT) 101, 221-224.

Leben, W. R. (1973) The role of tone in segmental phonology. In Consonant Types and Tone, ed. by L. N. Hyman. (Southern California Occasional Papers in Linguistics, No. 1), pp. 115-149.

Lehiste, I. and G. E. Peterson. (1961) Some basic considerations in the analysis of intonation. J. Acoust. Soc. Am. 33, 419-425.

Li, F. K. (1966) The relationship between tones and initials in Tai. In Studies in Comparative Austroasiatic Linguistics, ed. by N. H. Zide. (The Hague: Mouton), pp. 82-88.

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy. (1967) Perception of the speech code. Psychological Review 74, 431-461.

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

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

Lieberman, P. (1970) Towards a unified phonetic theory. Linguistic Inquiry 1, 307-322.

Lisker, L. and A. Abramson. (1964) A cross language study of voicing in initial stops: Acoustical measurements. Word 20, 384-422.

MacNeilage, P. (1970) Motor control of serial ordering of speech. Psychological Review 77, 183-196.

MacNeilage, P. and J. DeClerk. (1969) On the motor control of coarticulation in CVC monosyllables. J. Acoust. Soc. Am. 45, 1217-1233.

Maspéro, H. (1911) Contribution a l'étude du système phonétique des langues thai. Bull. de l'Ecole Française d'Extrême-Orient 19, 152-169.

Matisoff, J. A. (1973) Tonogenesis in Southeast Asia. In Consonant Types and Tones, ed. by L. M. Hyman. (Sourhern California Occasional Papers in Linguistics, No. 1), pp. 71-95.

Mohr, B. (1971) Intrinsic variations in speech signal. Phonetica 23, 65-93.

Noss, R. B. (1964) Thai Reference Grammar. (Washington, D.C.: U. S. Government Printing Office).

Ohala, J. J. (1970) Aspects of the control and production of speech. Working Papers in Phonetics (University of California at Los Angeles) 15.

Ohala, J. J. (1972) How is pitch lowered? J. Acoust. Soc. Am. 52, 124(A).

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

Pernkopf, E. (1963) Atlas of Topographical and Applied Human Anatomy, Vol. 1. (Philadelphia and London: W. B. Saunders Co.).

Pike, K. L. (1948) Tone Languages. (Ann Arbor: University of Michigan Press).

Port, D. K. (1971) The EMG data system. Haskins Laboratories Status Report on Speech Research SR-25/26, 67-72.

Sawashima, M. (1974) Laryngeal research in experimental phonetics. In Current Trends in Linguistics, 12, ed. by T. A. Sebeok et al. (The Hague: Mouton), pp. 2303-2348.

Sawashima, M., T. Gay, and K. Harris. (1969) Laryngeal muscle activity during vocal pitch and intensity changes. Haskins Laboratories Status Report on Speech Research SR-19/20, 211-220.

Schilling, R. (1937) Der Musculus sternothyreoidens und seine stimmphysiologische Bedeutung. Arch. f. Sprach-u. Stimmheilk 1, 65.

Shearer, W. (1963) Illustrated Speech Anatomy. (Springfield, Ill.: Charles C Thomas).

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

Simada, Z. and H. Hirose. (1971) Physiological correlates of Japanese accent patterns. (Research Institute of Logopedics and Phoniatrics, University of Tokyo Annual Bulletin) 5, 41-50.

Sonesson, B. (1970) The functional anatomy of the speech organs. In Manual of Phonetics, ed. by B. Malmberg. (Amsterdam: North Holland), pp. 45-75.

Sonninen, A. (1956) The role of the extrinsic laryngeal muscle in length-adjustment of the vocal cords in singing. Acta Otolaryngol. Suppl. 130.

Trager, G. L. (1957) Siamese phonemes: A restatement. Bull Inst. History Academia Sinica 29, 21-29

Van den Berg, Jw. (1957) Subglottic pressure and vibration of the vocal folds. Folia Phoniatrica 9, 65-71.

Van den Berg, Jw. (1958) Myoelastic-aerodynamic theory of voice production. J. Speech Hearing Res. 1, 227-244.

Van den Berg, Jw. and T. S. Tan. (1959) Results of experiments with human larynxes. Practica-Oto-Rhino-Laryngol. 21, 425-450.

Vanderslice, R. (1967) Larynx versus lungs: Cricothyrometer data refuting some recent claims concerning intonation and archetypality. Working Papers in Phonetics (University of California at Los Angeles) 1, 69-80.

Von Leden, H. and P. Moore. (1961) The mechanism of the cricoarytenoid joint. Arch. Otolaryngol. 73, 541-550.

Warotamasikkhadit, U. (1967) Some phonological rules in Thai. J. Am. Oriental Soc. 87, 541-574.

Whitaker, H. A. (1969) Stylistic tone-changing rules in Thai. Glossa 3, 190-197.

Zemlin, W. R. (1964) Speech and Hearing Science: Anatomy and Physiology. (Champaign, Ill.: Stipes).