ARIPUC VOL. 19

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1985

ts

ARIPUC VOL.

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN

This annual report is published on the basis of a grant from the Faculty of the Humanities (publications account), University of Copenhagen.

HUMANISTISK FAKULTET
REPRO AFDELINGEN

CONTENTS

ARIPUC 19. 1985. p. 1-42

SPL: A SPEECH SYNTHESIS PROGRAMMING LANGUAGE

Peter Holtse and

Anders Olsen *

This report describes the first version of a high level computer programming language for experiments with synthetic speech.

In SPL a context sensitive parser is programmed to recognize linguistic constructs in an input string. Both the structural and phonetic descriptions of the recognized structures may be modified under program control. The final output of an SPL program is a data stream capable of driving a parametric speech synthesizer.

The notation used is based on the principles known from Chomsky and Halle's "The Sound Pattern of English". This means that in principle all linguistic constructs are programmed in segmental units. However, in SPL certain macro facilities have been provided for more complicated units such as syllables or words.

1. INTRODUCTION

A special programming language for experiments with phonological rules in general and speech synthesis in particular is currently being tried out in our laboratory as part of a project to develop a Danish text-to-speech synthesis system (cf. Holtse(1982)).

---

* The Telecommunications Research Laboratory, Copenhagen

The first part of the present paper describes in broad terms the philosophy and concepts of the programming language while the last part contains the formal syntax definitions.

The idea of coding phonological rules directly into a digital computer is not new. of course. In some way or other special programming tools have been developed in most laboratories working with speech synthesis by rule. Thus. various rule testing programs have been described. e.g. Bobrow and Fraser (1968) and Basbøll and Kristensen (1974. 1975). Furthermore. several compiling and/or interpreting systems have been developed. for instance Hertz (1982). Carlson and Granström (1974). and Kerkhoff et al. (1984). These systems allow rules to be formulated in notations very like the style commonly used by linguists and phonologists.

While the early rule programs were mainly concerned with creating an environment for testing phonological rules. the later systems have tried to include facilities for direct manipulation of speech synthesizer control parameters via rule statements. Thus. the system described by Hertz (1982) is a complete interactive synthesis development system including a rule interpreter whereas Carlson and Granström describe a compiling language which can produce an entire text-to-speech conversion system.

The present paper defines a programming language (SPL for Synthesis Programming Language) very much like the one described by Carlson and Granstrom. In fact the whole project owes much to their ideas and experience. The language is. like most languages of this kind. based on the notation introduced by Chomsky and Halle (1968) in *The Sound Pattern of English.* The choice of this - maybe outdated - notational representation is not based on a firm conviction of the superiority of segmental descriptions over other more sophisticated ways of describing speech. But the segmental approach has the merit of being relatively widely accepted by linguists and phoneticians as one possible way of expressing phonological regularities. even if it is rarely the most elegant way. Also. since it is possible - with a bit of fiddling - to describe most phonologically relevant entities in terms of segments this approach is quite attractive as a basis for a general programming language so as to avoid hardwiring too many definitions into the language. And finally. the single level description is relatively easy to implement on a digital computer.

1.1. SCOPE OF THE LANGUAGE

SPL is intended as a tool for expressing regular phonological and phonetic rules. It could be used to write part of a text-to-speech system. or part of the hypothesis verification component(s) of a speech recognition system. Or it could be used directly for phonetic or phonological research.

Since an SPL program will accept normal orthography as input and is capable of directly driving a parametric speech synthesizer an entire text-to-speech system could, in principle, be coded in SPL. However, certain problems normally encountered in such systems are better dealt with in special orthographic preprocessors. Thus, since only regular expressions can be formulated in SPL single word exceptions to general rules or exceptionally weird orthographies are relatively expensive seing that they may need an entire rule to handle each word. Also, abbreviations, acronyms etc. can probably be dealt with more efficiently by general text preprocessors, although they may, of course, be handled via rules. Molbæk (1982) contains a detailed discussion of various strategies for handling these problems.

Finally, it should be noted that the SPL is not well suited for parsing larger linguistic units - primarily because it has no facilities for dictionary look up and cannot deal with even the most rudimentary kind of semantic information. This deficiency may turn out to cause difficulties with e.g. sentence stress in certain languages.

2. SURVEY OF THE LANGUAGE

This section gives a brief survey of the structure and concepts used in an SPL program. The chapter is intended as an informal description of the language - not a programmer's or user's manual. The complete formal definition of the SPL syntax is included as a separate section.

2.1. PROGRAM STRUCTURE

An SPL source program consists of two major components: A data declaration component and a rule description component. All data structures needed within an SPL program must be explicitly defined before they may be used either in rules or in other data declaration statements, i.e. forward references are not allowed. In the rule component the rules are described which transform an input text string first to a string of phonetic symbols and then to synthesis control parameters.

The actual transformation of a text string is effected in three stages:

During Stage One the input characters are translated to Phones - the internal representation used throughout the rule program. Then the Phones are loaded into a work buffer until all the characters of a complete sentence have been entered.

In Stage Two the rules of the rule component are applied in succession to the string of Phones in the work buffer. Each rule is applied to all the Phones of the buffer in a left to right fashion before the next rule comes into play. The rules may add additional Phones to the buffer or delete Phones from the buffer, or they may change the descriptions of the Phones in the buffer. In this way the contents of the buffer is gradually changed into a more and more detailed description of the utterance originally entered as text input.

When all the rules have been applied, the buffer should contain the equivalent of the acoustic segments of the utterance, and the program enters Stage Three. During this stage the segmental chunks now contained in the work buffer are interpolated and reformatted, and a file is output which contains the control code necessary to make the target speech synthesizer produce the utterance in question.

2.2. DATA TYPES

SPL attempts to impose as few restrictions as possible on the way a user can describe his linguistic theory. Therefore, the language contains no built in notions of what, for instance, a syllable or a word should look like. The only predefined phonological units within SPL are Distinctive Features and Phones, i.e. the system is basically segmental. However, quite complicated segemental sequences may be described via *structure types* (q.v.) and later referred to as syllables of various kinds etc. so that, to some extent at least, the limitations of using a segmental environment are removed. The important point is that the proper definition of such units is left entirely to the user/programmer.

2.2.1. Features, Scalars, Parameters and Phones

The basic unit within an SPL program is the *Phone* which at the input end is a segmental entity roughly corresponding to a letter or a phonetic symbol. During Stage Two, application of the rules, this description is gradually refined so that at the output stage each Phone corresponds to a separate acoustic segment. Thus, an aspirated stop will typically consist of three Phones at the output stage: closure, explosion, and aspiration.

A Phone consists of a structural part and an optional segmental part. The structural part serves to describe primarily the phonological properties of the Phone, while the segmental part contains a description of the physical properties associated with the realisation of the Phone.

The structural properties of Phones are described in terms of *Distinctive Features* (or just Features). Features are binary entities which assume values of *plus* or *minus* to indicate the presence or absence of a certain property within the Phone. Examples of Features are *consonantal, vocalic, syllabic,* or *labial.* For instance, the consonants *b, p,* or *m* might all be classified as [+labial] to indicate that they belong to the class of consonants articulated with lip closure.

Features are combined to form Matrices, e.g. [+voc, -cons, +syll]. Note, however, that while feature matrices in segmental phonology are traditionally written in columns, SPL Matrices are written in a linear fashion in reverence to the limitations imposed by most computer text editors.

Each Phone is, in principle, defined by a unique matrix of Distinctive Features. However, the values of certain Features may be irrelevant to a particular Phone. For instance, the value of the Feature "stress" could in some cases be considered irrelevant to consonants since stress may in some connections be regarded as a property associated with vowels. In cases like this the values of the irrelevant (or redundant) Features may be left undefined.

Certain properties of Phones cannot conveniently be expressed as binary values. To cope with these situations each Phone has associated with it a list of *Scalars,* which may be thought of as multivalued Features. The Scalars are, however, purely descriptive labels. They are not considered part of the definition of the Phone as such, i.e. two Phones may share the same combination of Scalar values, whereas each Phone must have a unique combination of Distinctive Features. *Duration* and *height* are examples of properties which could be expressed via Scalars. Technically, Scalars are integer variables capable of assuming the values of all integer numbers as defined by the implementation of the SPL compiler.

There are two classes of Scalars within an SPL program. The first class comprises the two predefined Scalars *DUR* and *RANK.* DUR specifies the duration of the Phone in time (expressed in milliseconds), while RANK is a control value used when the Phones are concatenated in the final output.

The second class comprises any user defined Scalars. The user defined Scalars have no direct influence on the physical characteristics of the output from the synthesis program. They may be used, as previously mentioned, to express multivalued structural conditions which can only with difficulty be expressed in binary distinctive features.

Both types of Scalars may be used in relational and logical expressions as part of phonological descriptions and conditions.

The actual acoustic phonetic realisation of the Phone is described in a table of *Parameters.*

Parameters are the physical control variables of the speech synthesizer which is eventually to use the output of the SPL program. The primary property of each Parameter is its *target.* The target value is an integer number which indicates, for instance, the frequency of a certain formant or the amplitude of a gate. Each Phone contains one target value specification for each Parameter associated with the speech synthesizer in question. Additionally, two transition times are associated with each target value: An internal and an external transition time. The transition times are dynamic properties of the Phone. They specify the speed with which the target values should be reached during execution of the synthetic utterance. (Holtse(1974) contains a more detailed description of the general strategy used during interpolation of parameters.)

Scalars or Parameters which need to refer to the same integer value in many places within a program may do so via a *Constant* reference. A Constant is an integer with a name to it.

A Phone may serve only phonological purposes and therefore have no physical realization of its own. In such Phones the Scalar and Parameter specifications need not be supplied. These Phones are known as *Pseudo Phones.*

The following is an example of the code needed to define a vowel named "alpha":

```
feature cons, voc, high, low, back, round .......
.
.
phone <alpha [Sa2033Q]>
                 [-cons, +voc, +low, -high, +back, ..... ]
                 F1  450, F2 1100, F3 2800, A0 30
```

The first name in the triangular brackets defines the name by which the phone will be known in any following rules in this source module. The character string in the first square bracket is an alternative name. This string will be printed during debugging instead of the internal name. In this way it is possible to code the source of the rule program using an ordinary text editor on any dumb terminal while the final synthesis program will be able to drive a rather more sophisticated terminal by taking advantage of, for instance, a phonetic character generator.

The matrix of binary features is the unique definition of the phone. And the last line is a description of the parameter default target values of the phone.

2.2.2. Structures

The Structure is a special descriptional aid which has been provided to circumvent the basically segmental nature of SPL. In principle it is simply a sort of short hand for a more or less complicated sequence of Phones. Thus, the segmental setup of, for instance, a syllable need only be declared once. From then on the declared name will automatically be expanded to the complete segmental syllable description every time the structure name occurs.

For example the following fragment of SPL code is one way of handling syllables:

```
feature cons, voc
.
.
phone <V [V]>  [-cons, +voc]
phone <C [C]>  [+cons, -voc]
.
.
structure  S   (C <0,3>) V (C <0,5>)
```

The first line declares the two names *cons* and *voc* to be of the type Feature. Then *V* and *C* are declared to be the names of two phones with the feature matrices contained in the square brackets. And finally *S* is declared to be a structure consisting of a vowel with from zero to three initial consonants and from zero to five final consonants. The expressions within the triangular brackets define the number of instances of the entity which are acceptable at that place.

Structure definitions may be used recursively so that the defintion above could be used in the declaration:

```
structure  W   (S<>) #
```

to declare that a word, *W*, is any number of syllables terminated by a word boundary symbol (which must, of course, also be defined as a Phone or a Structure). Please note that in the example above it is not necessary to compute the exact location of the syllable boundaries since all that is needed for the description to work is the "top" of each syllable.

2.3. DESCRIPTIONS OF TRANSFORMATIONS

Transformations are described in a context sensitive grammar and formulated in Rule Statements. Each Rule Statement contains a command word, a structural description of the string to be transformed: the Rule Kernel, a description of the context(s) in which the Kernel must occur for the Rule to apply, and a description of the changes to be made.

For example the rule:

```
change :  V / ["C", +lab] _____ # -> [+round]
```

could be one way of formulating that final vowels are rounded after labial consonants. As the example shows the Rule Kernel is separated from the Rule Context by a slash while the place of the Kernel within the Context is indicated by an underline. The right arrow points to the changes to be made.

Space, tab, newline and form feed characters may be inserted anywhere to improve readability. Thus, the whole rule may be written on one line or newlines and tabs may be used to provide special visual effects as in the example below:

```
change :      V        /       ["C", +lab] ___  #
        -> [+round]
```

The rule above will cause the synthesis program to find any occurrences of *V*, which presumably is a Pseudo Phone defined to match any vowel within the language being synthesized. Once a vowel has been located, the precontext, i.e. the context immediately preceding the Kernel, is scanned in reverse direction. In this example the precontext contains only one unit: *C*, which is probably any consonant within the language. Furthermore, the restriction is added that only *+lab* consonants are accepted.

When the precontext has been accepted the postcontext is scanned. The postcontext in this example consists of the (presumably) Pseudo Phone *#* - the usual sign for a word boundary.

If the structural descriptions of both pre- and postcontexts are matched the phone is changed as described in the last part of the statement. In this case the feature value *+round* is assigned to the vowel matching the Kernel, while all other feature values for that vowel are left unchanged. Furthermore, if the Distinctive Feature *round* is currently undefined for that vowel it will be marked as defined.

The Structures described above may be used to recognize more complicated conditions. For instance the rule:

```
changeall : S / ___  (S <2,2>)  #  ->  [+stress]
```

would, using the definitions from above, add the value *+stress* to all segments of the last syllable but three in a word of three or more syllables.

Modifications of acoustical descriptions may be programmed as in the following example:

```
change : ["V", +stress]
        /  _____          (S <num_syll= 0, 8>)    #

        ->     [(DUR = v_min +
                       ( (DUR - v_min) / (num_syll+1) )
                ))]
```

This fictitous rule will cause the number of following syllables in the current word, as previously defined, to be evaluated and placed in the variable 'num_syll'. The duration (DUR) of the stressed vowel just recognized will be set equal to the sum of the minimal vowel duration allowed (v_min as defined by the user) plus a correction component depending on the number of succeeding syllables in the word. The correction component is computed as the difference between the inherent duration of the vowel and the minimal vowel duration divided by the number of succeeding syllables as computed above. (num_syll is incremented by one before the division to avoid dividing by zero in words with stress on the last syllable.)

Consider finally the rule:

```
change :  ["V". -stress]
          /  ["V". +stress]  (C <num_c=0.7>)  _____

          ->  [(F1 += 0.1 * F1(-num_c-1))]
```

which causes a post tonic unstressed syllable to approximate the quality of the stressed syllable.

First. an unstressed vowel following a stressed vowel with from zero to seven intervening consonants is recognized. and the number of intervening consonants is placed in the variable *num_c.* Then the frequency of the first formant of the stressed vowel is obtained (F1(-num_c-1)): One segment further to the left than the number of consonants found. This frequency is multiplied by 0.1 and finally added to the frequency of the first formant of the unstressed vowel.

2.3.1. Rule Types

Various types of rules are recognized in SPL. In the examples above the difference between *change* and *changeall* rules has been shown: In an ordinary change rule each segment in the Change Field applies to a corresponding segment in the Kernel. while in a changeall rule the modifications described in the Change Field apply to all the segments of the Kernel - irrespective of the number of segments contained within the Kernel.

A third type of rule is the *replace* rule which has the form:

```
replace :  x y z  /  A  ___  B  ->  w q t
```

This type of rule will. under the conditions specified. replace the entire sequence of segments of the Kernel. irrespective of whether they are described in terms of Structures or Phones. by the sequence contained in the Change Field.

Thus. the change rule is used to modify the values of existing Phones in the buffer while the replace rule is used when entire Phones are to be substituted. Also. in a replace rule the number of Phones may differ in the Kernel and Change Fields so as to allow deletion and addition of Phones from the buffer.

Special cases of the replace rule are the *delete* and *insert* rules which are of the form:

delete : x y z / A ___ B ->

and

insert : / A ___ B -> x y z

These two types of rules require special command words as shown in the examples in order to improve error diagnosis.

2.4. INPUT CONTROL

Since the only unit recognized within an SPL program is the Phone any ordinary characters input to an SPL coded synthesis program must immediately be translated into an appropriate string of Phones. This translation is controlled via *graph statements.* The graph statement is of the form:

graph a : a1

which means that when the compiled program meets the character *a* in its input stream it must be translated to the Phone *a1* which must be a properly defined Phone or Pseudo Phone.

Alternatively, input may be complete matrix and parameter tables obtained from another SPL program. This facility allows the different phases of a complete rule system to be coded in independent programs in order to facilitate debugging.

2.5. OUTPUT CONTROL

Output from an SPL program may be provided in two ways. Either via a *print* command or via a *speak* command. The print command will cause the current contents of the buffer, including all feature matrices and parameter tables, to be output to the designated output stream. From here it may be redirected to a terminal or other printing device for inspection, or it may be used as input for another SPL program as explained above.

The speak command immediately causes the parameter tables of the work buffer to be interpolated. Interpolation is performed using a strategy very similar to the one described by Holtse (1974). This data stream is in a format acceptable for a parametric speech synthesizer.

2.5.1. Speech Synthesizers

SPL as such makes few assumptions about the type of synthesizer for which it is producing output and it may in fact be configured for a wide variaty of synthesizers - hardware or software implementations. Furthermore, SPL will recognize all the more usual parameter names, at least for synthesizers of the formant type, but it will, of course, only produce code for the synthesizer for which it is actually targeted. In the current version of SPL a separate compiler must be produced for each target synthesizer. This may, however, be changed.

2.6. PROGRAM DEBUGGING

SPL includes a *trace* facility which, when turned on, will print the output of any rule which has applied successfully. With proper use of the alternate phone representations, as mentioned in the section dealing with Phones, quite a detailed view of the actions of the synthesis program may be obtained.

3. IMPLEMENTATION

A first version of an SPL compiler has been developed as a joint effort between the Institute of Phonetics and the Telecommunications Research Laboratory, both of Copenhagen.

Currently, most of the defined facilities of the compiler and its corresponding run time system have been implemented on the VAX-11/750 computer under a VMS operating system at the Telecommunications Research Laboratory and on the PDP-11/60 computer under a UNIX operating system at the Institute of Phonetics. Our intentions are to keep the two versions as closely compatible as possible.

Also, various support facilities have been developed such as a special parameter editor which will allow very detailed interactive control of the synthesis control parameters. This is found to be a necessary tool for the development of proper Phone descriptions. Furthermore, special device drivers capable of producing phonetic script and detailed plots of control parameter traces are being developed.

Finally, driver tables for different speech synthesizers are under construction.

FORMAL DEFINITION
OF SPL

First a note of warning: Readers with a linguistic or phonetic background should observe that in this report words like *syntax* or *semantics* and their derivatives will refer to the syntax or semantics of the SPL language - *not* the syntax or semantics of any natural language.

4. NOTATION, TERMINOLOGY, AND VOCABULARY

The grammar of SPL is described in an adapted sort of Backus-Naur form. Thus, non-terminal constructs are denoted by English words surrounded by angular brackets: < and >. Terminal symbols are written in **bold** characters.

The production rule for non-terminal symbols consists of the non-terminal symbol itself followed by the symbol ::= (two colons and an equal sign). After this follows one or more terminal or non-terminal symbols.

Repetition of constructs is indicated by curly brackets: { and }. Alternative productions are separated by a vertical rule: |. Constructs surrounded by square brackets are optional.

The symbol <empty> denotes a sequence of zero symbols.

Non-terminal symbols may include an underlined part. The underlined part is an indication of a semantic subcategory - not part of the context-free syntax description. (For instance <feature name> means a <name> of the Type "Feature".

4.1. VOCABULARY

SPL programs are represented using the full set of printing ASCII characters as follows:

<letter> ::=

A | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | **J** | **K** | **L** | **M** | **N** | **O** | **P** | **Q** | **R** | **S** | **T** | **U** | **V** | **W** | **X** | **Y** | **Z** | **a** | **b** | **c** | **d** | **e** | **f** | **g** | **h** | **i** | **j** | **k** | **l** | **m** | **n** | **o** | **p** | **q** | **r** | **s** | **t** | **u** | **v** | **w** | **x** | **y** | **z**

<digit> ::=

0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9

<special character> ::=

any printing ascii character not mentioned above!

The following characters and character sequences are reserved symbols with special meaning to the compiler. They may not be redefined by the user:

<reserved symbol> ::=

feature | phone | integer | scalar | real | constant | structure | graph | change | changeall | replace | delete | insert | obl | opt | tg | ti | tx | trace | on | off | print | speak | file | include | identity | -> | /*

4.2. NAMES AND CONSTANT NUMBERS

Names denote variables, features, scalars, functions, constants, output units, identifiers, or parameters. The special class of Names used to identify Phones and Structures within rule formulations is known as Unitnames. Their syntax is identical to the syntax of ordinary Names, except that a Unitname may also contain special characters.

Each Name or Unitname must be unique and the reserved symbols previously mentioned may not be used.

<name> ::=

<letter> { <letter> | <digit> | _ }

<unitname> ::=

any sequence of printing ascii characters not containing the characters **(** *(left bracket),* **)** *(right bracket),* **[** *(square bracket begin),* **]** *(square bracket end),* **/** *(slash),* **<** *(angular bracket begin), or* **,** *(comma).*

Furthermore, a Unitname may not consist entirely of underline characters.

Examples of Names:

tot_dur
tempo
wb2

Examples of Unitnames:

p_asp
p
#
!8
t*
:

Numbers are decimal numbers. SPL supports integer and real constant numbers in the usual notation.

<integer> ::=

<digit> { <digit> }

<real> ::=

<integer>.<integer>
| .<integer>
| <integer>.

Examples of Integers:

117
2467

Examples of reals:

0.0
123.
117.99999

Lexical entries are delimited by the first character which is not a legal part of the entry.

4.3. CONSTANTS

Certain frequently used integers may be declared constant and referred to by name instead of the usual sequence of digits.

<constant> *::=*

constant <name><integer> { , <name><integer> }

A constant declaration consists of the reserved word *constant* followed by a Name and an Integer.

4.4. SPACES AND COMMENTS

Any number of white spaces, i.e. space, tab, or newline characters, may be inserted between lexical entries to improve readability and to separate entries which would otherwise flow together and cause syntactic ambiguities.

Comments are surrounded by sequences of **/*** and ***/**. Anything between (and including) these two symbols will be treated as a sequence of white spaces by the compiler.

5. PROGRAMS

A program consists of three parts: A data declaration part, an input part (the character conversion), and the program body which contains the Rule Statements and Auxiliary Commands.

<program> *::=*

<data declaration> <character conversion>
<program body>

<program body> *::=*

{ <rule statement> | <auxiliary statement> }

A Statement may occupy as many lines as desired. Blanks, tabs, or newlines may be used as previously explained to improve readability.

5.1. COMPILER DIRECTIVES

Directives look like ordinary statements. However, they are commands controlling the workings of the compiler itself whereas ordinary commands are commands to be incorporated in the program produced by the compiler.

Directives may be placed anywhere within a program.

5.1.1. Include Directive

<include directive> ::=

include " <filename> "

The Include Directive causes the compiler to include the named file of source text as if it had been part of the original input file at that point. After having read the included file, input will again be taken from the original file.

Filename may be any character string representing a legal filename within the operating system on which the compiler is implemented.

Include Directives may be nested to a reasonable depth.

5.1.2. Identity Directive

The Identity Directive is a sort of mock Data Declaration (q.v.), but it defines no new data structures.

<identity directive> ::=

identity <identity pair> { , <identity pair>}

<identity pair> ::=

(<name> , <name>) |
(<unitname> , <unitname>)

This directive causes the two names to point to the same data structure. The first name of a pair must be a previously defined Name or Unitname. The last Name of the pair will throughout the program be another way of writing the first Name.

Examples of Identity Directives:

```
identity (pluk, p)
identity (#, WB), (labial, lab)
```

6. DATA DECLARATIONS

Data types within an SPL program must, in principle, all be explicitly defined via a Data Declaration statement before they may be referenced in any other statement. However, certain types are predefined within the compiler. The predefined types are not part of the definition of the language as such, but they are implementation dependent since they are mostly concerned with the interface between the programming language and a specific type of hardware speech synthesizer.

<data declaration> ::=

{ <basic type declaration> }
{ <complex type declaration> }

6.1. BASIC TYPES

<basic type declaration> ::=

<basic type identifier> <name> { , <name>}

<basic type identifier> ::=

feature | **scalar** | **real** | **integer**

Examples of Basic Type Declarations:

```
feature front, back, high, low
scalar height
real tempo
integer contour
```

6.1.1. Features

Features are binary entities capable of assuming the values plus or minus to designate the presence or absence of a certain property in each Phone. Declaring a new Feature causes no direct storage allocation but reserves space for one binary Feature in all Phones defined later in the program.

6.1.2. Scalars

Scalars are integer variables associated with the Phones defined later in the program. Declaring a Scalar name causes no direct storage allocation, but storage will be allocated in connection with all Phone declarations.

Two Scalars, *RANK* and *DUR* are predefined within the compiler and cannot be redefined.

6.1.3. Reals

Reals are variables which may hold any real valued number within the range defined by the implementation.

6.1.4. Integers

Integers are variables which may hold any whole number within the range defined by the implementation.

Storage for Reals and Integers is allocated as they are declared.

Reals and Integers are known collectively as *variables* or - since they are accessible from all parts of a program as - *globals* or *global variables.*

6.1.5. Parameters

The fifth basic data type is the Parameter. Parameters are the physical control variables of the target synthesizer. Since, at least in the current version of SPL, a separate compiler must be generated for each target synthesizer, the Parameter Names are predefined in the compiler and cannot be changed by the user program (although synonyms may be created through Identity Directives).

However, since it is also desirable to allow the same SPL source code to be compiled for different target synthesizers, a rather generous supply of Parameter Names are known beforehand to the SPL compiler - irrespective of the actual target synthesizer. All such predefined Parameter Names will be accepted by any SPL compiler, i.e. no message of "unknown identifier" etc. will be produced. However, only the Parameters that are physically present in the target synthesizer will be reflected in the compiled object module. Non-active Parameters may be redefined so as to produce an identity between active and non-active Parameters in order to avoid having the SPL compiler ignore the statement containing the non-active Parameter.

The following Parameters are known to all SPL compilers:

F0 - Pitch or fundamental frequency of voice source.
F1 - Frequency of first formant.
L1 - Amplitude of first formant.
B1 - Bandwidth of first formant.
F2 - Frequency of second formant.
L2 - Amplitude of second formant.
B2 - Bandwidth of second formant.
F3 - Frequency of third formant.
L3 - Amplitude of third formant.
B3 - Bandwidth of third formant.
F4 - Frequency of fourth formant.

L4 - Amplitude of fourth formant.
B4 - Bandwidth of fourth formant.
C1 - Frequency of first consonant formant.
C2 - Frequency of second consonant formant.
FN - Frequency of nasal formant.
BN - Bandwidth of nasal formant.
FZ - Frequency of spectral zero.
BZ - Bandwidth of spectral zero.
A0 - Overall amplitude.
AV - Amplitude of voicing.
AS - Amplitude of sinusoidal voicing.
AH - Amplitude of hiss noise.
AF - Amplitude of fricative noise.
AN - Amplitude through nasal branch.
AB - Bypass path amplitude.
VO - Voicing switch.

This allowance of Parameter names ought to allow communication with the more usual formant synthesizers.

6.1.6. Characters

The set of all printable ASCII characters may be considered a sixth Basic Type. The character type cannot be declared, however, since it is an existing and closed corpus. Furthermore, characters may only appear in character conversion statements within an SPL program, i.e. they are removed as soon as they are brought into the program.

6.2. COMPLEX TYPES

The complex types are data types made up of combinations of other types. The two complex types are *Phones* and *Structures.*

<complex type declaration> ::=

<phone type> |
<structure type>

6.2.1. Phones

A Phone consists of two or three parts: A name, a matrix of Distinctive Features and an optional segmental part. A Phone without segmental description, i.e. without any direct acoustic manifestation, is known as a *Pseudo Phone.*

<phone type> ::=

phone <p-name><matrix>
[<segmental description>]

⟨p-name⟩ ::=

⟨ ⟨unitname⟩ [⟨script⟩] ⟩

⟨matrix⟩ ::=

[["⟨phone unitname⟩,"] ⟨feature bundle⟩]

⟨feature bundle⟩ ::=

⟨feature expression⟩ {, ⟨feature expression⟩}

⟨feature expression⟩ ::=

⟨feature value⟩ ⟨feature name⟩

⟨feature value⟩ ::=

+ | - | ?

⟨segmental description⟩ ::=

⟨segment field⟩ {, ⟨segment field⟩}

⟨segment field⟩ ::=

⟨scalar field⟩ |
⟨parameter field⟩

⟨scalar field⟩ ::=

⟨scalar name⟩ ⟨integer constant expression⟩

⟨parameter field⟩ ::=

⟨target definition⟩ [⟨transition definition⟩]

⟨target definition⟩ ::=

⟨parameter name⟩ ⟨target value⟩

⟨transition definition⟩ ::=

(⟨internal transition⟩,⟨external transition⟩)

<target value> ::=

<integer constant expression>

<internal transition> ::=

<integer constant expression>

<external transition> ::=

<integer constant expression>

<script> ::=

string of ascii characters

Examples of Phone Definitions:

```
phone <alfa[a2]> ["V". +low. -high. +back. -round]
        dur 10. rank 50. height 5.
        F1 650 (5.5).
        F2 1200. F3 2800. A0 60

phone <#[#0]> [-seg. +wb. +sylb]
```

The phone statement defines the properties to be associated with a given Phone. Each Phone has a Name (of the type Unitname) and a Script. i.e. phonetic transcription. The Name of the Phone is used in structural descriptions as an abbreviation for the complete feature matrix. For instance *p_asp* could be the Name of a special aspiration after [p].

The Script is a string of characters (including non printing characters in escape notation) to be used for debugging and other print out in symbolic form. It has no meaning to the internal workings of the SPL program but will be printed exactly as it is entered in the definition. This strategy allows the synthesis program to take advantage of any special character generators in printers or terminals while still retaining a measure of readability in the rule formulations.

The Feature Matrix is the combination of feature values which uniquely identifies that Phone. No two Phones may have the same combination of feature values.

The phone statement causes the Phone in question. together with the properties described in the statement. to be entered into the Symbol Table and the Phone Definition Table.

The data structure defined by of the matrix part of the Phone definition consists of two parts: A Definition Matrix in which the bit positions of the Features defined for that Phone are set and a Condition Matrix in which the bit positions of the Features having the value *plus* are set while Features having the value *minus* have their corresponding bits cleared.

When the matrix part of a Phone statement contains the name of a previously defined Phone the Definition and Condition Matrices are copied from that Phone and used as the basis for the new Phone. The special feature value *?* has the effect of *removing* a Feature from the definition of the Phone if it is already there. These facilities should save some typing efforts and errors in the definition part of the program.

When occurring in structural descriptions the Definition Matrix is used to mask out the undefined Feature positions so that only defined Features are matched to the input string. This strategy means that Features which are undefined for a given Phone cannot block the application of a rule.

The Segment Description defines the physical properties associated with the Phone. Thus each Parameter entry contains a target value and two transition times. The target is the frequency or amplitude to be reached during the Phone, while the transition times are the duration of the transitions external and internal to the duration of the Phone. (The duration of the Phone is contained in the Scalar *dur*.)

Transition times may be left undefined in a parameter field. In such cases they are by default set to zero.

Targets may be undefined for a given Parameter. In such cases the target value will be supplied from an internal default table depending on the target synthesizer.

While feature matrices may be "inherited" from previously defined Phones, Scalars and Parameter values must be explicitly declared for each segmental description.

6.2.2. Structures

The Structure concept is a sort of macro definition for commonly needed sequences of strings of Phones. Thus, Structure definitions may be used to simplify the formulation of complicated linguistic units such as syllables or words.

<structure type> ::=

structure <structure definition>
{ , <structure definition> }

<structure definition> ::=

<unitname> <structural sequence>

<structural sequence> ::=

{ <structural unit> |
(<structural sequence> <range expression>) }

<structural unit> ::=

<<u>phone</u> unitname> |
<<u>structure</u> unitname> |
<matrix>

Examples of Structure Definitions:

```
structure S (C<0.5>) V (C<0.5>)
structure W (S<ntsyl=1.>) #
```

The Structure definition statement consists of the reserved word *structure* followed by the description of one or more structures. Each description consists of a name and a listing of the units which make up the Structure. The elements of a Structure are Phones, feature matrices or other Structures.

The structure definition statement causes the list of elements for each Structure to be entered into the Structure Definition Table. In later rule formulations this list of structural descriptions are invoked every time the name of the Structure is used, i.e. it is a sort of macro facility for expressing commonly used complicated conditions. Technically, however, the Structure is expanded at compile time and therefore may not contain undefined or forward references.

The Range Expression is described in the chapter dealing with *Expressions*.

7. CHARACTER CONVERSION

Input to an SPL program is any string of ASCII characters. Internally in the program all operations are carried out on Phones - not on characters. The graph conversion rules define what the SPL program must do with the input ASCII characters when they are encountered in the input stream.

<character conversion> ::=

graph <character-phone map>
{ , <character-phone map> }

<character-phone map> ::=

<character> : <phone unitname>

Examples of Character Conversion Statements:

```
graph a : a1, A : a1
graph b : b_luk, B : b_luk
```

Each character-phone map defines a unique conversion from a given ASCII character to a previously defined Phone. Upper and lower case characters are different identities. Two different characters may be mapped to the same Phone. It is an error to map the same character to two different Phones.

Input characters which are not mapped to Phones are deleted from the input stream, i.e. they cannot be accessed within the program.

8. RULE STATEMENTS

Changes to the contents of the Work Buffer are made via Rule Statements. Each rule statement describes a set of conditions under which the Phones currently in the Buffer are modified. The changes may be deletions or additions of Phones or they may be modifications to the properties of the Phones already residing in the Buffer.

<rule statement> ::=

<rule command> <rule head> : <context field>
-> <change field>

A rule statement consists of a command word indicating the Rule Type followed by a Rule Head which is terminated by a colon. Then follows a description of the structural context to which the Rule applies and a right arrow pointing to the description of the changes to be made to the Buffer.

8.1. RULE TYPES

There are five types of rules in SPL.

<rule command> ::=

change | **changeall** | **replace** | **insert** | **delete**

Change Rules are used to modify the current contents of one or several Phones already residing in the work buffer of the synthesis program. This type of rule must state explicitly how many Phones are affected by the modification and how each Phone affected is to be modified.

Changeall Rules are used to apply the same modification to a whole family of consecutive Phones in the work buffer.

Replace Rules replace one or several Phones in the work buffer with a sequence of Phones obtained from the definition tables.

Insert Rules are used to enter additional Phones into the work buffer. The Phones inserted are taken from the definition tables.

Delete Rules are used to remove one or more Phones from the work buffer.

8.2. THE RULE HEAD

The Rule Head contains two fields, both of which may be empty: A Label and a Type Declaration.

<rule head> ::=

[<rule label>] [<rule type>]

<rule label> ::=

<rule class>.<rule number>

<rule class> ::=

<integer constant>

<rule number> ::=

<integer constant>

<rule type> ::=

(<type indicator>)

<type indicator> ::=

obl | opt

Examples of Rule Commands:

```
change 26.5(opt): ........
changeall 5.2: .........
insert (obl): .....
delete : ......
```

8.2.1. Rule Label

The Rule Label is used entirely for debugging purposes during program development. Thus the class and number digits are printed out every time the Rule applies successfully to a form and the Trace function is turned on. The specific numbers used have no meaning to the SPL program as such and need not be unique.

8.2.2. Optional and Obligatory Rules

SPL rules are either Optional or Obligatory. If the Type field is empty the rule is Obligatory.

Generally, the changes described in the Change Field of a Structural Rule are applied to any form which matches the structural conditions given in the Context field of the rule - This is an obligatory rule. If the Context Description of an *Optional* rule matches the input string the current state of the Work Buffer is saved in a special storage area - core or disk as the implementation prefers. The Optional Rule is then applied to the Work Buffer in the usual way, and execution continues as usual. When all the rules of the program have been applied the SPL run time system retreives the saved buffer version from its storage and applies all commands after the Optional Rule that caused the diversion. Thus two versions of the same input string are created: One with the effect of the Optional Rule included and another without.

Since every Optional Rule of a program may in principle cause a split of the Buffer this facility is not aimed at production versions of talking machines. Primarily, it is a research tool for trying out new rules. However, the ability to create several versions of the same utterance will also be needed for hypothesis building within Automatic Speech Recognition algorithms.

8.3. CONTEXT DESCRIPTIONS

The structural context to which the Rule applies consists of two fields. the Rule Kernel and the Rule Context.

<context field> ::=

<rule kernel> <rule context>

The Kernel describes the structural conditions of the string that is to be modified. while the Rule Context describes the conditions that must be met in the surroundings of the Kernel for the Rule to apply.

The unit used in the description of structural contexts is the *Context Sequence.*

<context sequence> ::=

{ <context unit> |
(<context sequence> <range expression> **)** }

<context unit> ::=

<<u>phone</u> unitname> |
<<u>structure</u> unitname> |
<context matrix>

The Context Sequence consists of a list of units which may be either Names of Phones or Structures or Context Matrices. Structures will be interpreted as a short hand form of a string of Phones with or without range indications as defined in the appropriate structure declaration statement.

<context matrix> ::=

[["<<u>phone</u> unitname>",] [<feature bundle>]
[**(**<expression>**)**] **]**

The Context Matrix is a matrix describing the conditions which must be met within a single Phone for the Rule to apply.

If the Matrix contains the name of a Phone it will be interpreted as equal to the complete feature specification of that Phone. Any specific Feature values after the Phone name will override the Feature values of the definition. Thus the matrix ["p". +voice] means: All Features defined for 'p' except that 'p' is voiced here. irrespective of its original definition.

The special Feature value *?* has the effect of removing a Feature definition temporarily from a Phone. Thus the matrix *["p", ?voice]* means: All the Features defined for *p* except that *p* is undefined for voice in this context.

The *expression* field of a context Matrix may be used to test for specific values of Scalars or Parameters within the Phone - or in the neighbouring Phones if relative addressing is employed - or it may be used to test for certain global conditions.

8.3.1. Rule Kernels

The Rule Kernel describes a sequence of Phones to which the changes described in the Change Field must be made. Basically, the Rule Kernel is just a Context Sequence:

<rule kernel> ::=

<context sequence>

This definition holds unconditionally for *changeall*, *replace*, and *delete* rules. Since the number of Phones affected by a *change* rule must be explicitly stated in the formulation of the rule, *Structures* and *range expressions* are illegal in the Kernel of a change rule, i.e. only *phone unitnames* and *context matrices* are allowed in the Kernel of a change rule.

Furthermore, the Kernel of an *insert* rule must logically be empty.

8.3.2. Rule Contexts

The Rule Context describes the conditions which must be met in the environment of the Kernel for the rule to apply.

<rule context> ::=

<empty> |
/ <context description> ____ <context description>

<context description> ::=

<empty> |
<context sequence>

The rule context is signalled by a slash. Then follows the description of the Phones that must precede a certain Kernel for the rule conditions to be met, the *Pre-Context*. The place of the Kernel itself within the context is signalled by one or more underline characters. Then follows the description of the Phones that must follow the Kernel, the *Post-Context.*

If the whole context field is empty, the rule applies unconditionally, i.e. regardless of the context in which the Kernel appears.

Context descriptions are evaluated from the Kernel and out - i.e. the Pre-context is evaluated from right to left while the Post-context is evaluated from left to right.

It should be noted that only Features, Scalars, and Parameters actually mentioned in a rule are taken into consideration when it is determined whether a given input string matches a particular rule. In any Matrix Features which are not mentioned are marked as "undefined". An already defined Feature may be declared as "undefined" through the *?* operator.

The following rules describe the conditions under which a description matches the description of a segment in the buffer:

(i) If a Feature in the description is marked as plus *(+)* or minus *(-)* the corresponding Feature in the segment under observation must also have a defined value.

(ii) If a Feature in the description is undefined, either because it has not been mentioned at all or thrugh an explicit "undefinition" *(?)* the corresponding Feature in the segment under observation may have any value.

Examples of Context Descriptions:

```
["C", -voice] ([+cons]<0,2>) V
(C) ["V", +back,+round]
(C <ntcons = 0, 4>) ["V", (dur>10)]
["V", -stress]
(["C", +dent (dur<20 !! dur>100)] <ntc=0,3>)
["C", -voice] ["C", (dur < dur(-1))]
```

8.4. RULE CHANGE DESCRIPTIONS

The Change Field is the last part of the Rule Statement. It describes the modifications to be carried out in the work buffer.

<change field> ::=

{ <absolute element> }

<absolute element> ::=

<<u>phone</u> unitname> |
<absolute matrix>

<absolute matrix> ::=

[["<<u>phone</u> unitname>" ,] [<feature bundle>]
[(<assignment field>)]]

<assignment field> ::=

<assignment statement> {, <assignment statement>}

The Change Field consists of a list of *Absolute Elements.* As were the case with the Kernel field there are certain semantic restrictions to the Change field.

Thus, in a *change* rule the number of absolute elements must agree with the number of Phones in the corresponding Kernel field so as to state explicitly how each Phone is to be modified.

In a *changeall* rule only *one* Absolute Element must appear, since the same modifications will be applied to all the Phones of the Kernel.

Finally, in a *delete* rule the whole change field must be empty.

An Absolute Element may be just the Name of a Phone, or it may be an expression involving explicit values of Distinctive Features with or without a Phone Name. If the special feature value *?* is used it will *un-define* the feature for that Phone. The *assignment field* is used when specific values are to be assigned to Scalars, Parameters, or to global variables (integers and reals). The syntax of the assignments is described in the next chapter.

If the absolute elements of a Change Description is the name of a Phone, the Feature Matrix of that Phone replaces the Matrix of the appropriate Phone in the Work Buffer. If the absolute element contains an absolute matrix the feature values of that matrix replace the corresponding feature values of the appropriate Phone in the work Buffer.

If any of the features concerned are currently undefined in the Work Phone these features become defined for that Work Phone.

Features having the value *?* in the matrix field of a Change Description should become undefined in the corresponding Work Phone if they are already defined.

If an absolute matrix contains an assignment field the receiving location(s) (or lvalue(s)) refer either to global variables or to Scalars or Parameters of the appropriate Work Phone.

Since Scalar and parameter lists are not copied into the Work Buffer until they are modified, assigning into a Scalar or Parameter which is currently not residing in the Work Buffer, will cause the appropriate Phone to be mapped back onto the definition table and the corresponding Scalar or Parameter list to be copied into the Work Buffer before the assignment actually takes place.

Relative addressing may be used to obtain values of Scalars or Parameters for comparison or copying between neighbouring Phones in the Work Buffer. The semantics for evaluating relative addresses in these (and all other) cases is governed by two general principles:

(i) Relative addresses are evaluated within the entire Context Field, i.e. Rule Kernel and Rule Context, before any modifications are applied to the Work Buffer.

(ii) Relative addresses in change Fields of Insertion Rules are evaluated when all new Phones have been entered into the Work Buffer.

Therefore in the following example of an Insertion Rule:

insert 1.1: / V ["C", (dur>dur(+1))] ___ C -> V

a vowel (V) will be inserted between two consonants (C) if the duration of the first consonant is greater than the last.

Also in the Deletion Rule:

delete 1.2: V / ["C", (dur<dur(+1))] ___ # ->

a word final vowel (V) will be deleted if its duration is greater than the duration of a preceding consonant (C), i.e. no attempt will be made to read the duration of the word boundary pseudo phone (#).

Finally consider the Insertion Rule:

insert 1.3: / V(C) ___ # -> V ["C", (dur=dur(-1))]

which will insert an extra VC-sequence word finally (before '#') after a vowel (V) with an optional consonant following. The duration of the inserted consonant will be set equal to the duration of the defined duration of 'V'.

Modifications are carried out from left to right as described in the Change Field.

Consequently the rule:

change 2.1: V C / -> [(dur += 10)] [(dur = dur(-1))]

will cause the duration of *C* to be equal to the duration of *V* - including the added 10 ms.

Whereas the rule:

change 2.2: V C / -> [(dur=dur(+1))][(dur+=10)]

will cause *C* to be 10 ms longer than *V*.

8.4.1. Assignment Statements

Assignment Statements move values to specified data locations.

<assignment statement> ::=

<lvalue> <assignment operator> <expression>

<lvalue> ::=

```
<integer name> |
<real name> |
<scalar name> |
<parameter specification>
```

<parameter specification> ::=

<parameter name> **[.<parsub field>]**

<parsub field> ::=

<empty> | **tg** | **ti** | **tx**

Lvalues are the receiving locations in assignment expressions. They may be global variables or segmental Scalars or Parameters. It should be noted that a Parameter consists of three fields. its Target and two transition times. If the parameter sub-field is left empty the expression is assumed to refer to the Target field of the Parameter. thus allowing expressions like *F1 = 250* or *F1. = 250* to mean what they appear to say.

Assignment statements containing non-active Parameters are currently ignored.

<assignment operator> ::=

= | += | -= | *= | /= | %=

The basic assignment operator is the equal sign, which simply causes the result of the expression following the operator to be left in the receiving location - the lvalue.

The other five assignment operators are *arithmetic* assignment operators. Thus, the operator *+=* causes the result of the expression following the operator to be *added* to the current contents of the lvalue, while the result of the addition is left in the same location.

The operations performed are: Addition (+=), subtraction (-=), multiplication (*=), division (/=), and the modulus operation (%=).

Real and integer values may be mixed in assignments. An assignment always converts to the type of the receiving location. Reals are converted to integer type by truncation.

Examples of Assignment Statements:

```
dur = ......
F1.tg *= ......
F2 = ......
```

9. EXPRESSIONS

In SPL there are two types of expressions: The ordinary *Logical* or *Arithmetic Expressions* and the special class of *Range Expressions.* Range Expressions are used within Context Sequences to compute the number of elements that satisfy the conditions described in the Sequence.

When expressions are evaluated over- or underflow is reported (except division by zero).

9.1. ARITHMETIC AND LOGICAL EXPRESSIONS

Constants, Scalars, Parameters, and global variables may be combined with operators in expressions to obtain new arithmetic or logical values.

Arithmetic expressions are used for ordinary computational purposes. However, the result of any arithmetic expression may be used as a logical value. Thus, any non-zero value has the logical value *true*, while a zero value is equal to the logical value *false*.

Real and integer values may be mixed in expressions. In these cases the contents of integer variables is converted to real before the result is computed. The result of a computation is always converted to the type of the receiving location. Reals are converted to integer type by truncation.

The order of precedence for the different operators is defined by the syntax:

<expression> ::=

[<logical sign>]<logical term> |
<expression> <alternative operator> <expression>

<logical term> ::=

<arithmetic expression> |
<arithmetic expression> <relational operator> <arithmetic expression>

<arithmetic expression> ::=

[<sign>] <term> |
<arithmetic expression> <additive operator> <term>

<term> ::=

<primary expression> |
<term> <multiplicative operator> <primary expression>

<primary expression> ::=

<constant name> |
<integer name> |
<real name> |
<scalar expression> |
<parameter expression> |
(<arithmetic expression> **)** |
<function call>

<function call> ::=

<function name> **(** <expression list> **)**

<expression list> ::=

<empty> |
<expression> { , <expression> }

<scalar expression> ::=

<<u>scalar</u> name> [<relative location>]

<parameter expression> ::=

<<u>parameter</u> name> [.<parsub
[<relative location>]

<relative location> ::=

(<<u>integer</u> expression>)

The Relative Location is used to specify Parameter or Scalar values from neighbouring Phones in the Work Buffer. Thus. *F1.tg(-1)* means 'the target value of the first formant in the Phone immediately preceding this one in the Buffer'.

It is an error to attempt to access Phones outside the Work Buffer.

9.2. OPERATORS

Operators are divided into four classes according to their order of precedence.

9.2.1. Multiplicative Operators

The three operators of the multiplicative class perform multiplication. division. and the modulus operation.

<multiplicative operator> ::=

*** | / | %**

9.2.2. Additive Operators

Operators of the additive class perform addition and subtraction. A term may be preceded by a plus or minus to indicate sign identity or sign inversion.

<sign> ::=

+ | -

<additive operator> ::=

+ | -

9.2.3. Relational Operators

Operators of the relational class compare two arithmetic expressions and return values of *true* or *false*.

<relational operator> ::=

== | != | < | <= | > | >=

The comparisons performed are: Equal to, Not Equal to, Less than, Less than or Equal to, Greater than, and Greater than or Equal to.

9.2.4. Alternative Operators

The two operators of this class combine the truth values of two or more (logical) expressions to produce one logical result of the operations Logical And, and Logical Or.

<alternative operator> ::=

&& | ||

A logical sign may be prepended a logical term to negate the truth value of the term:

<logical sign> ::=

!

9.3. FUNCTIONS

A limited set of the more usual mathematical functions are defined in SPL. Suggested list of basic arithmetic functions is: *log()*, *logn()*, *exp()*, *sqrt()*. This list may, however, be expanded as the need arises. Also, a special class of system functions is being considered. Thus, a function *lenght()* returning the number of Phones currently in the Work Buffer is needed.

Furthermore, a set of special input functions that will accept input from sense lines and knobs on a control panel are being considered. These could be used to make experiments with interactive modifications to parameter values.

9.4. RANGE EXPRESSIONS

Range Expressions are used to evaluate repetitions of constructs within structural descriptions.

<range expression> ::=

<empty> |
< [<lvalue> =] <minmax expression> >

<minmax expression> ::=

<empty> |
<min expression>, <max expression>

<min expression> ::=

<empty> |
<u><integer</u> expression>

<max expression> ::=

<empty> |
<u><integer</u> expression>

A Range Expression consists of two major fields: An assignment and a minimum-maximum expression field. The latter field defines the minimum and maximum number of repetitions of the structural entity in question which will satisfy the structural condition. The minimum and maximum number of repetitions allowed may not evaluate to a negative value.

The assignment field assigns the number of repetitions actually found in a successful match to the location described in the assignment field. If no match is found the contents of the location is unmodified.

If the assignment field is empty it is an indication that the user program will not need the output of the count. Consequently it will not be made available.

If the whole min-max field is empty, it means that any number of repetitions will satisfy the conditions.

If the minimum sub-field is empty it means: Zero or more.

If the maximum sub-field is empty it means: Infinitely many.

If the entire Range Expression is empty the construct is interpreted by a special convention to mean: Zero or One repetition. Thus, allowing the use of parentheses in a structural description to simply signify an optional string as in: *V(C)#.*

When a Context Sequence involving Range Expressions is evaluated the program will find the longest match which satisfies the structural conditions.

Due to the way Context Sequences involving range expressions are expanded certain structural descriptions may be overlooked. Thus the structural descriptions:

/ X (Y<0.5>) ___ or / ___ (Y<>) X

will not be correctly matched if the structural descriptions of *Y* match the descriptions of *X*: The *X-Phone* will be used as part of the range of *Y* before any matching of *X* can be done.

This situation is considered a programming error and will not be caught by the SPL compiler or run time system.

Examples of Range Expressions:

<n = 1, 2>
<0,3>
<,>
< >
<i=>

10. AUXILIARY COMMANDS

The Auxiliary Commands perform various functions extraneous to the linguistic description which is contained in the Rule and Data Statements. The functions are mainly debugging and input/output control.

10.1. TRACE FUNCTION

The Trace Command controls the trace debugging function.

<trace command> ::=

trace <trace function>

<trace function> ::=

on | **off** | <empty>

When the Trace function is *on*, the produced program will print out the contents of the Work Buffer from the start of the Buffer up to (and including) the current position of the Rule Kernel every time a Rule has modified the contents of the Buffer. The Rule Label (if it is there) will be prepended to the produced output string.

An empty function field or the reserved word *on* causes the debugging function to be turned on. It will stay on until the next off-command - or until the end of the program.

10.2. OUTPUT CONTROL

Three commands are available for controlling the output from an SPL program. They are *File*, *Print*, and *Speak*. The File statement defines the destination of the output produced, and should properly be considered part of the data dclaration. The Print command produces output in symbolic form, i.e. phonetic notation. And the Speak command produces output in a format capable of driving a parametric synthesizer.

10.2.1. File Statement

Since the File declaration interacts heavily with the operating system of the target computer of the SPL compiler the syntax described here may be considered a sort of guideline.

<file statement> ::=

file <output unit> , "<filename>"

<output unit> ::=

<name>

<filename> ::=

any legal filename

The File command causes the named output file to be associated with the output unit mentioned. If the file exists it will be opened for output. If it does not exist it will be created and prepared for output.

10.2.2. Print Command

The Print command produces output in symbolic form.

<print command> ::=

print <output unit>

The Output Unit must be a filename defined via a File statement.

The Print command causes the entire contents of the Work Buffer to be written to the appropriate output file in Symbolic Form.

10.2.3. Speak Command

The Speak command takes one optional argument.

<speak command> ::=

speak [<output unit>]

This command consists of the reserved word *speak* and an optional output destination. It causes the contents of the Work Buffer to be interpolated and the result to be output via the designated output unit. Using the Speak command without output designation causes output to be sent directly to the appropriate speech synthesizer. The format of the output is determined by the implementation.

The Speak command does not terminate the application of Rules, although the usual procedure will be to have a Speak command as the last Statement of the SPL program. If the program contains several Speak commands with one or more Rules between them, progressively more refined versions of the same input sequence will be produced.

In the UNIX implementation output from the Speak command is usually directed to the Standard Output Device from where it may easily be redirected to a file or piped to a driver program for a hardware synthesizer.

11. IMPLEMENTATION NOTES

11.1. Mapping the Contents of the Work Buffer

When a rule modifies the feature matrix of a Phone in the Work Buffer this modification will only affect the copy of the Phone in the Buffer - not the original definition of the Phone, which still remains intact in the definition table. At certain points during execution of the program the feature matrices will, however, need to be mapped back to the original definition table.

This happens first of all every time the buffer contents must be printed in symbolic form (through the trace or print command). In these cases the definition table is searched for matches with the Phones of the Work Buffer. And for every match found the corresponding Symbolname is printed. When no match can be found a special default symbol must be output (or the complete list of feature combinations?).

A more problematic mapping occurs when the Phones of the Work Buffer are expanded from consisting of only feature matrices to their full segmental descriptions, i.e. containing Scalars and Parameters. This mapping should, of course, be delayed for as long as possible so that changing, for instance, a vowel from [+back] to [-back] will cause another table of Parameter values to be used.

The first point when the mapping becomes necessary is when a Scalar or Parameter value in the Work Buffer need to be modified. Ultimately, i.e. at interpolation time, all the Phones must, of course, be mapped.

At the time of mapping the definition table is searched for matches with the appropriate matrices in the Work Buffer. When a match is found the segmental description is copied into the Work Buffer and Parameter, or whatever it was, is modified. Thus, by the end of the program the Work Buffer will hold all the modified versions of the Phones while Phones which were not modified by any rule may be taken directly from the definition table.

If no match can be found at the final mapping it is an error condition of which the user must be duly notified.

It should be noted, however, that once the segmental part of a Phone has been mapped into the Work Buffer no further changes in the feature composition of the Phone can affect the segmental properties of that Phone.

REFERENCES

Basbøll, H. and Kristensen, K. 1974: "Preliminary work on computer testing of a generative phonology of Danish". *Ann. Rep. Inst. Phon. Univ. Cph.* 8, p. 216-226

Basbøll, H. and Kristensen, K. 1975: "Further work on computer testing of a generative phonology of Danish". *Ann. Rep. Inst. Phon. Univ. Cph.* 9, p. 265-292

Bobrow, D. G. and Fraser, J. B. 1968: "A phonological rule tester". *Communications of the ACM*, 11, 11, p. 766-772

Carlson, R. and Granstrom, B. 1974: "A phonetically oriented programming language for rule description of speech". *Preprints of the SCL-1974*, 2, p. 245-253

Chomsky. N. and Halle. M. 1968: *The sound pattern of English* (Harper and Row)

Hertz. S. R. 1982: "From text to speech with SRS". *JASA* 72. 4. p. 1155-1170

Holtse. P. 1974: "Preliminary experiments with synthesis by rule of standard Danish". *Ann. Rep. Inst. Phon. Univ. Cph.* 8. p. 239-251

Holtse. P. 1982: "Speech synthesis at the Institute of Phonetics". *Ann. Rep. Inst. Phon. Univ. Cph.* 16. p. 117-126

Kerkhoff. J.. Wester. J. and Boves. L. 1984: "A compiler for implementing the linguistic phase of a text-to-speech conversion system". *Proc. Inst. Phon.. Catholic University. Nijmegen.* 8. p. 60-69

Molbæk Hansen. P. 1982: "The construction of a grapheme-to-phone algorithm for Danish". *Ann. Rep. Inst. Phon. Univ. Cph.* 16. p. 127-136

Molbæk Hansen. P. 1983: "An orthography normalizing program for Danish". *Ann. Rep. Inst. Phon. Univ. Cph.* 17. p. 87-109

ARIPUC 19, 1985, p. 43-93

AN APPRAISAL OF RESEARCH IN THE PHONETICS AND PHONOLOGY OF THAI*

JØRGEN RISCHEL

Since the Second World War there has been a considerable amount of research activity within Thai phonetics and phonology, first by foreign scholars but more recently also by Thai linguists and phoneticians. Thai being a language that plays a central role in connection with such theoretical issues as manner distinctions within stop consonants (VOT, etc.), or inherent pitch and tonogenesis, it was found expedient to take stock of the overall activity in this field. The present paper attempts to combine a survey of the field with some comments on controversial or neglected issues. The emphasis in this presentation is on descriptive and diachronic/comparative studies; work on speech disturbances, language acquisition, or language teaching is mentioned only occasionally.

I. INTRODUCTION

The topic of this paper is Thai phonetics and phonology. These terms are understood here in a broad sense, viz. as comprising not only descriptive study but also studies in diachrony (sound change) and linguistic reconstruction. One major reason for considering synchrony and diachrony together is that Thai linguistics is an outstanding example of the fruitfulness of combining these two "axes" of linguistic research. This means,

---

*) *This is a somewhat expanded version of a paper presented to the International Conference on Thai Studies in Bangkok, August 22-24, 1984, under the title "Achievements and challenges in Thai phonetics".*

on the one hand, carrying out descriptive work with a view to the "historical" implications of the results; on the other hand, it means doing comparative work and linguistic reconstruction on a firm descriptive basis and with a view to the possibility of defining interesting issues for the empirical study of extant languages and dialects.

For obvious reasons this review article must be confined to research on Thai proper, i.e. Standard Thai and Thai dialects. Thus, in principle, it disregards research on other Tai languages and dialects, even though the latter have to a considerable extent been studied with Thai as an (implicit or explicit) reference, and even though this research often provides data that are both typologically and genetically essential for Thai studies in the narrower sense. - Needless to say, evidence from other Tai languages and dialects plays a prominent role in the literature on the reconstruction of Proto-Tai; still it can hardly be questioned that Thai, particularly Central (or "Standard") Thai has been investigated in more detail than other members of the language family, so that a narrowing of the scope to Thai studies does not do injustice to the level of achievements of the field as a whole.

It must be admitted right away that the present report is hopelessly ambitious in scope, a.o. because there are reports of numerous (largely unpublished) theses and unpublished papers which have not been accessible to me. The remarks below are based on familiarity with a (somewhat randomly limited) subset of the literature; still, I have been audacious enough to give references (without comment) also to work I have not read myself, because I find the high level of activity in the field to be a highly distinctive feature in itself (which is, in a sense, as interesting as the "state of the art").

II. SEGMENTAL PHONOLOGY AND PHONETICS OF MODERN THAI

For ease of reference it may be expedient to start the section on segmental sounds by giving a brief survey of the immediately contrastive (surface-phonemic) consonant and vowel segments in Central Thai. These inventories, which can be elicited from standard textbooks, are given here in a broad transcription with a minimum of commitment to any particular phonological interpretation.

Syllable initial nonsyllabic segments:

p^h	t^h	c^h	k^h	
p	t	c	k	
b	d			
f	s			and: h ʔ
m	n		ŋ	
	l,r			
w		j		

Comments: (1) The consonants in the topmost row may be taken alternatively as aspirated stops or as clusters of stop + /h/, but the former solution is generally adopted (there being no strong arguments in favour of the cluster solution). - (2) There are a number of syllable initial combinations of an aspirated or unaspirated voiceless stop with a nonnasal sonorant (see the last two rows of the chart above), as in [kʰwaːj] *'buffalo',* [klaj] *'distant'; the inventory of such clusters will not be given here (see Noss 1964 for details). - (3) The glottal stop* [ʔ] *may be said to be predictable and hence it can be dispensed with in phonemic transcription:* /ʔɔɔk/ *or* /ɔɔk/ *'to go out'.*

<u>Syllable final nonsyllabic segments:</u>

p	t	k	
m	n	ŋ	and (at least): ʔ
w	j		

Note: the status of [ʔ] is controversial, see IIA below.

<u>Syllable medial syllabic segments:</u>

i	ɯ	u
e	ɤ	o
ɛ	a	ɔ

plus diphthongal sequences: iə ɯə uə

Comments: (1) There are various diphthongal or even triphthongal sequences ending in a labial or palatal glide; these may be taken to end in phonemic /u i/ *or in phonemic* /w j/ *(the latter solution is assumed here, for reasons mentioned in the text below). - (2) There is a clear-cut length contrast with simple vowels, the long vowels being interpretable as vowel + length or as sequences of two identical vowels:* /raːw/ *or* /raaw/ *'approximately' vs.* /raw/ *'we'. As for the complex items* /iə ɯə uə/ *these normally count as long (for occasional instances of short diphthongs see the text below).*

A. SEGMENTAL PHONOLOGY

The segmental phonology of the Thai syllable has been dealt with in numerous publications (see Bibliography) which cannot be reviewed here. The following remarks are confined to a few issues. (The overall pattern and the standard phonemicizations are treated e.g. in Henderson 1949, Haas 1964, and Noss 1964, which represent more or less different approaches. Also cf.

Vichin Chantavibulya 1959a, ch. X, for a detailed prosodic restatement of the sound pattern in the Southern Thai dialect of Songkhla.)

One major issue is the segmental or prosodic status of certain features of the FINAL PART OF THE SYLLABLE. A prosodic interpretation is proposed by various scholars, e.g. Hashimoto (1979): the final stops and nasals are variants, reflecting a "performance feature" of staccato (shorter syllable and stop ending) vs. legato (longer syllable and nasal ending).

There certainly is a fundamental difference between syllables with final stops and nasals, but this is part of an all-pervasive difference between "dead" and "live" syllables, i.e. between syllables checked by means of a final stop and all other syllables. The latter distinction is generally recognized as being useful both in descriptive and in comparative work. Marvin Brown (1965, 1976) argues that at least for Ancient Thai syllable final stops were in fact nasals plus a "dead tone". For Modern Thai he has come to a conclusion (1976, p. 33, 36) somewhat reminiscent of that of Hashimoto. He now finds that "deadness" is neither a property of tone nor of final consonant but of the syllable as such: spoonerisms and reduplication patterns suggest that it is a separate syllable component /ʔ/. Both analyses may remove a redundancy which is otherwise present for open syllables in a long vowel [V:] versus syllables in a short vowel [Vʔ]: these differ in "deadness" just as do syllables in [Vm] vs. [Vp], etc., and hence vowel length may be considered redundant in [C_0V:] and [C_0Vʔ] syllables.

It is indeed an interesting feature of Thai if there is a clear-cut dichotomy between syllables with a resonant termination (including open syllables) and syllables with a non-resonant termination. This dichotomy, then, combines with a phonotactic dichotomy between syllables with and syllables without a final consonantal segment. We may thus set up four syllable types resulting from the intersection of the two dimensions:

	resonant termination	non-resonant termination
with -C	VC^{nasal}	VC^{stop}
without -C	V: / VV	Vʔ

This scheme seemingly exhausts the general manner-of-articulation possibilities with regard to the final part of the syllable, that is, it specifies that there is (i) no possibility of syllables ending in consonantal resonants other than nasals, (ii) no possibility of a voicing or aspiration contrast of final stops, (iii) no possibility of final continuant (non-occlusive) obstruents. All of this is, incidentally, seen very clearly from the adaptation of loanwords, in which a final lateral is replaced by /n/, a sibilant by /t/ (in words such as *football*, *English*).

Phonologists working within the more phoneme-oriented tradition (like the present author) have to face the necessity of determining whether the consonant system should be regarded as defective in syllable final position, or whether one should speak of extensive neutralisation here. The former solution forces the analyst to choose between /p t k/ and /b d g/ as syllable final stop phonemes (incidentally, the "prosodic" solution outlined above does not in itself account for the lack of palatals finally). The /b d g/ solution, which has been advocated by Haas (1964, p. XI), has the obvious drawback that it introduces an otherwise unnecessary phoneme /g/. Moreover, it has been challenged by Abramson (1972), who observes that the final stops are unvoiced, so that /p t k/ rather than /b d g/ is an adequate transcription. This observation must be supplemented by information concerning the voicing conditions in case of adjacent stops in syllables such as /klàp bâan/, but it seems safe to state that the final stops are basically of /p t k/ type, and this is also the prevalent phonemicization (it is the phonemicization chosen also in Brown 1967 for didactic purposes).

The prosodic approach certainly has its merits when applied to syllable terminations in languages such as Thai. It provides for compact and structurally revealing descriptive statements which may be favourable also in a diachronic framework. Indeed, it may well be that "dead" and "live" syllables are psychologically real categories for language users. However, this does not mean that there is necessarily something special about the production of syllables in Thai (and typologically similar languages), as seems to be implied by Brown's (1965, 1976) references to what he calls "Control Phonology" and later to Action Theory.

Action theory is a very promising way of acquiring new insights into the way in which speech gestures may be planned and controlled centrally, but obviously the crucial issue here is to what extent the general format of such planning and control is structure-bound and language-specific. Is it likely, for example, that a Thai speaker produces a syllable such as [ʔim] in terms of quite different principles of syllable organization than a speaker of German does? (Thai /ʔim/ 'full' and German /im/ 'in the -', with a predictable initial glottal attack, sound sufficiently similar for this to be an interesting comparison.)

In fact, the case for Brown's and others' prosodic solution is not quite as strong as it may seem at first sight. This solution predicts that a short unchecked vowel cannot terminate a syllable, but what then about such syllables as the particle [kʰâ] without a final glottal stop? Brown himself actually gives an example of minimal contrast between final /ʔ/ and zero in his excellent AUA Thai course, viz. /háʔ/ vs. /há/ (as short forms of /kʰráp/ and /kʰá/, respectively, cf. Brown 1968, p. 139). One may say with Bee that *"final particles ... have their own 'particular' phonology"* (Bee 1975, p. 26, with explicit reference to the minimal pair /háʔ/:/há/), but why not allow for an extension of the syllable scheme to include the peripheral type /C_oV/ (or /C_oVh/??, cf. Rischel & Thavisak 1984, p. 245) with a short, unchecked vowel?

In modern Thai VOWEL LENGTH cannot be made entirely a function of syllable termination anyway, or at least it would be a rather strained solution in cases of vowel plus a final resonant, i.e. a nasal or a semivowel. Brown first seems inclined to handle such contrasts as /kan/:/kaan/ in terms of "delayed onset" under the dead tone analysis, but he ends up with what seems a straightforward length contrast for modern Thai. (For vowel length in a comparative/diachronic perspective, see also Brown 1979). As I see it, this logically entails that the analysis also accounts for the minimal contrast between, say, /kʰáa/ on the one side and /kʰá/ or /kʰáʔ/ on the other, that is, a potential distinction between long and short open syllables, for which the particles fill a gap (also cf. the remark on "linker syllables" below).

The only remaining skewness, then, is the absence of a contrast between /ʔ/ and zero finally after a long vowel, i.e. the absence of a contrast of the type /kʰaaʔ/:/kʰaa/ or /kʰáaʔ/: /kʰáa/. Although there is no such contrast open syllables may certainly have a glottalized termination associated with particular types of tone, i.e., we are in a sense back to the "prosodic" treatment of syllable final /ʔ/ (possibly as an aspect of phonation type, cf. Egerod 1971, p. 167-169).

An apparent or real difference of distribution or of distinctiveness of /ʔ/ after long versus short vowels is not very surprising. Such a situation may occur also in the analysis of other South-East-Asian languages, and it is not confined to tone languages. It may be a real crux for the analyst, and altogether it is quite appropriate that glottalization in Thai has been the object of much discussion and speculation, cf. the next section. (For the glottal stop in Thai phonology, see also Gandour 1974a.)

The PHONETIC DIPHTHONGS of Thai fall into two very different categories, viz. those ending in a palatal or velar glide ("falling" diphthongs, formerly - and still in some dialects - with a three-way distinction between [-i], [-ɯ], and [-u]) and those ending in a rather open central vowel. As for the latter there is general agreement on an analysis in terms of vowel complexes, viz. either as /ia ɯa ua/ or as /iə ɯə uə/. The former solution is probably preferable if the latter entails an identification of the final component with the single phoneme /ə/, which is otherwise a back vowel [ɤ] (both when it is long and when it is short, also see the reference to Henderson in section IIB below). On the other hand, the transcriptions /ia ɯa ua/, when used for didactic purposes (as in Brown 1967-68), may be slightly misleading if it is not explicitly stated that the termination is really of "schwa"-type, i.e. [-ə]. Phonetically, the most straightforward (though least economical) solution is to render the diphthongs as /iə ɯə uə/

with a status more or less as separate phonemic units though of a complex kind (see the remarks on "short" diphthongs below as support of this special status), and to render the back vowel [ɤ, ɤ:] as /ɤ, ɤɤ/.

The diphthongs /iə ɯə uə/ (or however phonemicized) behave in various respects like long vowels. This similarity is taken care of if both diphthongs and long vowels are represented as sequences of two vowel phonemes. However, in the overall system the analysis is complicated by the existence of marginally occurring short diphthongs, shortness here being, according to Henderson (1949), a feature proper to onomatopes and some foreign words. Haas (Spoken Thai I, p. 284) gives the example: [pʰía] (title for a person of a certain rank) vs. [pʰíaʔ] (imitative of a slapping or banging sound) in which she identifies a difference between high tone with a "strained effect" and high tone with a final glottal stop. In an alternative terminology the former example has the (normal) long diphthong, the latter the (exceptional) short diphthong. The question, then, is whether such pairs represent a difference between structures without and structures with final /ʔ/ (but possibly with some degree of glottalization in both cases as a concomitant feature of the high tone), or a difference between diphthongs counting as long vowels and diphthongs counting as short vowels. A third solution is chosen by Noss (1964, p. 15): Noss symbolizes the normal long diphthongs as /ia ɯa ua/ but the short ones as /iə uə/, which is of course technically possible.

As mentioned already, Henderson (1949) points out that the short diphthongs belong to a specific subset of syllables, which obviously have a marginal status in the overall system. This idea is elaborated for the Southern Thai dialect of Songkhla by Vichin Chantavibulya (1959). She finds that there are in that dialect pairs such as

ˊsiə ˊsiə

vs.ˊsiə ˊsiəʔ

both meaning something like 'don't do it that way or you'll soon spoil it' but having different status since the former expression may be used by a junior speaking to a senior member of the family, and the latter by a senior member of the family addressing a junior one. As to the phonetic difference between such syllables with and without the glottal stop she observes that the last element of the syllable is "markedly fronter if followed by ʔ: [ˊsiə ˊsiɛʔ]", which puts the question of a proper phonemicization of this difference into relief. - Except for such functional pairs Vichin Chantavibulya finds diphthongs with a glottal termination only in a few words (p. 100). She eventually chooses to set up a "secondary system" for deviating syllables including those with VVʔ, and finds "that the words concerned are always either sentence-final particles, phonaesthetic words, onomatopes, personal names, exclamations, or of foreign origin" (p. 119-121). By making such a separation between a primary (i.e. central) and a secondary (i.e. marginal) system she arrives at a solution which, she observes, "*avoids the necessity of setting up a quantity system for open syllables*" (p. 127).

To round off this discussion of syllable finals I shall mention also that the final components of falling diphthongs allow for alternative phonemicizations. From the point of view of general linguistics it is a commonplace that one may debate whether diphthongs with a palatal or labiovelar termination end in /i u/ or in /j w/. In the case of Thai there is in fact overwhelming evidence in favour of the usual VC interpretation of such diphthongs, since the final component sides with syllable final consonants in two important respects: (i) short and long vowels contrast before the second component (/raw/:/raaw/, etc.), (ii) diphthong plus final consonant is not a permitted structure, just as no syllable ends in a consonant cluster (hence the final consonants are deleted after diphthongs: /waj/ for wine, etc., see further Karnchana Nacaskul 1979, p. 157). It is, on the other hand, worth noting that the analysis which posits final /w j/ upsets the otherwise restrictive pattern of nonsyllabic terminations, which allows only segments specified as having oral closure ± nasality (stops and nasals). The phonemes /w j/ fall outside the general consonant pattern and must probably be granted status as a special set of semivowels occurring both syllable initially and syllable finally, as done by Haas (1964, p. XI).

The "true" diphthongs mentioned earlier combine with final /w j/ to form phonetic triphthongs: /iəw/, /ɯəj/, /uəj/ (as may be expected, there is no possibility of a length contrast here, unlike the sequences with monophthong plus /j w/: /aaj/ vs. /aj/, etc.). The existence of these triphthongs is a further argument for a differential phonemic representation of the various types of phonetic diphthongs, viz. as /VV/ in some cases and /VC/ in others, since the triphthongs are uncontroversially handled as /VVC/-structures under that analysis (which in fact is used by some but not all scholars).

Apart from the details of segmental analysis mentioned above I think the most interesting issue in segmental phonemics is the existence in polysyllabic words of syllables of REDUCED COMPLEXITY compared to "normal, full" syllables. Leaving aside the final particles there are unstressed word initial syllables as in /mapʰrâaw/ and word-internal "linker-syllables" such as the second syllable of /rátʰabaan/. Henderson (1949) deals with these in terms of "Prosodies of polysyllables"; a more recent, extensive account is given in Bee (1975). The existence of such reduced syllables as part of polysyllabic words is a topic which deserves further consideration (cf. the typological resemblance with "minor" syllables in Mon-Khmer languages).

B. SEGMENTAL PHONETICS

There are not many published studies specifically devoted to the instrumental-phonetic investigation of Thai VOWELS or

DIPHTHONGS. The authoritative treatise is the investigation of vowels and tones by Abramson (1962). The spectral characteristics of Thai are documented in Abramson's study, but there is no doubt that such data abound in several research centers, including some of the universities in Thailand (the Kay Elemetric sonagraph is available in more than one place), so that these data, if brought together, might provide even more solid evidence on the phonetics of a vowel system of the 3x3 type, for general phonetic reference.

Henderson (1975a) considers the place of articulation of the vowel series that is intermediate between the front unrounded and the back unrounded series and emphasizes that the intermediate ones (/ɯ/, etc.) are really back not central vowels, and that this should be reflected in the choice of symbols (cf. above p.48). It may be mentioned here that for the high vowel /ɯ/ both y and ʉ occur as symbolizations in language teaching materials (personally I think y is definitely the most natural choice, provided that it is made clear to the student that this symbol has nothing to do with either French or German front rounded [y] or English non-syllabic "y" [j]).

Vowel DURATION in Thai, in particular, has attracted the attention of both Thai and foreign scholars, cf. Kanda Sittachit (1972), Abramson (1974). One reason for considering durational data is that there is an interesting interplay between vowel duration and prosodic characteristics of the syllable, as pointed out in a comparative and diachronic perspective in Gandour's instrumental study (1977). Another important aspect of vowel duration is the variation accompanying rhythmical patterning (see section IV below).

The phonetic properties of Thai CONSONANTS have been studied more extensively. By far the most studied aspect is the manner of articulation of initial and final stop consonants. One reason for this interest is that the manner features involved are crucial in the context of hypotheses about tonogenesis (see later), but quite apart from this, Thai has come to be one of the languages referred to over and over again in connection with general phonetic theories about aspiration, voicing, and voice onset time (VOT). This applies specifically to the initial stops, of course.

There are, however, other reasons for taking interest in the language specific documentation of the nature of these consonants in Thai. What is the proper phonetic specification of the initial and the final stops? This question is of interest both as a prerequisite to scientifically based language teaching, and as a prerequisite to the proper placement of Thai in a language typology.

As for the INITIALS, the acoustic appearance of a three-way contrast of aspirated voiceless vs. unaspirated voiceless vs. voiced poses no inherent problems (it very nicely illustrates the descriptive expedience of the concept of VOT). However, there have been various suggestions about the laryngeal mechanism involved in the production of these stops. Various authors have suggested that there may in some instances be concomitant glottal closure involved.

Harris (1972) suggests that *"utterance initial voiced stops and approximants are usually preceded* [my underlining] *by glottal closure"*, which is interesting in a diachronic perspective, since there is very strong comparative evidence for positing Proto-Tai /*ʔb *ʔd/ as antecedents of Thai /b d/ in initial position, as argued by Li (1943 and later work). Initial /b d/ are very strongly voiced in Thai and invite a careful physiological investigation to ascertain what articulatory adjustments contribute to this strong kind of voicing as against the slighter voicing of the "b d g" series found word initially in some varieties of German and - bordering on unvoiced lenis articulation - in English. - Vichin Chantavibulya (1959, p. 67-70) working on the Songkhla dialect found that phrase internally, in examples such as /-dɔ:k, buə/ 'it's a lotus', *"voice is heard throughout but without any indication of constriction in the larynx"*. (In the same position she observes that /k/ initially in the second syllable is often voiced (sometimes fricative) if this syllable is unstressed.) - To me the initial, voiced stops in Thai sometimes sound slightly implosive, and their articulatory characteristics may be relevant to the general issue: how do we define the difference between implosive and non-implosive articulation of voiced stops?

As for the initial series /p t c k/, Brown (1965) and Harris (1972) speak of simultaneous oral and glottal closure (and release). This was not confirmed by preliminary observations by means of the fiberscope made by Rischel and Thavisak (1984): the glottis did not appear to be really firmly closed, which agrees with the assumptions of Gandour and Maddieson (1976, p. 187).

Another question is whether some of these stops are accompanied by a secondary articulation in the supralaryngeal tract. Egerod (1961, p. 65 and oral communication) has observed that /ii/ begins with what he describes as a velarized quality after /p t/ (to which he ascribes a velar pressure), and that there is also an audible modification of the beginning of /uu/ after these consonants. Harris (1972, p. 13) also speaks of velarization with /t/ before close front vowels. It seems to be the prevalent opinion among Thai scholars that the peculiar quality of these stops before high vowels is in fact a matter of velarization. Gandour and Maddieson (1976), however, have found that the larynx is sharply raised for the stop in such cases, and they assume that there is also a pharyngeal constriction which can explain *"the commonly observed 'dark' quality of vowels, especially the high front vowel, following this stop series"*. (Their argument against the assumption of closed glottis is that stops should sound ejective if the larynx raising were accompanied by glottal closure.) - The observations of Rischel and Thavisak clearly indicate that there is a narrowing in the low pharynx, appearing as a retraction of the epiglottis, i.e. a (low) pharyngealization. One would not expect such a gesture to be accompanied by velarization in a narrow sense, so the question is whether there is at all such a thing as velarization of /p t/ initially in Thai (or whether the auditory assessment of "velarized" simply is not selective enough, cf. a similar issue with regard to "emphatic" consonants in Arabic). The acoustic effects of constrictions in the back oral cavity and the pharynx are deceptive, so it takes physiological investigation to settle this issue in a definitive manner.

If this is strictly a matter of tongue-root retraction, it is interesting in a geographical perspective, cf. the extensive discussion of tongue root articulation as a feature of register in Mon-Khmer languages.

From the point of view of Thai phonology it is noteworthy that the feature tongue root retraction seems to be always most prominent with /p t/ (this is confirmed by fiberoptic observation). One might speculate whether this has something to do with the fact that /p t/, unlike /c k/, participate in a voicedness contrast (with /b d/): is it the case that "epiglottalization" serves to enhance this contrast, whereas it is less essential with the retracted points of articulation (both because of the lack of contrast here and because voicing occurs less willingly with non-anterior articulation)? As pointed out by Egerod (personal communication), the assumption of velarization would provide a straightforward answer: this feature is auditorily "effective" only with consonants having anterior articulation, and it is indeed questionable whether "velarization" is from a general phonetic point of view a possible secondary articulation with /c k/.

As regards stops in SYLLABLE FINAL position, it has been established, as mentioned above, that these are basically unvoiced. It is also assumed that these stops are glottalized, cf. Harris (1972, p. 11ff.). The question of glottalization (or possibly laryngealization?) in final stops is crucial in a diachronic perspective (see below). - It would be useful to have access to published data on the behaviour of the final consonants (with regard to voicing assimilation and presence or absence of glottalization) in a variety of environments ranging from the position before pause to the position immediately before a voiced stop in a following stressed syllable (of the same stress-group).

As for the oral articulation of Thai consonants there is an abundance of valuable impressionistic and, in part, instrumentally based information in the literature, e.g. in Harris' paper (1972). Some of this information refers to dialects other than Central Thai (Standard Thai) but is often very suggestive also for the articulatory description of consonants in Central Thai, cf. the numerous palatograms and the detailed descriptions in Panupong (1972). Consonant articulation has been described also as a sociolinguistic variable (cf. Beebe 1976, Tanwattananun 1982).

III. PROSODIC FEATURES OF THE SYLLABLE IN MODERN THAI

Before going into the intricacies of Thai prosody it seems expedient (as with the segmental items above) to present first the immediately contrastive items on the syllable level. The tabulation will be limited to the "tones" (tonal contours) occurring on isolated syllables containing a long syllabic or a voiced final (or both); the more restricted sets of tones occurring in other types of syllables or in connected speech require special comment (see text).

Central Thai has five tones, which are generally symbolized by diacritics place over the (first) vowel symbol in the transcription of the syllable:

(not marked)	:	"mid" (slightly falling at end)
`	:	"low" (slightly falling)
´	:	"high" (rising, ± laryngealization at the end)
ˇ	:	"low rising"
ˆ	:	"high falling"

Note that the tone marks occurring in the <u>Thai script</u> have no straightforward relation to these five tones of modern Thai, though the tone category of a monosyllable is generally uniquely predictable from the choice of consonant symbols plus the use of tone marks (this will be touched upon later in connection with diachronic studies).

Tone is the phonological characteristic of Thai <u>par excellence</u>. The five tones of Central Thai have been the object of study above all by Abramson (1962 and later, see Bibliography), who has given detailed acoustic descriptions and studied the tones also from the point of view of perception (also cf. Gandour 1978). Basic phonetic research has also been done by others; it should be mentioned in particular that Gandour and Erickson both deal with the production of Thai tones in a general theoretical framework (theses and various papers, see Bibliography).

It is well established that the tonal system of Thai is a contour tone system though involving not only rising and falling but also more or less level tones. The latter are found to be the ones that are most easily confused (in the case of mid and low tone) since the most important perceptual cue may be relative pitch level in this case (cf. Abramson 1975b, 1976). The "high" tone is not just high but high rising or high rising-falling, often with audible laryngealization at the end; Henderson (1982) observes that the manifestation of this tone has been changing during this century, tending now towards a more purely rising contour (there is, nevertheless, no major risk of confusion with the "rising" tone, which in fact is slightly falling at the beginning and rises only comparatively late in the course of the syllable).

From the phonological view one of the much debated issues is the possibility of a componential or feature analysis of Thai tones. Leben (1971 and elsewhere) advocates the possibility of a reductive analysis of tones, likewise Gandour (1975). Abramson (1978) challenges the idea of splitting tone contours into consecutive levels, one argument against such an analysis having to do with the behaviour of the tone shapes when they are reduced in connected speech.

As I see it, it is essential to distinguish between at least three categories of arguments if one wishes to advocate a componential analysis of tone: (1) In some languages (not Thai) there is a strong case for such a solution because of morphophonemic processes, composite tones arising from the combination of morphemes with simple tones: low + high → (low) rising, etc. (2) Phonetic and phonotactic evidence may support the analysis of some tones as composite, others as simple. (3) A componential analysis may give an expedient taxonomy, e.g. for dialect geographical purposes (3 levels: high, mid, low giving theoretically 3x3 = 9 possible contours with two components). Such a taxonomy is of course fruitful only if the analysis is reasonably adequate from a purely phonetic point of view; to take an example: is it satisfactory to label the tone of Central Thai lέεw "high" or [hi]+[hi] without any further qualification?

This leads over to another issue: how can instrumentally recorded tone curves be specified in terms of a finite number of numerical values? Is it best to state the time and frequency coordinates of the start, the end, and whatever major tonal break (a maximum or a minimum) there may be in between? To what extent is a specification of start, middle and end sufficient? Should the time coordinate be given in centiseconds, or in percentage of total duration? Should the frequency coordinate be given in absolute values (Hz) or in terms of tonal intervals (semitones)? What parameters are useful when dealing with tone in context (possibly involving either truncation or shrinkage of the total, unperturbated contour)? The literature is rich in solutions to such problems, but I do not think a simple answer can be given that covers all types of data and all uses of the tone descriptions.

There are by now several descriptive studies which deal with the tones of Standard Thai as well as those of other Thai dialects. Some of these are listed in the bibliography. In his pioneering work on comparative Tai (see later) Fang Kuei-Li has devoted several papers to the diachronic relationship between tones and initial consonants (cf. Li 1962, 1966, 1970); his descriptive work on tonal systems (and other aspects of phonologies) is concerned with Tai languages and dialects outside Thailand in the first place. Within Thai dialectology in a narrower sense there are quite a few studies of the tonal systems of individual dialects and also some comparative or contrastive studies, an early paper in this latter category being Haas (1958). Egerod (1961) and Brown (1965) gave comprehensive surveys of the tonal systems of a variety of dialects (including samples of Southern Thai, Central Thai, Northeastern Thai and Lao, Northern Thai, and Shan); Brown's monumental work remains the principal source of reference for tonal systems in Thai dialects. For references to other work on the dialects and their tones (predominantly by Thai scholars) see the bibliographical survey paper by Kalaya Thingsabadh (1984, p. 7-9).

The Thai dialects are found to differ significantly both in the realization of the individual tones and in the overall number of contrastive tones. Comparative work also discloses differences in the distribution of the tones on individual lexemes, which has become a major cue in genetic classification of Thai dialects (see later). The "tone chart", which is designed to bring out the systematic features of the distribution of tones on syllable structures, therefore plays a prominent role in these papers and monographs. This tends to make much of the literature on tones in Thai (dialects) less accessible to general linguists or phoneticians without some knowledge of the framework developed for comparative Thai studies, so much more since the "tone chart" refers not to phonological structures of modern Thai but to <u>reconstructed</u> structure types (fortunately for scholars who, like the present author, have a bad memory, these structures are largely retrievable from Thai orthography).

Also Tai languages and dialects not belonging to Thai proper have been studied extensively from the tonal point of view. Although these studies are, on the whole, kept outside the scope of the present paper, I wish to mention that Kanchana Ngourungsi (Patamadilok) in her work on the Tai Yai dialect (Lic.Phil. thesis, University of Copenhagen) observed what seems to be a coexistence of different tonal systems, possibly correlated with sex (the Tai Yai dialect is found in a small pocket in Northern Thailand). It is known from tonal studies elsewhere that interference of dialects with other dialects or regional norms may - at least in a transitional stage - tend to produce slightly different effects for (the majority of) men and (the majority of) women, probably because of differences in their pattern of social interaction with speakers of other language norms.

Another interesting issue in connection with tone is the impact of the intrinsic pitch of vowels, and the effect of consonant type on vowel pitch and hence on tone contour. There has been some study of this, also for Thai, particularly as regards the effect of different types of initial stop consonants (such as /ph p b/) on the pitch contour of the syllable (cf. e.g. Gandour 1974b, Erickson 1975). Amon Thavisak has also made some acoustic measurements of these aspects of tone in Thai at the University of Copenhagen. All evidence suggests that the pitch starts lower after voiced stops than after voiceless stops (which is universally true), whereas the picture is anything but clear as regards aspirated versus unaspirated voiceless stops. Pitch perturbation caused by segmental syllable composition seems to tend to be less in tone languages than in non-tone languages, but it certainly plays a role in Thai, both with regard to consonantal influence and with regard to the intrinsic pitch of vowels (the general rule of thumb being that high vowels are accompanied by slightly higher pitch than low vowels, everything else being equal).

These features of tone perturbation must be taken into consideration in all tone study involving acoustic measurements. Strictly speaking, this means that the contours of different tones are not comparable unless the syllables are segmentally identical, and that one may have to truncate the tone curve - or make some numerical compensation in order to arrive at the canonical tone shape if the initial consonant is of a type expected to have a significantly perturbating influence on pitch.

The interrelations of vowel length and glottalization with tone have been mentioned earlier and will not be taken up here. Dynamic and rhythmical aspects of syllable prosody will be dealt with in the next section.

IV. FEATURES AND MODIFICATIONS ASSOCIATED WITH CONNECTED SPEECH

In recent years there have been quite a few studies dealing with prosodic aspects of Thai phrases and utterances, although the bulk of empirical data is unpublished.

From the general linguistic/phonetic point of view one of the most interesting issues is: to what extent do tone languages exhibit a SENTENCE INTONATION superimposed, as it were, on the individual tonemes? This aspect of Thai grammar and phonology is covered by the recent study of Sudaphorn Luksaneeyanawin (1983) (the contents of which are only known to the present author through a two-page abstract and a short paper (1984)).

The study of SENTENCE INTONATION involves a number of complex issues, both phonetically and phonemically, but generally speaking the primary task is to come to grips with the relationship between intonation and syntax/semantics, a topic which transcends the boundaries of the present report. Rhythm and intonation are interwoven as signals of the division of utterances into smaller units (possibly a hierarchy of units of different size); this aspect will be taken up below. It may be expedient, however, to refer here to the analysis of Noss (1964, p. 21). As part of his extensive analysis he sets up two "intonation phonemes" having to do with the way intonation contours are linked together: /./ = Pause, and /↑/ = an element meaning that a new intonation contour begins on a high pitch line (examples of how these function in clause constructions are found on p. 22 and 38-40 in Noss 1964).

Across languages intonation, and particularly the final part of the intonation contour, serves to express modalities (such as statement versus interrogation), and intonation is one of the major factors in signalling attitudes of the speaker. It is an interesting issue how intonation works in tone languages, of course. Abramson (1979b) recognizes three terminal pitch contours for "non-emotive" sentence prosody, partly on the basis of the work of Panninee Rudavanija (1965). Henderson (1949) focuses on the information carried by final particles, and describes a variety of types of "sentence tone" associated with these to express command, interrogation, etc.

A related issue is the existence of EMPHASIS as a prosodic category. Emphasis in the most general sense, as something to do with "underlining" (putting into relief) for insistence, for contrast, or just for focus, is found in the most diverse languages, with manifestations involving extra high (or extra low) pitch and possibly dynamic and durational features as well. Thai certainly has "intensification" manifested tonally, as in the first syllable of /dí́i dii/ 'very good' (see Haas 1946) but also other differences of enhancement of syllables. Like intonation proper, this complex of types and functions of syllable enhancement poses a descriptive problem in general, but it may be particularly interesting to study these matters in a language in which tone already has a considerable lexical load. Several authors recognize at least a categorial, binary difference between stressed (or: accented) and unstressed (or: unaccented) syllables, and the function of this dichotomy in relation to grammar has been investigated by Samang Hiranburana (1971).

A subject which has enjoyed considerable attention on the part of both Thai and foreign scholars, is the greater or lesser stability of lexical tones in positions of TONE COARTICULATION, i.e. before a closely succeeding prominent syllable. The literature comprises both impressionistic and instrumental studies, a major issue (perhaps first pointed out by Henderson 1949) being to what extent tonal neutralizations occur. Among contributions relevant to this issue are those by Noss (1964), Whitaker (1969), Samang Hiranburana (1972), and Abramson (1979a,c). S. Hiranburana sets up a taxonomy defining the set of "unaccented" syllables in Thai (loc. cit., p. 25-26) and finds that the tone changes occurring in these syllables cause a collapsing of the five distinctive contours of lexical tones into three level pitch contours: "mid", "modified low", and "high". Abramson (1979c) challenges the view that all syllables should be considered to bear a phonemic tone; he finds that the pitch imposed on particles *"seems to be determined by the intonation of the whole sentence"*, and that although the results of this can sometimes be aligned with the lexical tones of Thai phonology, they are more often deviant. As for the preservation of tones in running speech, the general picture is that the shapes of tones in isolation undergo severe modifications in running speech, but, says Abramson, *"as I look at the contours and listen to the speech, I find preservation of the full system of five tones in running speech"*, although particles must be excluded from this statement, and other *"frequently used function words, such as modals and pronouns, often undergo tonal replacement"* (p. 386). Also see Abramson and Katyanee Svastikula (1983).

One further study must be explicitly mentioned here, a.o. for its extensive discussion of the approach to instrumental analysis, viz. Gsell (1979). As for neutralizations in colloquial speech, Gsell only recognizes two "Architonèmes" in unaccented position (p. 69), as against Abramson's inventory of three. As for language typology, Gsell notes that Thai cannot be said to have tonal sandhi, but only coarticulation.

Several studies deal specifically with RHYTHM, i.e., the clustering of syllables into larger units, and the phenomena of enhancement and timing serving to cue this clustering. References are legion, but it may be appropriate to single out for specific reference the studies of Noss (1972, with a useful survey of earlier literature) and Theraphan L. Thongkum (1976a, b, 1977, 1984).

In the grammar of 1964 Noss specifies rhythmic patterns as having six relative syllable durations. These are analysed in terms of the intonation phoneme /./ (pause) vs. /,/ = phrase boundary, the stress phoneme /:/ (sustained stress), plus two extra phonemes /-/ and "space". All these prosodic phonemes when occurring alone or in mutual combinations specify the relative duration of the syllable preceding the symbol(s), the longest duration occurring before /: ./, successively shorter durations before other symbols down to /–/, and syllables not followed by any such symbol (including space) being very short. - Noss sets up three stress phonemes, the Sustained Contour /:/ mentioned already, plus Loud Onset /!/ and Normal Onset /'/ (both written before the syllable in question), as in /!paj/ 'Let's go!' vs. /!paj:/ 'Sure (he) went!' vs. /'paj/ 'Yes (he) went' vs. /'paj: kan/ 'They went' vs. /'ɔ̀ɔg: paj/ '(He) went out' vs. /'ɔ̀ɔg paj: khráb/ '(He) went out, sir' (p. 21). These stress phonemes or phoneme combinations are also employed in a careful specification of tone allophones, Noss recognizing well-defined differences in contours under varying stress (p. 18-20).

This whole descriptive system is posited with a wealth of illustrations and interesting applications to grammar (ch. II), but with no theoretical discussion of the analysis. In his paper of 1972, however, Noss has a principled discussion of rhythm and stress. He states (p. 37) that syllables do have discernibly different relative lengths, i.e., as he puts it, that *"rhythm is a phonetic feature of Thai"*. He also notes that there is general agreement on some kind of phonetically marked unit which is larger than a syllable and smaller than an utterance, though there is disagreement on the status of this unit (rhythm-group, stress-group, pause-group?).

His paper is a comment on the standing issue whether rhythm and stress are independently phonemic properties of Thai, or whether one depends on the other (either so that rhythm is determined by stress, or so that stress is determined by rhythm). While suggesting that instrumental research is desirable, Noss himself has used a slow-speed playback technique to assess relative differences of syllable length by ear. He thus arrives at contrasts like the following (numbers indicating relative duration, 1 being longest and 5 shortest):

	tham	maj	maa	aw	sìi	mooŋ
(a)	3	3	2	4	3	1
(b)	4	4	4	2	3	1

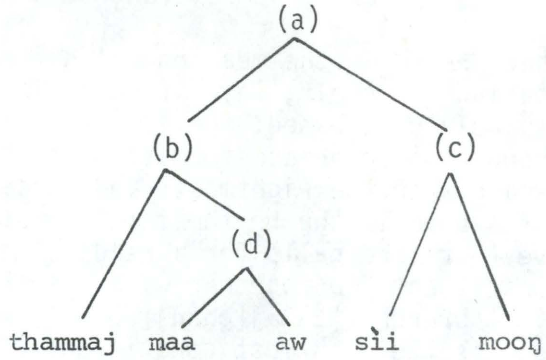
which, with durational pattern (a), has the reading: 'why do you choose to come at 4 o'clock?', but with pattern (b): 'why come to get it at 4 o'clock?'.

Noss' point is that such contrast give evidence that *"rhythm in Thai, if not phonemic, is at least interesting"* (p. 41).

It is not difficult to see that there is something interesting going on here, but it remains an open question how to handle such contrasts. This largely depends on how one defines STRESS, and on whether it is considered desirable to account for rhythmical clusters with reference to a category of stress, like this is done for the so-called "stress-timed" languages (like English). On reading (b) of the utterance above, it is obvious that the surface rhythm reflects a semantico-syntactic clustering of /maa/+/aw/ into one complex unit; on a more "surfacy" level, however, there is a further clustering with the lexical item /thammaj/. The items /sìi/+/mooŋ/, in turn, form a semantico-syntactic unit reflected phonetically. We thus get the following hierarchical structure (disregarding for simplicity the internal structure of /thammaj/):

$$[_a[_b[\text{thammaj}][_d[\text{maa}][\text{aw}]]_d]_b[_c[\text{sìi}][\text{mooŋ}]]_c]_a$$

or, in the visually more expedient tree structure notation:



Now, provided that every branching is assigned a rhythmical feature of final weight, and provided that "final weight" is interpreted (in Thai phonetics) primarily in terms of duration, we can in fact generate the relative durations of Noss' example by a simple algorithm (which shall not be dealt with here). - Similarly with example (a), provided that the hierarchical structure is now supposed to have its major (highest) branching after /maa/:

$$[_a[_b[\text{thammaj}][\text{maa}]]_b[_c[\text{aw}][_d[\text{sìi}][\text{mooŋ}]]_d]_c]_a$$

This kind of analysis raises the immediate question: to what extent are such analyses semantico-syntactically supported? To the extent that there is agreement between rhythm and syntactic structure posited on independent grounds, there is a pay-off both ways: syntax helps to "explain" (in the sense of providing a basis for generating) the rhythmical aspect of sentence prosody, and the latter may be adduced as support of a certain phrase structure analysis. To the extent that there is disagreement, there may be residues in syntactic structure which have not been taken properly care of hitherto, and there may also be syntax-independent rhythmical principles at play. Or the whole analysis may be inadequate.

As seen e.g. from the examples analysed by Noss, there are cases which do not yield to an exhaustive analysis of the type outlined above, e.g.

	thâa	chán	sǔuŋ	ìik	nít	nʉŋ
(a)	4	4	2	3	1	3
(b)	4	2	3	4	1	3

meaning in both cases: 'if X (i.e. /chán/) were just a little taller/higher', the difference being that X is understood to mean 'I' under reading (a) but to mean 'the shelf' under reading (b). Such contrasts may involve both differences of stress and of phrase assignment, as noted by Noss, but it is a question on what level of description it is relevant to speak of stress.

One may claim that /chán/ in the meaning 'I' differs from /chán/ in the meaning of 'shelf, layer (etc.)' in that the former is underlyingly unstressed, the latter stressed. There must then be a condition on phrase formation saying that an unstressed item cannot be the rightmost (and hence heaviest) constituent under a node in the rhythmical tree-structure. Reading (a) above therefore calls for a readjustment so that the first two constituents go together with the third one to form one rhythmical branch [[thâa][chán][sǔuŋ]], whereas on reading (b) /chán/ is the rightmost constituent of [[thâa][chán]], and /sǔuŋ/ is free to go together with [[ìik][nít]] to form one other branch (under a higher node than that separating [ìik] and [nít]). The next question, then, is whether the alleged "stress" difference between the two words /chán/ (a) and /chán/ (b) is lexical (inherent) as a phonological feature, or whether it reflects a difference between a "major" lexical category (including such nouns as /chán/) and a "minor" lexical category (including such pronouns as /chán/).

Incidentally, the pair of clauses above also illustrates another important kind of "residue" found with the iambic (rightmost-constituent-heaviest) conception of rhythmical trees. This residue has to do with phrase final "particles" in the widest possible sense of this word. Obviously, /nɯŋ/ in the examples above upsets the possibility of setting up a requirement to the effect that <u>every</u> branch in a rhythmical tree must be terminated by at least one "heavy" constituent placed as much to the right (under the node in question) as possible. If, however, such a principle cannot be upheld, the whole principle of analysis collapses. This means that the only way to save the analysis is to introduce a special rule for certain particles, stating that they are, or may be, "extrametrical" (to use a term coined within recent metrical theory in phonology), i.e. that they may not count in the building up of the tree-structure. Sentence final particles such as /khâ/, /khráp/ (and variants such as /hâ/, /háʔ/) obviously belong here, together with /ná/ and some others, and this helps to put their deviation from other lexical items (in terms of segmental complexity) in its proper perspective. Particles are not the only short syllables; other syllables as well may occur in reduced versions (of the type /C_0V/ with a short vowel not followed by glottal stop) but only as <u>non-final constituents</u> under a node in the rhythmical tree. Particles, however, may be extra-metrical, and those that always are may have a structure which would not permit them to ever occur as the heavy (i.e. rightmost) constituent under a metrically counting node: this is true of both syllables of /ná há/, for example.

The approach outlined above is based on an "iambic" rhythm principle, which seems to be operative in a process-like fashion on constituents such as /tɔ̀ɔ 'paj/ 'next', which before a stressed syllable in the same phrase exhibits either "iambic reversal" or loss of stress on the whole sequence /tɔ̀ɔ paj/, as in /'tɔ̀ɔ paj 'níi/ or /tɔ̀ɔ paj 'níi/ 'from now on; following; (as) follows' (from Haas 1964, p. 188 - Haas indicates stress by an acute accent after the syllable in question).

There is, in the present author's opinion, a strong case for this principle at least on an abstract level of Thai phonology. The question is whether it holds in surface phonology/phonetics.

Theraphan L. Thongkum (cf. above) works within the opposite framework, as it were. She posits a *foot* with the "beat" on the *first* not the last syllable. This means that light syllables in final position fall in place and do not have to be regarded as extra-metrical. On the other hand, there will be a residue of light syllables occurring *before* the first full beat, and obviously the occurrence of such initial syllables will be an entirely normal situation even for structures not containing words of "minor" lexical categories, cf. the non-final syllables of such sequences as /phɔ̂ɔ mɛ̂ɛ/ 'father and mother', /klàp bâan/ 'return home', adverbial /thâa jàŋŋán/ 'if so', etc. etc., which certainly need not be preceded by any lexical material. The solution to this in her framework of description is to use the Abercrombian idea of a silent beat preceding the seemingly pretonic syllables so that these are in fact *posttonic*, viz. belonging to a foot without a manifested head syllable:

$$\underbrace{\overset{\text{Beat}}{(0)}\ \text{klàp}}_{\text{Foot}}\quad \underbrace{\overset{\text{Beat}}{\text{bâan}}}_{\text{Foot}}\quad \text{etc.}$$

If this is the appropriate "surface" solution, and it may well be, then there is a discrepancy between underlying and surface organization of the prosodic structure. There is, however, nothing particularly controversial in that (a similar discrepancy has been noted for Danish in recent work).

Th. L. Thongkum has done acoustic measurements of duration, which seem to support the validity of the Abercrombian parsing of syllable sequences into feet, but probably more research is needed before it can be decided with certainty whether light syllables go exclusively with either the preceding (silent or segmentally supported) beat or the following beat, i.e., whether they are to be regarded as exclusively posttonic or pretonic or both ("tonic" being understood here to refer to the placement of the beat). This is indeed an empirical issue, which can be approached at different levels of analysis, e.g. by acoustical measurement (as done by Th. L. Thongkum), but also by perceptual studies.

It is certainly of interest also to find evidence for more abstract psychological patterns having to do with rhythmical parsing. One of the very interesting fields of study in this context is poetic METRE and the accentuation of syllables in renderings of verse. I have entirely refrained from including considerations of the Thai literature on this subject here because of personal ignorance about the performance of Thai poetry (quite generally I find this a difficult field of study because composition and performance of poetry often reflect traditions associated with a specific style of speech).

A central issue for Th. L. Thongkum is to what extent durational relationships support the notion that Thai is a SYLLABLE-TIMED language, or to what extent they point toward STRESS-TIMING. There is much debate in the international phonetic literature on the role of either the single syllable or the foot (or whatever term may be appropriate for a cluster of syllables) as the basic unit of measure: are syllables spaced relatively evenly within a sentence, or is this rather true of feet? If languages differ significantly on this point, then tone languages such as Thai are a priori expected to be candidates for the former type of behaviour. However, both Th. L. Thongkum's measurements, and general observation of speech performance, indicate that Thai cannot be called truly syllable-timed (like Lisu), nor truly stress-timed (like English) but represents a mixed type: syllable-stress-timed rhythm.

V. DIACHRONY AND RECONSTRUCTION

Tai (understood as a wider term including Thai as well as some other languages) is one of the language families to which the comparative method has been applied with much success in this century. The genetic relatedness among the languages dubbed "Tai languages" is well established, whereas the relationship of Tai to other Southeast or East Asian languages is a controversial issue.

It would go beyond the scope of the present paper to review the literature on comparative Tai phonology. The bulk of scholarly work in this field will be taken into consideration here only to the extent that it is indispensable for a proper understanding of issues in the comparative and diachronic study of Thai proper, "Thai" being understood in a narrow sense, i.e. to a first approximation, as a common denominator for dialects of the language of Thailand (though Lao and Shan etc. are often included).

Within Thai proper, the pioneering comparative and diachronic work was done by Egerod (1961) and Brown (1965), the approach of the latter being to trace the sound shifts leading from an assumed common ancestor: Ancient Thai to the modern dialects (also cf. Jones 1965b). In recent years several descriptions and comparative studies of dialects in Thailand have been performed (see the survey by Kalaya Thingsabadh 1984), for the most part by Thai scholars and students. At the same time the general comparative and diachronic research has been continued by scholars such as Brown (1975), Gedney (1972, 1973), Chamberlain (1972b), and Strecker (1979a, 1979b), just to name some of the work done.

One of the things that make it complicated to view the individual Thai dialects in a comparative and historical perspective is that it is not self-evident how to define what dialects belong to one and the same language, as it were. One would hardly a priori expect the linguistic boundary between Thai and non-Thai dialects to coincide with the present national boundary, but even if the concept of "Thai dialects" is defined on a purely descriptive-linguistic basis, one cannot be certain that the resulting cluster of dialects derive directly from a common ancestral language, to the exclusion of other dialects. The classification of Tai languages in general is quite controversial (cf. Chamberlain 1975), and this is also true of the "Southwestern branch" of Tai to which the Thai dialects belong, according to the classification of Li (and Chamberlain). From a synchronic point of view the most obvious conflict between different classificatory principles is found with the so-called Northeastern Thai dialect in Thailand, which is in essential respects just a variety of Lao.

These complications, however, do not make it irrelevant to perform studies for which the scope is (in the first analysis) strictly limited to the appropriate dialects within Thailand, viz. "Northern Thai", "Northeastern Thai", "Central Thai" (to which Standard Thai belongs as a normative sociolect), and "Southern Thai". There may well be uniting features that make it interesting to speak of a Thai dialect geography in this narrow sense, since in relatively recent time there has been an influence exerted by Central Thai on certain other dialect areas within Thailand, as well as mutual contacts among these other areas. In fact, much scholarly activity is being devoted these years to such dialect geographical work within Thailand proper, and not surprisingly, transitional dialect areas are coming into focus. These transitional areas are a challenge to dialect classification, but they are highly interesting also from a historical point of view (being evidence of earlier cultural contact or migration routes, etc.). This is true, e.g., of the Thai Isan - Thai Khorat area investigated by Vichin Panupong (1983), cf. Brown's characterization of Thai Khorat as "central Thai with a Lao accent" (Brown 1965, p. 23).

When considering dialects (or languages) in a historical perspective it is always the linguist's delight if it is possible to set up a "Stammbaum" with an ancestral language from which all modern dialects spring as separate branches (the greater or lesser mutual relatedness among dialects being reflected in the hierarchy of branchings). Strictly speaking, however, this is only likely to work with dialects that have been geographically separated from each other ever since (maybe before) the dialect split. Obviously, Thailand is a place where migrations and cultural and political dominance have to a large extent had the opposite effect, i.e. to cause dialects to influence each other. This raises the basic question to what extent one can pinpoint what is "original" (or: pure) Northern Thai, Northeastern Thai, etc.

A. PHONOLOGICAL RECONSTRUCTION

Since the early sixties the conceptual framework of research on Thai language history has been somewhat reminiscent of that of Romance linguistics, the gross features of the reconstructed ancestral language being taken essentially for granted, and the phenomena of modern dialects being derived from this source, much as Romance languages and dialects are ultimately derived from Latin. Needless to say, in the case of a reconstructed source language the explanatory advances in "historical" linguistics will in actual practice go both ways: inferences from modern data serve to refine the reconstruction, i.e. the present is used to explain the past, just as the confrontation of different (attested or reconstructed) chronological stages entails the use of the past to explain the present (with a very free quotation from W. Labov).

It is, however, no simple matter to determine what chronological stage to reconstruct in order to account for the modern dialects of Thai. As for the attested older stages, the language of the Ayudhaya period (14th to 18th century A.D.) has a special provenance, and cannot be set up as a common denominator. It is rather different with the language of the preceding Sukhothai period, which is archaic enough to be a useful point of reference. King Ramkanhaeng the Great created the Thai script in 1283 (on the basis of Mon and Khmer scripts), thanks to a long stone inscription from 1292 written in this script (as well as later material) the structure of Old Siamese of 700 years ago is well documented. Egerod (1961, p. 74) takes this stage of the language as his point of reference in dialect comparison, though *"a few features, especially in Southern Thai, seem to antedate Old Siamese"*. The Sukhothai language (Old Siamese) he takes to be *"the direct ancestor of Central Thai of today"*.

Needless to say, it takes an interpretation of the phonological status and the phonetic value (or values) of each letter of the Sukhothai script to arrive at a transcription of Old Siamese which can be used directly for reference in work on sound change and dialect splits in Thai. If Central Thai is taken to be a continuation of Old Siamese, reconstruction of the latter naturally involves a backward projection from the pronunciations of words in the modern language, as well as a consideration of the discrepancies between phonology and orthography in the modern language which may be taken to reflect phonological changes having taken place since King Ramkamhaeng's time. The latter aspect (conservative spellings in modern Thai) is highly relevant because the orthography is quite faithful to tradition. Most of the changes in question appear as phonological mergers, e.g. between voiced and voiceless sonorants, or between the old diphthongs */aɨ/ and */aɯ/ (with the consequence for modern Thai that the pronunciation is largely predictable from spelling but not the other way round).

Another important source of knowledge (or of qualified assumptions) about the phonology of Old Siamese is the orthographical rendering of old loanwords, e.g. from Sanskrit. These are not a priori proof of the standard values of the Thai letters at the time the loanwords were borrowed, however. For one thing, it might be the case that some letters of the Sukhothai script were used with special phonetic values in Indic loanwords (possibly more in accordance with the values of the related letters in the Indic script in question), so that these do not count in assessing the values of the letters in genuine Thai words of that period. Secondly, we cannot know a priori to what extent the pronunciation of such loanwords in Old Siamese differed from their pronunciation in the lending language. However, it is known that the creator of the Sukhothai script was well versed in Sanskrit and Pāli, and it is most natural to assume that the letters had largely the values one would expect from the renderings of loanwords, though this creates something of a gap between the phonology of Modern Thai and the phonology of Old Siamese. To take just one (important) section of the phonology, there is in fact overwhelming comparative evidence corroborating the assumption that of the letters that are used to represent an aspirated stop in Modern Siamese only some had this value in Old Siamese, others representing a voiced stop, cf. loanwords such as /phút(tha-)/ 'Buddha', /thêep, theewaa/ 'divinity, god' in Modern Thai with obvious cognates with voiced initial in Indic.

On this and on several other points Old Siamese seems to have been closer to Proto-Tai, as reconstructed by Haudricourt, Li and others, than to Modern Thai. The old system of stop consonants, if exemplified by the labials, is supposed to have looked as follows with its modern reflexes:

Old	Modern
*ph	ph
*p	p
*b	ph
*ʔb	b

the reflex of old /*ph/ being written with a so-called "high consonant" in Thai, and the reflex of old /*b/ with a so-called "low consonant". In the Thai script there is a similar distinction between "high" and "low" consonant letters for the voiceless fricatives, which likewise reflects the old manner distinction.

In the case of sonorants there was likewise a distinction between two categories. One category was constituted by initials combined with voicelessness (i.e. /*m̥/ or /*hm/, etc. - the

Thai script spells these as sequences with initial "high" h followed by the ordinary "low consonant" letter representing the sonorant in question). The other category was constituted by voiced initials (/*m/, etc., represented in Thai script as single, "low" consonants). With /*j/ there are even reflexes of a three-way contrast between voiceless, plain voiced, and glottalized (/*j̊/ or /*hj/; /*j/; /*ʔj/).

So much for initials. Looking now at the remainder of the syllable we find three kinds of components expressed in writing, viz. a vocalic nucleus, a final consonantal part (which is not an obligatory component of the syllable), and a component expressed by presence or absence of the so-called "tone marks". It is generally assumed that these tone marks, of which the more common ones are "máj èek" (the Arabic figure "1" as a diacritic over the consonant letter before the vowel) and "máj thoo" (the Arabic figure "2" placed in the same fashion), were used in Old Siamese to represent different (to a greater or lesser extent tonal) prosodies from which the present tonal systems of the various dialects have developed (see section B below). According to the authoritative Proto-Tai reconstructions of scholars such as Haudricourt and Li, these prosodies are reflexes of an original system comprising three main categories: A (no tone mark in Thai script), B ("máj èek"), and C ("máj thoo"), each of which defines a laryngeal state or phonation type in the final part of the syllable.

The distinctions reflected by the tone marks in Thai script intersect with a distinction between so-called "live" and "dead" syllables, the latter being syllables ending in a stop consonant, and also with a distinction between long and short vowels (cf. the remarks on syllable structure in Modern Thai in section II A above). For "Ancient Thai", a proto-language (much antedating Old Siamese of the Sukhothai period) which Brown (1965) sets up as the frame of reference for his diachronic study, Brown himself posits an integrated system comprising four final laryngeal components plus a distinction of length versus shortness, which combine to form a total of five components: whisper ("w"), voice ("v"), creaky ("c"), glottalization with length ("longstop", "q"), and glottalization with shortness ("shortstop", "k"). Now, how do these compare to the phonetic interpretations suggested for the Proto-Tai categories A, B, and C by other scholars? Haudricourt (in the "Additional Note" to the 1972-version of his paper on tonal splitting, as formulated by Court) finds that there is partial agreement between Brown's reconstruction and his own in that they both have glottalization for category C, but that they differ on assigning features or components to categories A and B. Brown's "whisper" occurs in the case of category A, and his "voice" in the case of category B, whereas Haudricourt reconstructs rather the opposite for cognates in Austroasiatic and Proto-Miao: "voiced final vowel or sonorant" for category A, and "final -h or other fricative" for category B. - This discrepancy is indicative of the rather hypothetical status of these phonetic interpretations of reconstructions.

B. DIACHRONY: THE ORIGIN OF THAI TONES

Although the diachronic literature deals with quite a number of changes from the proto-language to Modern Thai (some of which are dialect specific and others more general), the only issue to be considered in detail here is the loss of manner distinctions in consonants and the concomitant development of tones in Thai. This is a challenging issue because of the different tonal developments in the various Tai languages and also within the Thai dialects in the narrower sense. The presentation here will deal only summarily with the diachronic reorganization of syllables by which consonantal distinctions were replaced by tonal ones, and the attempts that have been made within the last two decades to explain the changes in terms of phonetic and psychological processes. However, the survey is intended to be explicit enough to pinpoint what is obvious and what is controversial with regard to TONOGENESIS in Thai (for the concept "tonogenesis", see Matisoff 1973, Henderson 1982). Put differently, the summaries and comments in this and the following section are directed at questions such as (1) How did the Thai tones come into being? and (2) is the development in Thai illustrative of universal mechanisms causing tonogenesis to happen?

When looking at the possible connection between laryngeal features of syllable initials + finals and tonogenesis, it may be useful to keep in mind that in principle tonal contrasts or specific tonal manifestations can arise in a number of ways. For one thing there may sometimes be phonetically different pitch contours associated with long and short vowels, and under certain conditions such contours may conceivably come to acquire the status of different tones; likewise, contraction of two consecutive syllables into one may be the source of a tonal contrast with items that were monosyllabic "from the start". Such sources of tones may be relevant in explaining the tones of Proto-Tai and very early Thai (ideas along these lines have been expressed in passing by various scholars but not worked out in detail for Thai).

To mention quite another type of source, specific shapes of tonal contours may be borrowed from one dialect into another (cf. Chamberlain 1972a). It has been argued (Brown 1965, p. 157) that coalescences or splits in tonal systems are not borrowed, but I am not at all convinced (in spite of Chamberlain 1972a) that this holds true as a principle; it is at any rate clear - as evidenced par excellence by South-East Asian languages - that non-tonal languages may become tonal by diffusion of an areal feature of tonemicity. At the very least, one must admit that the tendency to develop tones out of other properties of syllables may spread as an areal phenomenon; it is an important empirical issue how far this idea (also cf. Brown 1965, p. 62) can explain the facts without the assumption of direct borrowing of tonal contrasts.

However, as mentioned already, everybody agrees that there is - in Thai and many other Asian tone languages - a direct connection between the development of lexical tone and the loss of certain contrasts among initial consonants. This is what Brown (1975) refers to as the GREAT TONE SPLIT. I think the best way to state briefly what this is all about, is to cite the introductory passage from Brown's paper:

> *"The great tone split was a sound change that swept through China and northern Southeast Asia nearly a thousand years ago. It was probably the 'greatest' sound change we have record of today, for it affected almost all of the words of almost all of the languages of this vast area (...) Simply put, voiced, glottal, and aspirate initial consonants split all existing tones in two (or three) and then partially coalesced, thereby shifting some laryngeal distinctions from initials to tones. A typical example is shown below.*
>
> | | | |
> |---|---|---|
> | *phaa* | *pháa* | *pháa* |
> | *paa* | *paa* | *paa* |
> | *baa* | *bàa* | *phàa* " |

(Brown 1975, p. 33).

There are three points to be made about the general nature of this great tone split: (1) It should be noted that the split into different tones (e.g. high, mid, and low, as suggested by the transcription above) has an obvious affinity to the perturbating effect of initial consonants on the pitch (fundamental frequency) of a following vowel which is well-known from numerous acoustic-phonetic studies of living languages (cf. references in section III above); hence the relationship between tonogenesis and mechanical pitch perturbation is a crucial issue. - (2) It should be kept in mind that what occurs with the great tone split is not necessarily tonogenesis in the typological sense that a language becomes tonal; rather, one must allow for two diachronic possibilities: (a) that an already existing tone system was multiplied, as it were, by the tone split (this is assumed for languages such as Chinese and Thai), or (b) that a previously non-tonal language became tonal. - (3) The great tone split does not always work in the phonetically transparent way illustrated above; on the contrary, the tonal reflexes may be totally at variance with the phonetic predictions. (This crux, which is referred to in this paper as the "tone split paradox", will be dealt with in section C below.)

To return to the general origin of the tones of modern Thai in its dialects, the idea is that (as mentioned already in section A) the tones stem from the interplay between the pitch-perturbating initial components and various properties of the remainder of the syllable, viz. properties that are expressed in Thai writing as short versus long vowel, "live" syllable (resonant termination, including open syllable) versus "dead" syllable (non-resonant termination), and the choice of "tone mark". The intersections of these dimensions create a roster of slots which are filled differently for different dialects (with or without mergers between the tonal reflexes in the various slots, and with different manifestations of the individual tones). The filling out and confrontation of such tone charts (on the basis of Gedney 1973) has become one of the major concerns of comparative Thai dialectology.

As a result of the complex interplay of various components it holds true for all modern dialects that the tones are only very indirectly linked to the old tone marks. In Central Thai, for example, "máj èek" has both falling and low tone as reflexes (depending on the etymological status of the initial consonant(s)), "máj thoo" has both high and falling reflexes (likewise depending on the initial), and "zero" (absence of tone mark) may correspond to any of the five tones in the modern language, i.e., mid, rising, high, falling, or low tone (depending on vowel length, on the specific combination of initial and final consonants, and on the etymological status of the initial). In other dialects the details are more or less different but in principle the picture is similar.

Because of the conservative character of Thai orthography the tone of a particular syllable is normally retrievable from its spelling (according to the orthoepic rules of the dialect in question), at least for Central Thai. The orthographical representation of tones, on the other hand, is only in part predictable because of mergers among initials, such as

$$\left.\begin{array}{l} {}^{*}\mathrm{ph} \\ {}^{*}\mathrm{b} \end{array}\right\} > \mathrm{ph} \qquad\qquad \left.\begin{array}{l} {}^{*}\mathring{\mathrm{j}}\ (\mathrm{hj}) \\ {}^{*}\mathrm{j} \\ {}^{*}\mathrm{ʔj} \end{array}\right\} > \mathrm{j}$$

which combine with partial mergers among the prosodies represented by tone marks. These latter mergers are dialect specific. In Standard Thai (Central Thai), for example, it is so that the combination "high" consonant (± "low" consonant) + "máj thoo" (in "live" or "dead" syllable) has merged with the combination "low" consonant + "máj èek" or "dead" syllable, the result being in all cases a falling tone. Thus sequences such as /phâa/, /nâa/, /phâap/ are in principle (and often in actual practice) genuinely ambiguous with regard to their etymological source and their spelling.

After this lengthy introduction to the tonogenesis issue, the possibility of tracing and of explaining the actual mechanism of diachronic change will be dealt with in the next section.

C. TONOGENESIS MECHANISMS AND THE TONE SPLIT PARADOX

Taking for granted that the present tones of Thai reflect the interplay between laryngeal properties of the syllable onset and laryngeal properties of the remainder (tones or other old prosodies, etc.) in the proto-language the next question is how to trace what really happened in the process of tone split in Thai (and possibly in all of northern Southeast Asia and China).

Brown (1965) describes the development from Ancient Thai to modern Thai dialects in terms of REGISTER and CONTOUR. Register, he says, developed *"as the initials unloaded distinctions unto the tones"*, Modern Thai having three such registers, viz. R1 characterized by low pitch, R2 characterized by mid pitch, and R3 characterized by high pitch. These registers (not to be confused with what is now called "register" in Mon-Khmer studies!) he assumes to be controlled by the crico-thyroid and thyro-arytenoid muscles (which are indeed the major pitch-controlling muscles). Contours he supposes to have developed gradually and to have become associated with specific registers; it might then happen that register distinctions were lost, and the distinction was carried by contours alone (p. 58). He distinguishes three kinds of contour for Modern Thai, viz. C1 appearing as low dull tone, C2 appearing as mid normal tone, and C3 appearing as high bright tone. These contours are supposed by Brown to be produced by different degrees of contraction (due to rotary movements of, or pressure on, the arytenoids as controlled by the lateral crico-arytenoid muscles): strong contraction = C1, mid contraction = C2, and weak contraction = C3. The idea, then, is that the actual pitch movements of tones in Thai dialects reflect combinations of contour and register with syllable final components ("endings") although, as he points out himself, the phonetic reality cannot be derived very directly from such a componential representation.

As for the emergence of tonal differences among dialects, Brown claims that endings were the most stable components, contours and registers varying much more from one dialect to another.

This whole analysis is ingenious and interesting, but the integration of very mechanistic phonetic considerations with a rather abstract componential analysis is in my view a problematic undertaking, both in the case of modern dialects (for which instrumental analyses are highly desirable) and much more in the case of the reconstructed proto-language.

The question is to what extent it is possible to predict the tonal effect of various kinds of <u>syllable onsets</u> and <u>syllable terminations</u>. As for the latter <u>there has been</u> <u>comparatively little research</u> on the effect of such differences as presence versus absence of a <u>final voiceless</u> consonant. This phonetic issue is particularly <u>relevant in</u> connection with attempts to explain the remote origin of the three (or more) prosodies in Proto-Tai and in other proto-languages of the area, which is a topic outside the scope of the present paper (see references to Haudricourt and others in section A above). It is, however, also relevant to the understanding of the later development in Thai (cf. the importance of the concepts "dead" and "live" syllable), as is the possible effect of the presence or absence of a glottal syllable termination (glottal constriction or closure), see e.g. Egerod (1971). Here is a field in which more instrumental research on present-day languages of the area is called for.

It is at first sight different with the effect of syllable onsets on the pitch of the remainder of the syllable. The general pitch-perturbating effects of resonants versus obstruents, or of voiced versus voiceless consonants, in syllable initial position are well-known, cf. the survey of these effects in Lea (1973). As for Thai, there has recently been a considerable amount of research serving, a.o., to show the interrelations between consonant articulation, durational features, and pitch movements on the basis of contemporary insights into speech physiology and with the use of modern apparatus. This research (see e.g. Abramson 1975a, Erickson 1975, Gandour 1974b, Gandour and Maddieson 1976) of course deals with contemporary Thai but is in part done with a direct view to the diachronic perspectives of the findings.

The general phonetic literature clearly points to voiced stops having a pitch-lowering influence on the onset of a following vowel as compared to voiceless stops, although Painter (1978, p. 263, 265, 273-274) finds that *"the actual difference in frequency between the sets is very small compared to the difference between high and low tones* [in tone languages such as Yoruba, JR]" and expresses scepticism toward the current explanation of tonogenesis as due to this effect of consonants on vowel pitch.

It is much less clear whether aspirated voiceless stops cause a higher pitch than do unaspirated voiceless stops. There is simply conflicting evidence (also cf. section III above). One conceivable reason for this is that unaspirated stops may be produced in different ways in different languages, and perhaps even within one language or dialect: with or without glottalization, with or without supraglottal constriction (e.g. the type of tongue root retraction observed for Thai /p t/), and with or without vertical movement of the larynx (cf. Gandour and Maddieson 1976). Unless such parameters - which have an aerodynamic influence - are totally under control it seems hard to make precise predictions about the pitch perturbating effect of consonant articulation. I think this should be kept in mind when the origin of tones as mechanical pitch perturbations is at issue, especially because it may be hard to reconstruct such details of articulation (unless one renounces on independent evidence and allows for complete circularity, in which case the reconstruction is of limited explanatory interest, of course).

To judge from the phonetic literature, then, there is nothing strange with the tonal development depicted schematically by Brown (1975) in the quotation above (section B). Languages or dialects in which the tonal reflex after old voiced consonants (e.g. /*b/ > /ph/ in Thai) is low-pitched, i.e. what Brown calls "voiced-low" dialects, behave according to expectation. But then the above mentioned tone split paradox presents itself: some dialects, on the contrary, have a high-pitched tonal reflex after old voiced consonants, i.e., they are "voiced-high". In his 1965-monograph Brown ventured a physiological explanation of the paradox that - in the framework of his description - the initial components aspirated, glottalized, and voiced have caused respectively high, mid, and low register in some dialects but respectively low, mid, and high in others. Brown suggests that there were two possible ways for the vocal cord adjustment to react to aspiration; as I understand him the idea is that the vocal cords might be in a state of tightness, and a pitch rise would then result from the opening of the glottis, or they might be slackened so as to give in, with a resulting pitch drop in the latter case. These two options being available some dialects chose one, others the other option.

It is hard to evaluate the physiological plausibility of such an explanation, although it does not seem *a priori* unlikely that some of the variability in the general phonetic results concerning pitch perturbation after consonants have to do with the existence of different production mechanisms for aspirates. However, as long as there is not independent evidence for such differences in laryngeal adjustment (with exactly the desired effect) to have existed in the proto-language, it seems that the question must be left totally open.

In his 1975-paper Brown approaches the question from a quite different angle. He no longer claims that there were variable production mechanisms underlying the differential tonal developments of the dialects; instead he states that

> *"There seems to be complete agreement in the literature on how different consonant types go about splitting tones: voiced initials tend to lower tones and aspirate initials tend to raise them; and the tones with glottal initials* [i.e., consonants of /*p/-type as well as consonants of /*ʔb/-type, cf. Strecker 1979a, p.77] *get drawn one way or the other (giving two different levels for the old tone) or else stay in between (giving three)."*

(Brown 1975, p. 33).

Nevertheless he finds that most Thai dialects, with the exception of Southern Thai dialects, show definite voiced-high correlations, that is, exactly the opposite of the pattern one would expect from the universal phonetic tendencies.

Before going into the crucial issue: Brown's new explanation of the tone split paradox, I shall summarize how he and Strecker

(1979b) explain a number of more specific discrepancies among the tonal systems of the various dialects (within the "voiced-high" and the "voiced-low" group, respectively). Brown refers to the operation of two forces working on register and contour (tonal components in his analysis, see the beginning of this section), viz. "the principle of least effort" and "the need to maintain distinctions" (p. 40). In the extensive and impressive follow-up paper by Strecker (1979b) it is pointed out that the latter force (maintenance of tonal contrast) may cause tones that are too close to move further apart either by changes in overall pitch or by modifications of pitch contours, such modifications (rises or falls) affecting either the final or (e.g. in the Chiengmai dialect of Northern Thailand) the initial part of the tones. - It is immediately seen that the principle of least effort and the maintenance of distinctions (or, in Strecker's terminology, the principle of "psychological distance") may be in conflict within individual tonal systems; this leaves the linguist with a powerful tool since there is considerable freedom in explaining why the tones came out differently in different dialects, viz. as a consequence of one or the other force taking the lead in each individual case. The whole explanatory frame becomes even more powerful because it is necessary to assume some kind of tonal reorganization (so that in principle dialects which are now voiced-low may be originally voiced-high, and vice versa).

In addition to the more abstract forces mentioned above Strecker further finds evidence for two general tendencies: high-falling tones tend to fall more than medium- or low-falling tones, and conversely, low-rising tones tend to rise more than medium- or high-rising tones. These tendencies are documented with a wealth of material including a wide variety of Tai dialects inside and outside Thailand. - Interestingly enough, there seems to be a correlation with the voiced-high/ voiced-low categorization: a high-falling tone falls relatively much in a voiced-high dialect but less in a voiced-low dialect. This, of course, fits neatly into the general idea of maintaining psychological distance between tones. However, as mentioned in Strecker's Appendix (p. 73), Southern Thai dialects actually run counter to the generalization by being voiced-low but having two falling tones of which the higher falls longer than the lower one (an observation which Strecker incidentally mentions as a possible stumbling-stone for the whole explanatory framework).

After this digression I shall return to the main issue: why are there "voiced-high" in addition to "voiced-low" dialects? Brown (1975) now suggests that the explanation may be a psychological one having to do with the way listeners focus their attention when perceiving tonal contours. His explanation is based on the contention that the pitch in syllables beginning with voiced consonants is higher throughout most of the syllable although the pitch of the voiced consonant is lower, compared to syllables beginning with voiceless consonants. *"The natural working of the larynx in any synchronic description is"*, says Brown (p. 44), *"for tones or intonations in* *syllables with voiced initials to start lower but be higher"*. He claims that this is seen from the pitch contours for Thai in Erickson (1974) and for English in Lea (1973).

If it is indeed the case for Thai that the tonal contour found after a voiced initial (for a given tone) typically rises above the contour after a voiceless initial, we may have here an interesting case of compensation: *"The tone signal is such as to produce the desired pitch regardless of environmental factors"* (Brown 1975 p. 44). This may well be essential in a tone language (as against a non-tonal language). Now, Brown further speculates that the listening strategy with regard to pitch may proceed in either of two ways, causing dialects to become voiced-high and voiced-low, respectively: speakers of the former dialects are assumed to "start at the center of a syllable and work out", whereas speakers of the latter dialects "start at the point where the vowel begins and work out" (p. 45). A dialect may thus become voiced-high if a listener narrows his focus down to the syllable center and hears the pitch as higher after the voiced consonant (by successive imitation, the argumentation continues, pitch allophones emerge which are dissociated from the initials and reinterpreted as phonemic tones). - Strecker (1979) tends to accept Brown's hypothesis (which, by the way, Brown himself carefully refers to as "sheer speculation", p. 45), and provides arguments against other conceivable explanations, some of which are taken up below on the basis of Strecker's excellent survey.

Before looking in detail at Brown's specific suggestion one may address the more general question: is it likely that a relatively low pitch after voiced initial can be reinterpreted categorially as a higher tone than the pitch after voiceless initial, and are there perceptual mechanisms and strategies which may account for such a categorization? The important paper by Abramson and Erickson (1978), using synthetic stimuli presented to Thai listeners, suggests that the categorization is non-trivially related to the choice of initial consonant, so this is definitely an area that deserves further consideration in tonogenesis studies. - To return to the alleged cross-over of pitch contours after voiced and voiceless initial, Brown refers a.o. to the graphs in Lea (1973). I do not find Brown's interpretation of these compelling, however, and altogether the phonetic literature is not indicative of a general tendency for the lower initial pitch to be followed by a higher final pitch after voiced than after voiceless initial. Moreover, Brown's ingenious explanation replaces one crux by another: is it likely that listeners begin to focus on differences between parts of the pitch contour if the alleged cross-over is not mechanically caused but is perceptually motivated in the first place, viz. as a means to ensure perceptual sameness?

Finally, the assumption that a speaker may listen selectively to either part of the pitch contour or the whole pitch contour of a syllable and perceive the selected part as the overall pitch is, if not a priori inconceivable, at any rate highly radical if seen on the background of current assumptions about pitch perception.

Thus, as I see it, we are still not much closer to a real explanation of the tone split paradox.

One alternative to the across-the-board hypotheses of Brown and Strecker (attempting to explain the scenario on the basis of principles affecting the system as such), is to assume that tone shifts are not totally dependent on the system: the voiced-high situation may arise through unsystematic changes affecting individual tones. Brown and Strecker (p. 55) are most decidedly not in favour of such an explanation, but considering the general limitations on predictability in historical phonology I think we have to admit that this may be precisely what happened in voiced-high dialects.

Another interesting possibility is that the developments have to do with the chronology of events; Strecker (p. 56) cites an unpublished paper by Jay Fippinger for the following suggestion:

> *"(...) certain consonant changes took place before the tones underwent splitting, and (...) one set of consonant changes took place in the voiced-low dialects and a different set took place in the voiced-high dialect. Therefore the consonantal environment affecting the tones was different in the two types of dialects, which explains why the tones underwent a different type of splitting in each case."*

I wish to mention one aspect of the consonant development which may be relevant here. One of the developments assumed to be involved in the tone split is the change from initial /*b/ to /ph/ (and similarly for the other points of articulation). Now, such a change of a voiced stop into a voiceless aspirate is in itself a real crux. Why did /*b/ not coalesce with /*p/ instead of bypassing it, as it were? One possibility is that /*p/ was set apart by being glottalized; this, however, does not account for the phonetically strange line of development [b] > [ph] (via [p], or what?). Egerod (1961, p. 76-77) speculates whether we may have instead a development from a voiced aspirated stop via a voiceless stop with voiced aspiration to the voiceless aspirated stop of Modern Thai, that is, something like [bɦ] > [pɦ] > [ph]. (There seems to be a parallel to the suggested development of voiced into voiceless aspiration in Chinese.) I am not sure that there is an easy phonetic explanation for such a chain of events, but at least it represents a unidirectional process having to do with increasingly delayed voicing.

As far as I can see, it must remain an open question whether stops of /p/-type were glottalized or not at the time of the great tone split (in spite of Strecker's insistence that the reflexes of these consonants "are accompanied by simultaneous closure of the glottis" in most or all modern Tai dialects for which detailed phonetic descriptions are available, 1979, footnote 15; see section II B above for counter-evidence). *If*, further, it is likely that the consonants generally *reconstructed* as /*b/ etc. were, during a long time span, aspirates but with voicing specifications differing from one chronological stage to the other, one should be a little careful with a mechanistic application of the phonetic principle of pitch lowering after voiced as against voiceless stops. We simply do not know enough about the relevant parameters of consonant articulation in the proto-language, or at chronologically successive stages in the history of Thai.

Strecker (1979b, p. 57-58) also communicates that John Grima (personal communication) has proposed a phonetic explanation which ties the tone split to the development of *register* Understood as a split of *vowels* into two sets) *in Mon-Khmer* languages. One might, *according* to Grima, assume that the development in the Tai languages went through a register stage, viz. according to the following scheme (where "a̰a̰" stands for creaky, "a̤a̤" for breathy voice, and Arabic figures for components of a rising tonal contour):

	Consonant difference	>	Register difference	>	Tone difference
Voiced-high	paa 24		pa̰a̰ 24		paa 13
	baa 24		paa 24		paa 24
Voiced-low	paa 24		paa 24		paa 24
	baa 24		pa̤a̤ 24		paa 13

According to this hypothetical scheme the pitch contour is lowered in some dialects after voiceless initial via creaky phonation, in others after voiced initial via breathy phonation. Strecker mentions the observation that breathy voice still occurs after originally voiced stop in certain Southern Thai dialects (according to Egerod 1961) in favour of this hypothesis.

As I see it, both the relative chronology hypothesis and the register hypothesis are in principle open to verification or refutation by external evidence and worthy of careful consideration in future work. At the same time it must be taken for granted that the tone split paradox *may* ultimately turn out to be explicable only in rather *abstract* psychological terms referring to the way a tonal system is mentally represented and reorganized as a function of specific structural properties of the language and of general "forces" operating on tonal systems (if such forces can be empirically shown to exist).

D. ADDITIONAL REMARKS ABOUT SOUND CHANGES AND DIALECT DIFFERENCES

Although the preceding sections have been devoted to problems directly or indirectly associated with the "Great Tone Split" there were also other major changes which affected the phonological structure of Thai syllables. As for vowels, there were shifts particularly in the diphthongs (taken in the wide sense). Not only did the peculiar diphthong /aɯ̯/ merge with /aj/ (/ai̯/) in most areas (outside the Northeast), and the vowel /ɯ/ merge with /i/ (completely or partially) in Northern Thai dialects, but the spelling of the sequences now pronounced /aj aw iə ɯə uə/ is clearly suggestive of major changes from Old Siamese (also cf. the particular, diphthong-like spelling of the vocalic complex now pronounced /ɔ́ʔ/). Further, there were (more or less dialect-specific) changes in vowel length, so that it is now only partly predictable from spelling what length to assign to a vowel of the set /ɛ a ɔ/ (see further Hartmann 1976, Brown 1979).

In initial consonant clusters there is a widespread loss of the liquid in the sequence muta cum liquida (/paa/ for /plaa/, etc.). Altogether, the phoneme /r/ has been the most vulnerable member of the initial consonant inventory, changing either to /l/, /h/, or (in the just-mentioned clusters) zero. There is considerable dialectal variation associated with the /f/-/khw/ complex. It may be mentioned also that /*ɲ/ changed to /j/ in most dialects. These phenomena are being studied in Thailand in a comparative dialectologist framework and in a sociolinguistic perspective (cf. Beebe 1975, Tanwattananun 1982).

The most pervasive changes took place in syllable-final position, all finals (in inherited words as well as loanwords) eventually merging into a very small set of possible consonant types (cf. section IIa above on permissible finals). As with tonogenesis we are here faced with an important areal phenomenon posing a host of inherently interesting questions (How old are the present severe limitations on possible finals?, How are the old loanword spellings to be interpreted phonologically with reference to Old Siamese pronunciation?, etc.). However, this issue is not debated in the current literature on Thai to the same extent as the tonogenesis issue, and the present author does not feel in a position to give a concise and adequate appraisal of the Stand der Forschung in this area.

VI. FINAL REMARKS

There has been an impressive activity within phonetic and phonemic research on Thai within the last few decades. This research was initiated by predominantly Western scholars but Thai scholars are now doing extensive work in Thai dialect geography and Thai phonetics (which may not be so apparent from the present paper because of its emphasis on issues with a general linguistic perspective, and on presentations of these issues in print). Apart from historical phonology, for which the recent study of a variety of dialects is extremely fruitful though little known outside the sphere of Thai specialists, there are several such general issues which invite further study in a Thai context. Thus the pitch-perturbating effect of initial consonants deserves a cross-dialect study, since many Thai dialects seem basically very alike in segmental phonetics though they differ widely in tone shapes. As regards dialects, I wish to mention quite another type of inherently interesting issue, viz.: what is the relative importance of tonal versus segmental differences for the mutual intelligibility among dialects (in cases where lexical differences can be eliminated from consideration)? As the tones of the various dialects differ very much in distribution and phonetic shape (e.g., high for low, and vice versa), this provides an ideal case for studying language users' ability to "normalize" the speech input from other dialects even in case of drastic differences in tone pattern, and thus to understand each other after some minimum of adaptation. (The situation is the more interesting since segmental dialect differences are surprisingly few in comparison with the tone differences.)

VII. BIBLIOGRAPHY

The bibliographical references given below are supposed to illustrate the high level of activity within the synchronic and diachronic study of Thai phonetics and phonology within the last few decades. For this reason I have not limited the list strictly to references which I consider central to the issues raised in this paper, and I have not even limited it to items I have had access to myself: quite a few items are secondary references only. This is true in particular with regard to several of M.A. or Ph.D. theses (dissertations) listed here. The reader should be aware that the addition of the word thesis or diss. in parentheses signals that the work may be unpublished and not easily available (and that the reference may perhaps be inadequate). The reason for this somewhat risky listing of (possibly spurious) references is that the very existence of such work is illustrative of the considerable interest and zeal devoted to this field by Thai students and others. (Other unpublished manuscripts, however, are as a rule disregarded, even in the case of important congress papers.)

The list includes not only work on Central Thai but also some items concerning other dialects of this language. On the other hand, I have considered it appropriate to limit the scope - by and large - to Thai proper, i.e. not to include work dealing with other Tai languages, even such that are spoken in Thailand, to any considerable degree. On the other hand, the list includes several items within the field of comparative Tai study, because the comparative aspect, and the reconstruction of Proto-Tai, are of paramount importance for the ways in which research on Thai phonetics and phonology has developed.

Needless to say, the bibliography makes no claim whatsoever with regard to completeness of coverage of any aspect of Thai studies; it is absolutely only meant to be illustrative of what is going on. (Also see Note p. 93.)

Abramson, A.S. 1962: "The vowels and tones of Standard Thai: acoustic measurements and experiments", Bloomington Indiana Univ. Res. Center in Anthropology, Folklore and Linguistics, Publication 20 (= IJAL 28.2, Part III)

Abramson, A.S. 1971: "Tonal experiments in whispered Thai", in: A. Valdman ed.: Studies in French and general linguistics and phonetics to the memory of Pierre Delattre (The Hague: Mouton) (also cf. Haskins SR-19/20, 1969, p. 37-57

Abramson, A.S. 1972: "Word-final stops in Thai", in: Harris & Noss 1972, p. 1-7 (also Haskins SR-28, 1971, p. 193-201)

Abramson, A.S. 1974: "Experimental phonetics in phonology: vowel duration in Thai", Pasaa IV.1, p. 71-91

Abramson, A.S. 1975a: "Pitch in the perception of voicing states in Thai: diachronic implications", Haskins SR-41, p. 165-174

Abramson, A.S. 1975b: "The tones of Central Thai: some perceptual experiments", in: Harris & Chamberlain 1975, p. 1-16

Abramson, A.S. 1976: "Thai tones as a reference system", in: Gething, Harris & Kullavanijaya 1976, p. 1-12 (also Occ. Papers 17, Univ. of Essex 1976, p. 1-14, cf. Haskins SR-44, 1975, p. 119-126)

Abramson, A.S. 1978: "The phonetic plausibility of the segmentation of tones in Thai phonology", Haskins SR-53.1, p. 73-78

Abramson, A.S. 1979a: "The coarticulation of tones: an acoustic study of Thai", in: Thongkum et al. 1979, p. 1-9 (also Haskins SR-44, 1975, p. 119-126)

Abramson, A.S. 1979b: "Lexical tone and sentence prosody in Thai", Proc. 9th Int. Congr. Phon. Sc. II, p. 380-387 (Inst. of Phonetics, Univ. of Copenhagen) (also see Haskins SR-59/60, 1979, p. 153-160)

Abramson, A.S. 1979c: "The noncategorical perception of tone categories in Thai", Haskins SR-51/52, p. 91-100 and Lindblom & Öhman eds.: Frontiers in Speech Communication Research, p. 127-134 (London)

Abramson, A.S. & Donna M. Erickson 1978: "Diachronic tone splits and voicing shifts in Thai: some perceptual data", Haskins SR-53.2, p. 85-96

Abramson, A.S. & Katyanee Svastikula 1983: "Intersections of tone and intonation in Thai", Haskins SR-74/75, p. 143-154, also cf. Preprints, Working Group on Intonation, The XIIIth Int. Cong. of Linguists, Tokyo 1982, p. 13-22

Abstracts of thesis dissertations written by faculty & students of the Department of Linguistics, Chulalongkorn University, published (in Thai and English) 1983

Bee, Peter 1975: "Restricted phonology in certain Thai linker-syllables", in: Harris & Chamberlain 1975, p. 17-32

Beebe, Leslie M. 1975: "Consonant cluster simplification in Bangkok Thai", Int. J. of the Sociology of Language 5, p. 43-61

Beebe, Leslie M. 1976: "Social conditioning of grooved and flat fricatives in Thai", in: Gething, Harris & Kullavanijaya 1976, p. 13-27

Bradley, C.B. 1911: "Graphic analysis of the tone-accents of the Siamese language", JAOS XXXI, p. 282-289

Brière, E.J. & Sinuan Chiachanpong 1980: "An investigation of Thai interference in selected American-English phonemes", Papers and Studies in Contrastive Linguistics 11, p. 101-118 (Poznań)

Brown, J. Marvin 1965: From Ancient Thai to modern dialects (Soc. Sc. Found. Press of Thailand, Bangkok)

Brown, J. Marvin 1967 & 1968: A.U.A. Language Center Thai Course, Book 1 & 2

Brown, J. Marvin 1975: "The great tone split: did it work in two opposite ways?", in: Harris & Chamberlain 1975, p. 33-48

Brown, J. Marvin 1976: "Dead consonants or dead tones?", in: Gething, Harris & Kullavanijaya 1976, p. 28-38

Brown, J. Marvin 1979: "Vowel length in Thai", in: Thongkum et al. 1979, p. 10-25

Brudhiprabha, Prapart, Achara Wangsotorn, & Joe D. Palmer eds.: A Festschrift for Richard B. Noss = Pasaa VI.1-2, Spec. Issue

Bunnag, Suriya 1970: Contribution à la recherche du Thai commun à travers les inscriptions de l'époque Sukhodaya (thesis, Univ. of Paris)

Chamberlain, James R. 1972a: "Tone borrowing in five North-eastern dialects", in: Harris & Noss 1972, p. 43-46

Chamberlain, James R. 1972b: "The origin of the Southwestern Tai", Bull. des amis du Royaume Lao 7-8, p. 233-244

Chamberlain, James R. 1975: "A new look at the history and classification of the Tai languages", in: Harris & Chamberlain 1975, p. 49-68

Chantavibulya, Vichintana 1959a: The phonology of the syllable in Songkhla, a Southern Thai dialect (M.A. thesis, SOAS, London)

Chantavibulya, Vichin 1959b: "wítthájaaniphon rûəŋ khwaamtὲεk-tàaŋ ráwàaŋ phaasǎa kruŋthêep lέ' phaasǎa sǒŋkhlǎa" [a thesis on the differences between the language of Bangkok and the language of Songkhla]

Court, Christopher 1972: "The tones of the Traat dialect", in: Harris & Noss 1972, p. 47-49

Court, Christopher 1975: "The segmental and suprasegmental representation of Malay loanwords in Satun Thai: a description with historical remarks", in: Harris & Chamberlain 1975, p. 67-88

Davies, Somchit 1979: "A comparative study of Yong and Standard Thai", in: Thongkum et al. 1979, p. 26-48

Debavalya, Kesmanee 1982: A linguistic borderline between Central Thai and Southern Thai: a tonal study (thesis), see Abstracts .. Chulalongkorn Univ. 1983, p. 14

Deepuengton, Pawadee 1982: Distortion of tones in Thai country songs (thesis), see Abstracts .. Chul. Univ. 1983, p. 15

Diller, Anthony 1976a: "A note on PT *ml-", in: Gething, Harris & Kullavanijaya 1976, p. 39-46

Diller, Anthony 1976b: Towards a model of Southern Thai diglossic speech variation (diss., Cornell Univ.)

Diller, Anthony 1979: "Tones, segments and Thai regional society", in: Thongkum et al. 1979, p. 60-93

Donald, S. Lea 1977: "Phonetic and auditory aspects of adaptation: evidence from Thai voicing contrasts", Haskins SR-51/52, p. 175-182

Drans, J. 1942: "Contribution à l'étude de la phonétique siamoise, les consonnes nasales initiales", Publ. Maison Franco-Japonaise B.I. (Tokyo)

Egerod, S. 1961: "Studies in Thai dialectology", Acta Orientalia 26, p. 43-91

Egerod, S. 1971: "Phonation types in Chinese and South East Asian languages", Acta Ling. Hafniensia 13, p. 159-171

Egerod, Søren 1972: "The tones of Phuket and Ko Samui: a correction", in: Harris & Noss 1972, p. 50-51

Erickson, Donna 1974: "Fundamental frequency contours of the tones of Standard Thai", Pasaa 4.1, p. 1-25

Erickson, Donna 1975: "Historical implications for an historical account of tonogenesis in Thai", in: Harris & Chamberlain 1975, p. 100-111

Erickson, Donna 1976: A physiological analysis of the tones of Thai (diss., Univ. of Connecticut)

Foreit, K.G. 1977: "Linguistic relativism and selective adaptation for speech: a comparative study of English and Thai", Percept. & Psychophysics 21, p. 347-351

Fowler, M. & T. Isanasena 1952: The total distribution of the sounds of Siamese. A topographical analysis (Univ. of Wisconsin Press)

Gandour, J.T. 1974a: "The glottal stop in Siamese: predictability in phonological description", Pasaa 4.2 (also UCLA WPP 27, p. 84-91)

Gandour, J.T. 1974b: "Consonant types and tone in Siamese", J. of Phonetics 2, p. 337-350 (also UCLA WPP 27, p. 92-117)

Gandour, J. 1975: "On the representation of tone in Siamese", in: Harris & Chamberlain 1975, p. 170-195 (also UCLA WPP 27, 1974, p. 118-146)

Gandour, J.T. 1976a: "A reanalysis of some phonological rules in Thai", in: Gething, Harris & Kullavanijaya 1976, p. 47-61

Gandour, J. 1976b: Aspects of Thai tone (diss. UCLA, Abstract in UCLA WPP 33, p. 20-22)

Gandour, J. 1976c: "'Counterfeit tones' in the speech of Southern Thai bidialectals", UCLA WPP 33, p. 3-19

Gandour, J. 1977: "On the interaction between tone and vowel length: evidence from Thai dialects", Phonetica 34.1, p. 1-80

Gandour, Jackson T. 1978: "The perception of tone", in: V.A. Fromkin ed.: Tone, a Linguistic Survey, p. 41-76

Gandour, J.T. 1979: "Tonal rules for English loanwords in Thai", in: Thongkum et al. 1979, p. 94-105

Gandour, J. & R. Dardarananda 1982: "Voice onset time in aphasia: Thai I. Perception", Journal of Brain & Language 17.1, p. 24-33

Gandour, J. & I. Maddieson 1976: "Measuring larynx movement in Standard Thai using the thyroumbrometer", Phonetica 33, p. 41-67 (also UCLA WPP 33, p. 160-190)

Gedney, W.J. 1947: Indic loanwords in spoken Thai (Ph.D. diss., Yale University)

Gedney, W.J. 1972: "A puzzle in comparative Tai phonology", in: Harris & Noss 1972, p. 52-57

Gedney, W.J. 1973: "A checklist for determining tones in Tai dialects", in: Smith ed.: Studies in linguistics in honor of George Trager, p. 423-437 (The Hague: Mouton)

Gething, Thomas W., Jimmy G. Harris, and Pranee Kullavanijaya 1976: Tai linguistics in honor of Fang-Kuei Li (Chulalongkorn Univ. Press, Bangkok)

Gillette, J.A. 1955: Prosodic features in Bangkok Thai (M.A. thesis, Georgetown Univ.)

Grima, John A. 1982: "A velar for alveolar substitution in Thai child language", Studies in Language 6, p. 175-192

Gsell, R. 1979: Sur la prosodie du thai standard: tons et accent (Univ. de la Sorbonne Nouvelle, Institut d'études linguistiques et phonétiques, Paris)

Gsell, R. 1980: "Sur la réalisation chuchotée des tons du thai standard", Travaux III, Sorbonne Nouvelle, p. 217-246

Haas, Mary R. 1946: "Techniques of intensifying in Thai", Word 2, p. 127-130

Haas, Mary 1958: "The tones of four Thai dialects", Bull. of the Institute of History and Philology, Academia Sinica 29, p. 817-826

Haas, M. 1964: Thai-English student's dictionary (Stanford Univ. Press)

Haas, Mary R. and Heng R. Subhanka 1945: Spoken Thai I-II (repr. 1978) (Spoken Language Series, Inc.)

Harris, J. 1972: "Phonetic notes on some Siamese consonants", in: Harris & Noss 1972, p. 8-22

Harris, Jimmy G. 1975: "A comparative word list of three Tai Nüa dialects", in: Harris & Chamberlain 1975, p. 202-231

Harris, Jimmy G. and James R. Chamberlain 1975: Studies in Tai linguistics in honor of William J. Gedney (Office of State Universities, Bangkok)

Harris, Jimmy G. and Richard B. Noss 1972: Tai phonetics and phonology (Mahidol University, Bangkok)

Hartmann, John F. 1976: "The waxing and waning of vowel length in Tai dialects", in: Gething, Harris & Kullavanijaya 1976, p. 142-159

Hashimoto, Mantaro J. 1979: "Tones and ending consonants in monosyllabic languages", in: Thongkum et al. 1979, p. 278-285

Haudricourt, André-G. 1948: "Les phonèmes et le vocabulaire du thai commun", Journal Asiatique, p. 197-238

Haudricourt, André 1952: "Les occlusives uvulaires en thai", BSL 48.1, p. 86-89

Haudricourt, André 1956: "De la restitution des initiales dans les langues monosyllabiques: le problème du thai commun", BSL 52

Haudricourt, André 1972 (1961): "Two-way and three-way splitting of tonal systems in some Far Eastern languages", in: Harris & Noss 1972, p. 58-86 (also, in French, in BSL 56, 1961, p. 160-172)

Haudricourt, André G. 1975: "A propos du puzzle de W.J. Gedney", in: Harris & Chamberlain 1975, p. 252-258

Haudricourt, André 1979: "Le traitement des spirantes lors de la mutation des initiales", in: Thongkum et al. 1979, p. 106-110

Henderson, Eugénie J.A. 1949: "Prosodies in Siamese: study in synthesis", Asia Major, New Series 1, p. 189-215 (repr. in F.R. Palmer ed.: Prosodic analysis, 1970, p. 27-53, and in Jones & Laver eds.: Phonetics in linguistics: a book of readings, 1973, p. 127-153)

Henderson, Eugénie J.A. 1951: "The phonology of loanwords in some Southeast Asian languages", Transactions of the Philological Society, p. 131-158 (repr. in F.R. Palmer ed.: Prosodic analysis, 1970, p. 54-81)

Henderson, Eugénie J.A. 1959: "The tones of the Tai dialect of Songkhla", Bull. of the Inst. of History and Philology, Academia Sinica 30, p. 233-235

Henderson, Eugénie J.A. 1964: "Marginalia to Siamese phonetic studies", in: D. Abercrombie, D.B. Fry, P.A.D. MacCarthy, N.C. Scott, and J.L.M. Trim eds.: In Honor of Daniel Jones, p. 415-424

Henderson, Eugénie J.A. 1975: "Phonetic description and phonological function: some reflections upon back unrounded vowels in Thai, Khmer and Vietnamese", in: Harris & Chamberlain 1975, p. 259-270

Henderson, Eugénie J.A. 1976: "Thai phonetics sixty years ago: gleanings from the unpublished notes of Daniel Jones", in: Gething, Harris & Kullavanijaya 1976, p. 162-170

Henderson, E.J.A. 1982: "Tonogenesis: some recent speculations on the development of tone", Transactions of the Philological Society, p. 1-24 (London)

Hiranburana, Samang 1971: The role of accent in Thai grammar (Doct. diss., Univ. of London)

Hiranburana, Samang 1972: "Changes in the pitch contours of unaccented syllables in spoken Thai", in: Harris & Noss 1972, p. 23-27

Jones, Robert B. 1965a: "Phonological features of Northern Thai", in: J.R. Hanks, L.H. Hanks, and L. Sharp eds.: Ethnographic notes on Northern Thailand (Cornell Thailand Proj. Series, No. 9, Cornell Univ.), p. 10-14

Jones, R.B. 1965b: "On the reconstruction of Proto-Thai", Lingua 14, p. 194-229

Jones, Robert B. 1976: "Tone shift in Tai dialects", in: Gething, Harris & Kullavanijaya 1976, p. 171-178

Karnchanathat, N. 1977: An instrumental study of Thai speakers' English intonation (Ph.D. diss., Univ. of Michigan)

Khanittanan, Wilaiwan W. 1973: The influence of Siamese on five Lao dialects (Ph.D. diss., Univ. of Michigan)

Khanittanan, Wilaiwan W. 1975: "phaasǎa thîn "phûuthaj" nakhɔɔn phanom" [The Phuthai dialect of Nakhorn Phanom] in: Harris & Chamberlain 1975, p. 377-386

Kloster-Jensen, M. 1965 (review): "The vowels and tones of Standard Thai ... af Arthur S. Abramson", Word 21.1, p. 163-166

Kroll, M.E. 1956: Suprasegmental phonemes of Thai (Bangkok dialect) (M.A. thesis, Georgetown Univ.)

Kruatrachue, Foongfuang 1960: Thai and English: a comparative study of phonology for pedagogical applications (Ed.D. thesis, Indiana Univ. Press)

Lea, Wayne A. 1973: "Segmental and suprasegmental influences on fundamental frequency contours", in: L. Hyman ed.: Consonant Types and Tone, p. 15-70 (Univ. of Southern California)

Leben, W.R. 1971: "On the segmental nature of tone in Thai", MIT Res. Lab. of Electronics Quarterly Progr. Rep. 101

Lee, C.K. 1980: "La nasalisation vocalique en thai", Travaux III, Sorbonne Nouvelle, p. 121-136

Lewis, Paul 1972: "Phonetic problems involved in teaching Thai to speakers of Lahu and Akha", in: Harris & Noss 1972, p. 105-112

Li, Fang-Kuei 1943: "The hypothesis of a pre-glottalized series of consonants in primitive Tai", Bull. of the Inst. of History and Philology, Academia Sinica 11, p. 177-188

Li, Fang-Kuei 1954: "Consonant clusters in Tai", Language 30, p. 368-379

Li, Fang-Kuei 1960: "A tentative classification of Tai dialects", in: S. Diamond ed.: Culture in history, Essays in honor of Paul Radin, p. 951-959 (New York)

Li, Fang-Kuei 1962: "Initial and tone development in the Tai dialects", Bull. of the Inst. of History and Philology, Academia Sinica 34, p. 31-36

Li, Fang-Kuei 1966: "The relationship between tones and initials in Tai" in: N. Zide ed.: Studies in Comparative Austroasiatic Linguistics, p. 82-88 (The Hague: Mouton)

Li, Fang-Kuei 1968: "Tai dialectology", Zeitschrift für Mundartforschung, Beiheft, Neue Folge 4,II, p. 534-537

Li, Fang-Kuei 1970: "Some tonal irregularities in the Tai languages", in: R. Jakobson & Sh. Kawamoto eds.: Studies in General and Oriental Linguistics, p. 415-422

Li, Fang-Kuei 1973: "Some dental clusters in Tai", BSOAS 36.2, p. 334-339

Li, Fang-Kuei 1975: "Some words for Thai dialectology", in: Harris & Chamberlain 1975, p. 271-273

Li, Fang-Kuei 1977: A handbook of comparative Tai (Univ. Press of Hawaii)

Luksaneeyanawin, Sudaporn 1976: "Some phonetic observations on the speech of Thai children", in: Brudhiprabha et al. 1976, p. 120-125

Luksaneeyanawin, Sudaphorn 1983: Intonation in Thai (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 28-29

Luksaneeyanawin, Sudaporn 1984: "The tonal behaviour of one-word utterances: the interplay between tone and intonation in Thai", Work in Progress (Dept. of Ling., Univ. of Edinburgh) 17, p. 16-30

Maspero, Henri 1911: "Contribution à l'étude de la phonétique du système phonétique des langues thai", Bull. de l'Ecole Française d'Extrême-Orient XI, p. 153-169 (Hanoi)

Matisoff, J. 1973: "Tonogenesis in Southeast Asia", in: L. Hyman ed.: Consonant Types and Tone, p. 71-95 (Univ. of Southern California)

Miller, Roy A. 1956: "Phonology of the Thai dialect of Nakhonsitthammarat", Orbis 5, p. 250-258

Minnigerode, B. and A. Vichapun 1969: "Informationstheoretische Untersuchungen an der thailändischen Sprache", Phonetica 20, p. 41-56

Nacaskul, Karnchana 1979: "A note on English loanwords in Thai", in: Thongkum et al. 1979, p. 151-162

Nacaskul, Karnchana 1981: rabòp sǐəŋ phaasǎa thaj [The sound system of Thai] (Bangkok)

Noss, Richard B. 1964: Thai reference grammar (Washington D.C. Foreign Service Institute)

Noss, Richard B. 1972: "Rhythm in Thai", in: Harris & Noss 1972, p. 33-42

Noss, Richard B. 1975: "How useful are citations forms in synchronic Thai phonology?", in: Harris & Chamberlain 1975, p. 274-284

Ongrabieb, Preou-Payom 1974-75: Etude comparative des systèmes phonologiques et phonétiques du français et du thai (Thèse de 3e cycle, Univ. de la Sorbonne Nouvelle)

Painter, Colin 1978: "Implosives, inherent pitch, tonogenesis, and laryngeal mechanisms", Journal of Phonetics 6, p. 249-274

Palmer, A. 1969: "Thai tone variants and the language teacher", Language Learning 19, p. 287-299

Panka, Kanchana 1980: The phonological characteristics of Lao dialects in Amphoe Muang, Nakhon Pathom (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 12

Panupong, Vichin (Chantavibulya) 1956: Differences between the Bangkok and the Songkhla dialects (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 33

Panupong, Vichin (Chantavibulya) 1959: The phonology of the syllable of Songkhla, a Southern Thai dialect (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 35

Panupong, Vichin 1972: "laksanáʔ kaan'ɔ̀ɔk sǐəŋ khɔ̌ɔŋ khamphûut naj phaasǎa sɔ̌ŋkhlǎa" [Some aspects of Songkhla pronunciation], in: Harris & Noss 1972, p. 87-104

Panupong, V. 1976: "khwaamtɛ̀ɛktàaŋ rawàaŋ phaasǎa thîn" [Differences among dialects], in: Gething et al. 1976, p. 242-254

Panupong, V. 1983: "Word geography in Nakhon Ratchasima: a pilot project", Proceed. of the XIIIth Int. Congr. of Ling., p. 986-989 (Tokyo)

Piyatham, S. 1970: A feature approach to the phonology of Phuket, a Southern Thai dialect (M.A. thesis, London)

Purnell, Herbert C., Jr. 1972: "Toward contrastive analysis between Thai and hill tribe languages: some phonetic data", in: Harris & Noss 1972, p. 113-130

Richards, Jack 1966: Pronunciation features of Thai speakers of English (thesis, Victoria Univ. of Wellington)

Rinprom, Chalida 1976: "The phonemic system of the Khorat dialect" (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 4

Rischel, Jørgen and Amon Thavisak 1984: "A note on work in progress: secondary articulation in Thai stops", Ann. Rep. Inst. Phon., Univ. Copenhagen 18, p. 243-254

Rudaravanija, Panninee 1965: An analysis of the elements in Thai that correspond to the basic intonation (Ed.D. diss. Columbia Univ.)

Sagarik, P., see: Rudaravanija

Sarawit, M. 1973: The Proto-Tai vowel system (Ph.D. diss., Univ. of Michigan, repr. Univ. Microfilms, Ann Arbor)

Sarawit, Mary 1975; "Some changes in the final component of the Tai syllable", in: Harris & Chamberlain 1975, p. 316-328

Sarawit, Mary 1976: "Assimilation in child language acquisition", in: Gething, Harris & Kullavanijaya 1976, p. 205-213

Scovel, Thomas 1975: "The role of aphasia research in Thai linguistics", in: Harris & Chamberlain 1975, p. 329-341

Senisrisant, Rachanee 1982: A comparative study of some phonetic and phonological aspects in the speech of Maapplakhao Laophuan speakers of different age groups (thesis), see Abstracts ... Chulalongkorn Univ. 1983 p. 25-26

Sittachit, Kanda 1972: "khɔ̂ɔ sǎŋkèet kìəw kàp sarà' sân sarà' jaaw naj phaasǎa thaj" [Observations on short and long vowels in Thai], in: Harris & Noss 1972, p. 28-32

Somnuk, Jariya 1981: "The phonology of Nakhon Sithammarat sub-dialects" (thesis), see Abstracts ... Chulalongkorn Univ. 1983, p. 6-7

Strecker, David 1979a: "A preliminary typology of tone shapes and tonal changes in Tai: the Lān Nā A-tones", in: Thongkum et al. 1979, p. 171-240

Strecker, David 1979b: "Higher falls more: A tonal sound change in Thai", Computational Analyses of Asian and African Languages 11, p. 30-84

Suntharagul, B. 1962: "nùaj sǐəŋ khɔ̌ɔŋ phaasǎa chiaŋmàj [The phonemes of Chiang Mai] (thesis, Chulalongkorn Univ.)

Surintramont, A. 1973: "Some aspects of underlying syllable structure in Thai: evidence from KHAMPUAN - a Thai word game", Studies in the Linguistic Sciences 3.1, p. 121-142

Suwattee, Duangduen and Pranee Kullavanijaya 1976: "A study of อย, ย, หญ and ญ", in: Gething, Harris & Kullavanijaya 1976, p. 214-224

Tanwattananun, Penporn 1982: The change of unaspirated stops in Chieng Mai dialect and some social variables (thesis), see: Abstracts ... Chulalongkorn Univ. 1983, p. 19

Thampradit, Jaroenquan 1980: The use of the phonetic characteristics of high vowels for a dialect survey of Trang, Krabi, Phangnga, and Phuket (thesis), see: Abstracts ... Chulalongkorn Univ. 1983, p. 8-9

Thaweesomboon, Sanit 1955: Syllable juncture within stress groups in spoken Thai (M.A. thesis, Univ. of London)

Thaweesomboon, Sanit 1969: [An analysis of Thai students' problems in pronouncing the vowels and consonants of American English (Bangkok), in Thai]

Thongkum, Theraphan L. 1976a: "Rhythm in Thai from another viewpoint", in: Brudhiprabha et al. 1976, p. 144-159 (for Th. L. Thongkum 1976b, see Tongkam, Th.L.)

Thongkum, Theraphan Luang 1977: Rhythm in Standard Thai (thesis), see Abstracts ... Chulalongkorn Univ. 1983 p. 30-31

Thongkum, Theraphan L. 1979: "Iconicity of vowel qualities in Northeastern Thai: reduplicated words", in: Thongkum et al. 1979, p. 247-260

Thongkum, Theraphan Lüang 1984: "Rhythmic groups and stress groups in Thai", in: Preprints of the Int. Conference on Thai Studies, Bangkok, 22-24 August 1984, volume Language and Literature

Thongkum, Theraphan L., Pranee Kullavanijaya, Vichin Panupong, and M.R. Kalaya Tingsabadh 1979: Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie J. A. Henderson (Chulalongkorn Univ. Press, Bangkok)

Tingsabadh, M.R. Kalaya 1980: "A phonological survey of the Thai language of Suphanburi province" (thesis), see: Abstracts ... Chulalongkorn Univ. 1983, p. 11

Tingsabadh, Kalaya 1984: "Thai dialectology up to the year 1984", in: Preprints of the Int. Conference on Thai Studies, Bangkok, 22-24 August 1984, volume Language and Literature

Tongkam, Theraphan L. 1976b: "Relative durations of syllables in Siamese connected speech", in: Gething, Harris & Kullavanijaya 1976, p. 225-232

Tuaycharoen, Pintip 1977: The phonetic and phonological development of a Thai baby: from early communicative interaction to speech (Ph.D. thesis, Univ. of London)

Tuaycharoen, Pintip 1979: "An account of speech development of a Thai child: from babbling to speech", in: Thongkum et al. 1979, p. 261-277

Van Lancker, Diana and Victoria A. Fromkin 1973: "Hemispheric specialization for pitch and "tone": evidence from Thai", Journal of Phonetics 1, p. 101-109

Velder, Christian 1962: "Tonsilbenanalyse des Thai", Zeitschrift für Phonetik und allg. Sprachwissenschaft 15, p. 177-206

Warotamasikkhadit, Udom 1967: "Some phonological rules in Thai", JAOS 87.4, p. 541-574

Warotamasikkhadit, Udom 1968: "Stresses and tone rules in Thai", Papers of the CIC Far Eastern Language Inst., p. 169-175 (Ann Arbor, Mich.)

Whitaker, H.A. 1969: "Stylistic tone-changing rules in Thai", Glossa 3.2, p. 190-197 (also: UCLA WPP 10, 1968, p. 191-199

Note: For earlier work on Tai linguistics there is the standard source:

Shorto, H.L., J.M. Jacob & E.H.S. Simmonds 1963: Bibliographies of Mon-Khmer and Tai linguistics (Oxford Univ. Press)

As for the special subject of aphasia studies dealing with Thai language, see the following papers with ample references:

Gandour, Jack and Rochana Dardarananda 1984: "Voice onset time in aphasia: Thai. II. Production" and "Prosodic disturbance in aphasia: vowel length in Thai", Brain and Language 23, p. 177-205 and 206-224, respectively.

As for recent work within Thai dialectology (primarily by Thai scholars and students) important additional references can be found in Kalaya Tingsabadh (1984).

ARIPUC 19, 1985, p. 95-141

FUNDAMENTAL FREQUENCY AND LARYNX HEIGHT IN SENTENCES AND STRESS GROUPS

NIELS REINHOLT PETERSEN

The paper reports experiments investigating a) the relation between larynx height and the slow overall fundamental frequency movement (declination) in sentences of varying length and type, and b) the relation between larynx height and Fo in the prosodic stress group. The aim of the experiments was to see in general to what extent the extrinsic laryngeal muscles could be assumed to be involved in the production of Fo, and as part of this to consider more specifically the question whether the fundamental frequency declination is actively controlled by the speaker in a linguistically purposeful manner, or whether - as has been suggested - it is an automatic consequence of the functioning of the pulmonic system. The results can be summarized as follows: In declarative sentences an overall decline was observed for Fo as well as for larynx height, and for both the slope of declination varied with sentence length, being steeper in short than in long sentences. In interrogative sentences Fo as well as larynx height showed either no decline or a slight rise over the sentence. These results are taken to indicate that Fo declination is actively controlled by the speaker, and it is suggested that it may be attributed primarily to the activity of the extrinsic laryngeal muscles. In the stress group the relation between larynx height and Fo seems to be far more complex, showing patterns which vary both among speakers and (to a lesser degree) within the individual speaker, and which have to be explained as the result of an intricate interaction of the activity of both intrinsic and extrinsic laryngeal muscles.

I. INTRODUCTION

One way of describing the fundamental frequency (Fo) variation in speech is to regard it as the composite result of the superposition of several components related to units of varying temporal scope, ranging from the text, through the utterance, sentence and phrase to the prosodic stress group. In addition to these components a microprosodic component will have to be specified, taking care of the segmentally conditioned Fo variation, such as inherent Fo levels of vowels and (voiced) consonants and effects exerted by consonants on the Fo of adjacent vowels.

The present paper will examine the association between fundamental frequency and larynx height in a Danish material as seen in relation to two of the above mentioned components, namely the sentence component and the stress group component. The basis for taking larynx height into consideration is the existence of a statistical relationship between LH and Fo (see e.g. Shipp 1975, Ohala 1973 and 1978, Ewan 1979) and the ability of the extrinsic laryngeal muscles to affect the vertical position of the larynx in the neck.

Apart from segmentally conditioned effects, the fundamental frequency variation associated with units of short temporal scope, such as the stress group, is generally believed to be caused by the activity of the intrinsic laryngeal muscles, particularly the cricothyroid muscle (see e.g. Collier 1975, Atkinson 1978, Honda 1983, Gelfer, Harris, Collier, and Baer 1983, and Ohala 1978, who gives an extensive review of the literature). But there is also evidence, although less consistent, perhaps, that the extrinsic laryngeal muscles take part in the control of local Fo movements. (See further Ohala 1970, Sawashima, Kakita, and Hiki 1973, Collier 1975, Erickson and Atkinson 1976, Atkinson and Erickson 1977, Erickson, Liberman, and Niimi 1977, Atkinson 1978, Honda 1983.) In an earlier paper (Reinholt Petersen 1984), I presented data showing that for two of the three subjects examined, a larynx rise accompanied the Fo rise from stressed to first posttonic syllable which is found in Standard Copenhagen Danish (see Thorsen 1979). The third subject had the opposite pattern, his Fo rise being accompanied by a larynx lowering. I shall not here enter into a detailed interpretation of these results, since they were essentially similar to those of the present study, which will be discussed below. Suffice it here to say that it was assumed that the extrinsic laryngeal muscles had been involved in the control of local Fo movements, and that the speakers had used two different strategies producing identical Fo movements.

In units of greater temporal scope, such as the sentence, there seems to be a universal tendency for the fundamental frequency in declaratives to fall gradually as a function of time. One of the central issues concerning this phenomenon, which has been termed Fo declination or downdrift, is whether it is to be explained as an automatic, i.e. not actively controlled, consequence of constraints in the speech production apparatus, or whether it is under active speaker control and part of the linguistic system.

The explanations adhering to the former view are based upon the fact that the lung volume gradually decreases as a function of time during speech. One such hypothesis takes its point of departure in Lieberman's (1967) theory which claims subglottal pressure to be the major source of Fo variation. The theory in its general form is hardly tenable (see e.g. Ohala 1978), but Collier 1975, and Gelfer, Harris, Collier, and Baer 1983, who investigated the EMG activity of various laryngeal muscles together with subglottal pressure, present data suggesting that in their material Fo declination could be accounted for by the gradually declining subglottal pressure which they observed, and which was assumed to be related to decreasing lung volume (see also Cohen, Collier, and 't Hart 1982).

Another explanation referring to the pulmonic system has been advanced by Maeda (1974). On the basis of the observation that the vertical position of the larynx - like Fo - shows an overall decline over the utterance, it is hypothesized that the decreasing lung volume causes a continuous lowering of the sternum, which via ligaments and muscular tissue pulls the larynx downwards with a fundamental frequency decline as the result. Other researchers, however, have not found the larynx lowering reported by Maeda; on the contrary, Gandour and Maddieson (1976) and Ewan (1979) report that - if anything - the larynx rises during the utterance.

Maeda (1979) gives a modified version of the hypothesis. In the material presented he also fails to find the Fo decline to be consistently accompanied by a larynx height decline, but he finds that the laryngeal ventricle shows an overall tendency to shorten during the utterance. Under the assumption that ventricle length can be used as an estimator of vocal fold length, the following mechanism is hypothesized: The decreasing lung volume exerts a pull upon the trachea. This pull is transferred to the cricoid cartilage and - if the thyroid cartilage is fixed - is thought to tilt the cricoid cartilage in relation to the thyroid cartilage, whereby the vocal folds are shortened and, consequently, Fo lowered.

The mere fact that fundamental frequency declination is so widespread in terminal declaratives among the languages of the world could be taken to speak in favour of the 'automatic' hypotheses outlined above. However, observation of Fo declination under various linguistic conditions provide strong evidence in support of the view that declination is under active speaker control as part of the linguistic system: Fo declination can convey information about sentence type and function as is the case in Danish (Thorsen 1979 and 1980a), where terminal declarative sentences have the steepest declination, syntactically and lexically unmarked questions have no declination, and non-final periods and interrogative sentences with word-order inversion and/or interrogative particle have intermediate degrees of declination. Furthermore, the slope of declination seems to vary with sentence length, being steeper in short than in long sentences (see e.g. Maeda 1974, Sorensen and Cooper 1980, Thorsen 1980b, 1981, and 1983). And as a final example, resetting (partial or complete) of the declination line has been shown to take place without intervening inhalation (Sorensen and Cooper, 1980, Thorsen 1980b, 1981, and 1983). Data of this kind are not easily accounted for by explanations basing themselves upon physiological constraints in the pulmonic system (or in any other part of the speech production apparatus for that matter).

In the study referred to above on Fo and larynx height (Reinholt Petersen 1984), results were reported indicating that Fo declination is somehow related to larynx height, the larynx as well as the fundamental frequency showing a gradual decline as a function of time through the sentence. But since sentence length and type were not varied (only declarative sentences of three stress groups were examined), these findings cannot be taken to support either of the views on fundamental frequency declination outlined above.

Therefore, the purpose of the experiments reported below was to investigate the relationship between Fo declination and larynx height variation in sentences of varying length and type in an attempt to gain information relevant to the question whether Fo declination is actively controlled or not, and to the discussion of the physiological mechanisms underlying it. Another purpose was to examine the Fo and larynx height variation attributable to the stress group component in a greater number of speakers, with the particular aim of looking for different strategies and, further, to expand the scope of investigation to include not only the relation between stressed and first posttonic syllables (as was done in Reinholt Petersen 1984) but the entire stress group, i.e. up to the next stressed syllable.

II. METHOD

A. MATERIAL

Two types of speech material were employed in the investigation. One (henceforth referred to as material A) was primarily designed with the purpose of examining overall Fo and LH variation in sentences of varying length and type (declarative versus interrogative). It consisted of the nonsense word ['fi:fi] which was inserted in carrier sentences containing from two to five stress groups as follows:

2.1 gi ['fi:fi] forklaring
2.2 forløbet i ['fi:fi]

3.1 i ['fi:fi] forkortes vokalen
3.2 vokalen i ['fi:fi] forkortes
3.3 vokalen forkortes i ['fi:fi]

4.1 i ['fi:fi] forkortes vokalen i starten
4.2 vokalen i ['fi:fi] forkortes i starten
4.3 vokalen forkortes i ['fi:fi] i starten
4.4 vokalen forkortes i starten i ['fi:fi]

5.1 i ['fi:fi] forkortes vokalen for meget i starten
5.2 vokalen i ['fi:fi] forkortes for meget i starten
5.3 vokalen i starten i ['fi:fi] forkortes for meget
5.4 vokalen forkortes i starten i ['fi:fi] for meget
5.5 vokalen i starten forkortes for meget i ['fi:fi]

The sentences have numbers indicating the number of stress groups in a sentence and the the position of the test word, e.g. 5.3 is the 5-stress-group-sentence where the test word occurs in the 3rd stress group. All sentences have an initial unstressed syllable followed by the appropriate number of stress groups, each consisting of a stressed syllable with a long vowel plus two unstressed syllables, except for the last stress group, which contains only one unstressed syllable. This design of the material made it possible, for each sentence length, to treat the measurements as if they had been extracted from only one sentence with that length. (It would have been asking too much of the speakers to have them read e.g. one five-stress-group sentence with five repetitions of ['fi:fi] in it.) All sentences were to be rendered with a neutral declarative intonation, but in order to obtain information about the effect of sentence type on Fo and larynx height the three-stress-group-sentences were also to be spoken on an interrogative intonation. Material A could also, of course, be used for looking into the relation between Fo and larynx height in stressed versus unstressed (lst posttonic) syllables. The material was spoken by three subjects.

The second material (henceforth material B), which was to be spoken by a larger group of subjects, was made up by the following sentences:

2.1 gi ['fi:fi] forklaring
2.2 forløbet i ['fi:fi]

5.1 i ['fi:fi] forkortes vokalen for meget i starten
5.5 vokalen forkortes for meget i starten i ['fi:fi]

4.1/4 { ['fi:fifi] er kortere i starten end ['fi:fi]
['fi:fifi] er kortere i starten end ['fi:fi]

4.2 vokalen i ['fi:fifi 'fi:fəs] i starten

The numbering of the sentences follows the same rules as in material A (note that the two identical 4-stress-group-sentences are both called 4.1/4, meaning that test words occur in the first and in the fourth stress group).

Material B was designed with three objects in mind: 1) As a supplement to material A, it was to give information about Fo declination and larynx height in sentences of varying length.

This requirement was met by the sentences 2.1, 2.2, 5.1, and 5.5, which were identical to the shortest and the longest sentences of material A, and also in essentials by the two identical sentences 4.1/4, although they have a slightly different rhythmical structure compared to the 2- and 5-stress-group-sentences.
2) It was intended to make it possible to observe Fo and larynx movements over the entire stress group. This could be done in the nonsense word portions of sentence 4.2, and in addition stressed and first posttonic syllables could be observed in all ['fi:fi] words in the material and second posttonics in the ['fi:fifi] words in the type 4.1/4 sentences.
3) The third object of the material was to gain information about larynx height in pauses in order to see whether it would differ from the larynx height observed during speech. Since it was considered desirable to have identical sentences adjacent to the pauses in which larynx height was to be observed, the two type 4.1/4 sentences, which always occurred in succession during recording, were included in the material.

The two materials were arranged in eight random orders in two separate reading lists.

B. RECORDING

The recording equipment consisted of a television camera (Sony AVC-3250 CES) and a video-recorder (Sony U-Matic type 2630). The frame frequency of the equipment was the normal 50 frames per second. The speech signal was recorded on the sound track of the video-tape via a Sennheiser MD 21 microphone placed about 15 cm from the subject's mouth. In order to synchronize speech and video signals a timer signal was recorded on the video-tape using a timing device (FOR-A CO. type VTG 33). On playing back the tape, the timer signal was displayed on the monitor screen in minutes, seconds, and centiseconds and it could, moreover, be registered together with the speech signal on an ink writer as pulses for seconds and centiseconds. In this manner it was possible to relate each TV-frame to the speech signal.

During recording the subject was seated in a dentist's chair with a fixed head-rest. The camera was placed at the level of the subject's thyroid prominence and at right angles to his mid-sagittal plane at a distance which allowed the area between the subject's chin and sternum to be covered by the field of vision. The subject was wearing a light but firmly fitting headgear to which a measurement scale was attached in such a manner that it formed the background of the front of the neck and the laryngeal prominence on the TV picture. The scale was divided into units which corresponded to millimetre units at the subject's midsagittal plane. The two materials were recorded at separate recording sessions.

C. SUBJECTS

Recordings were made of altogether nine male speakers. Three of them (PD, PM, and NR, the author) read both materials. The other six speakers (CH, JB, JR, ND, OB, and PH) read material B only. All subjects were speakers of Standard Danish, but with varying degrees of dialectal and/or regional influence.

D. REGISTRATION AND MEASUREMENTS

The following acoustic curves were made: duplex oscillogram, two intensity curves, and an Fo curve. The timer signal was also registered on the mingograph. Fo and larynx height were measured at the midpoint of all vowels in the test words (except for the vowel [ə] in ['fi:fəs] in material B, sentence 4.2). Together with Fo and larynx height the distance in time from the beginning of the sentence was also recorded for all points of measurement. The beginning of the sentence was defined as the onset of the first vowel in the sentence. In the pause between the two identical sentences (4.1/4) in material B larynx height was measured at points 4, 6, and 10 cs apart, depending on the distance from the sentences, so that measurements were more closely spaced near the sentences. Furthermore, for the purpose of comparison the larynx height (but not Fo) was measured in all vowels in the two sentences. The locations on the video-tape of the frames in which larynx height was to be measured were determined from the acoustic curves and the timer curve. Since the interval between frames was 2 cs (the frame frequency being 50 Hz), the temporal inaccuracy of a frame in relation to the corresponding point of measurement as determined from the acoustic curves was ± 1 cs. The video recorder was equipped with a step function which made it possible during play-back to "freeze" the picture and step forward frame by frame and read off larynx height in the frames selected for measurement. The vertical position of the larynx could be measured with an accuracy of ± 0.5 mm.

III. RESULTS

Mean values for fundamental frequency and larynx height in material A and B are given in appendix I together with standard deviations and number of observations. Note that in material B the measurements for the two identical sentences (4.1/4) have been pooled, so that the maximum number of observations is 16 instead of 8. The means are displayed graphically as a function of time (i.e. the distance from the onset of the first vowel in the sentence) in figs. 1 to 3 (material A) and figs. 4 to 10 (material B).

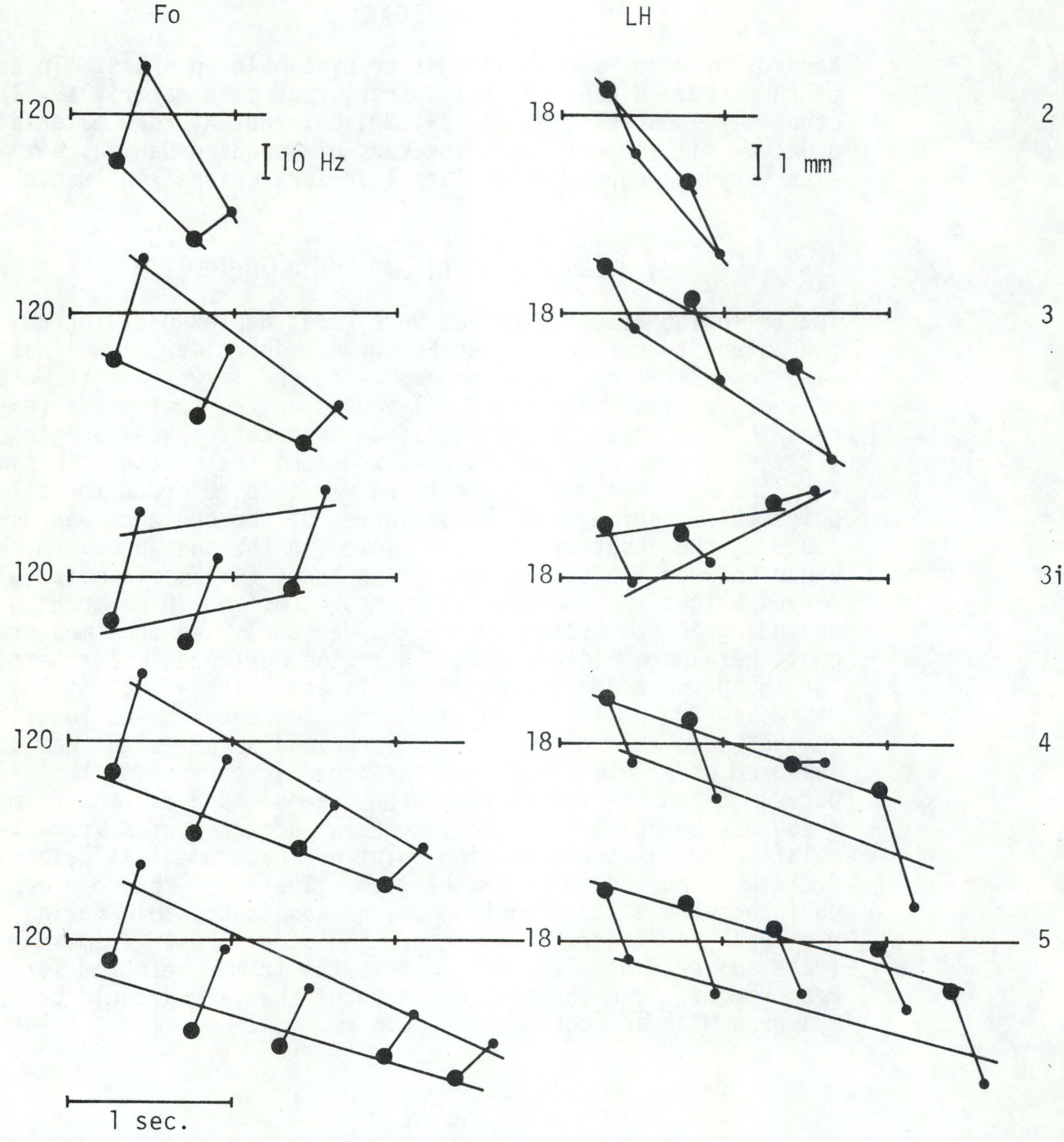


Figure 1

Material A, subject NR. Fundamental frequency (Fo) and larynx height (LH) as a function of time in sentences consisting of 2 through 5 stress groups (3i indicates the 3 stress group sentence spoken in an interrogative intonation). The level of the horizontal reference line is given in Hz and mm, respectively, using an arbitrary zero for larynx height. Large dots indicate stressed syllables, small dots unstressed ones. The lines through the data points are regression lines fitted to stressed and unstressed syllables separately using the method of least squares.

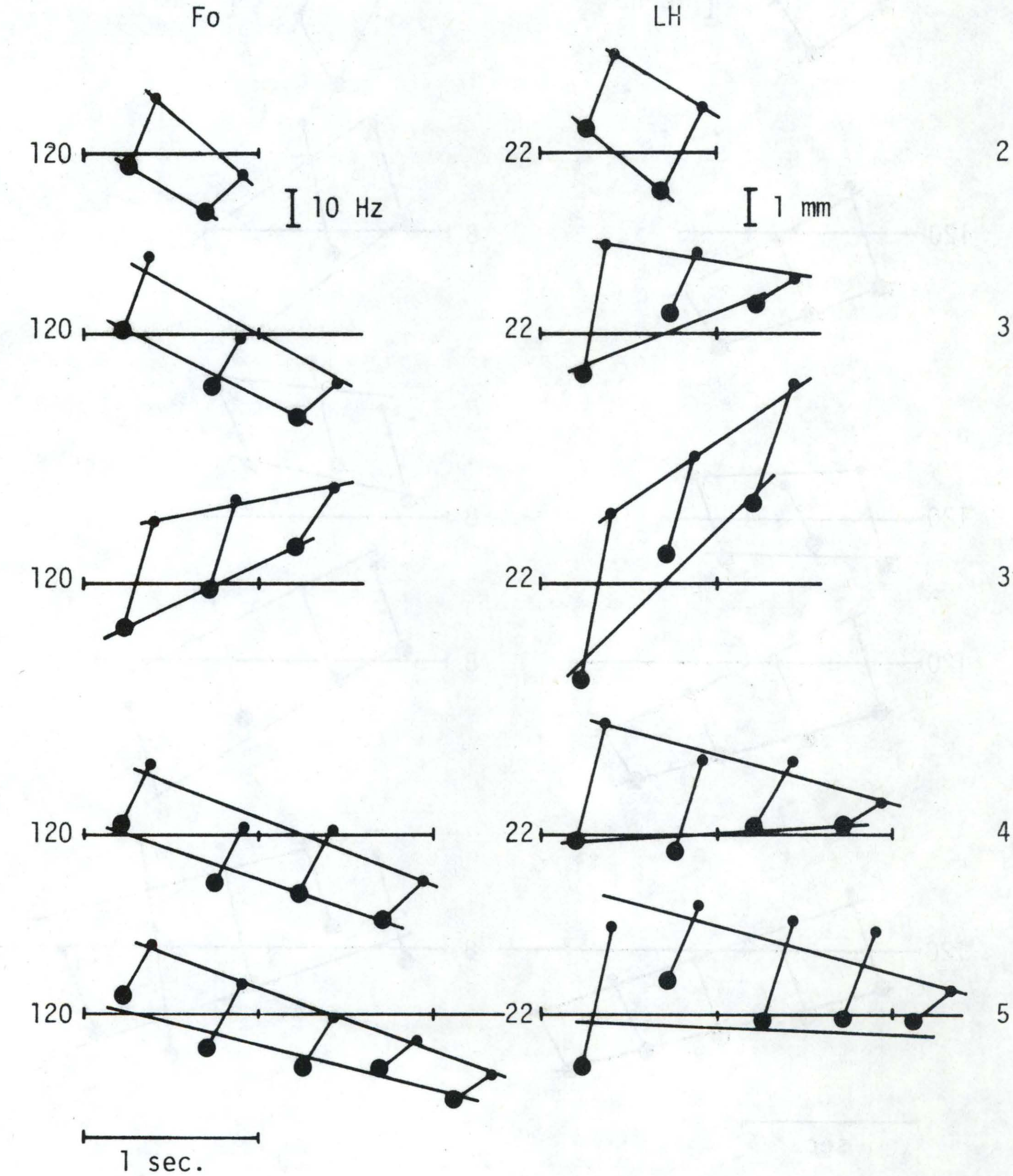


Figure 2

Material A, subject PD. Fo and larynx height as a function of time. See further legend to figure 1.

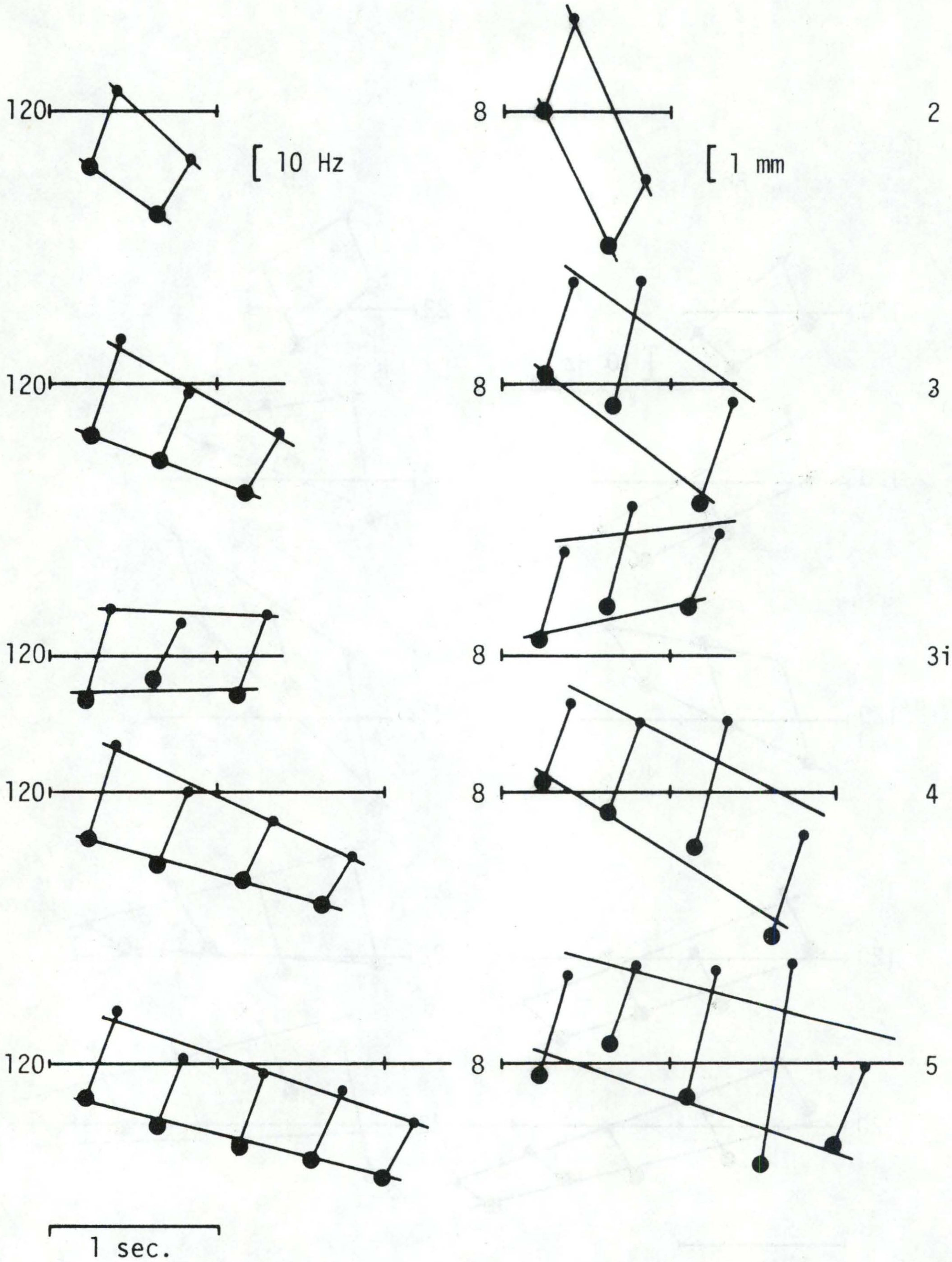


Figure 3

Material A, subject PM. Fo and larynx height as a function of time. See further legend to figure 1.

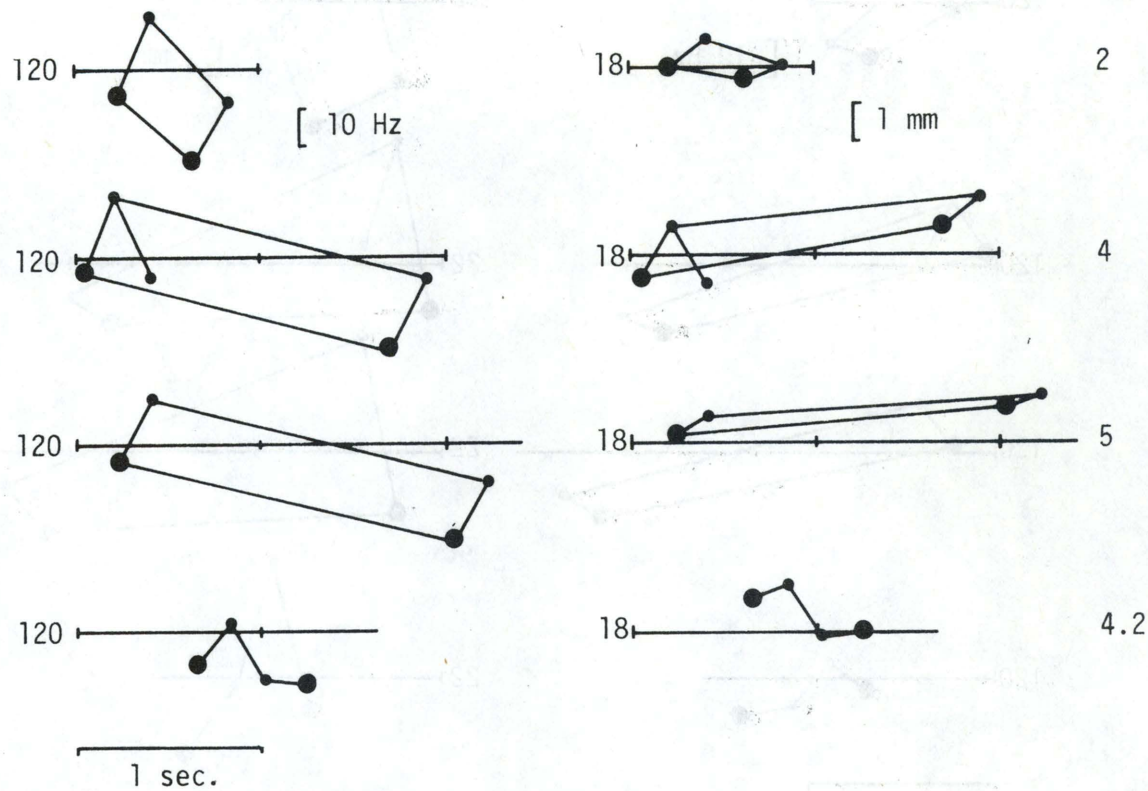


Figure 4

Material B, subject CH. Fo and larynx height in first and last stress groups in sentences of 2, 4, and 5 stress group sentences (upper 3 rows) and in the 2nd stress group in a 4 stress group sentence. The level of the horizontal reference line is given in Hz and mm, respectively, using an arbitrary zero for larynx height. Large dots indicate stressed and small dots unstressed syllables.

A. FUNDAMENTAL FREQUENCY DECLINATION AND LARYNX HEIGHT

1. MATERIAL A

In order to obtain a quantitative estimate of the overall Fo and larynx height variation over the sentence, regression lines and correlation coefficients of these variables versus time were computed. The computations were made on the basis of the raw data, and stressed and first posttonic syllables were treated separately. The regression lines are inserted in the data plots in figures 1 to 3, and slopes (in Hz and mm per second) and correlation coefficients are given in table I.

It is seen that in all declarative sentences there is a clear fundamental frequency decline over the sentence. The correlation of Fo versus time is significantly negative in all cases ($p<0.01$). In interrogative sentences, on the other hand, Fo shows either no decline (subject PM), or it has a slightly rising slope (NR and PD). It is also apparent from table I and

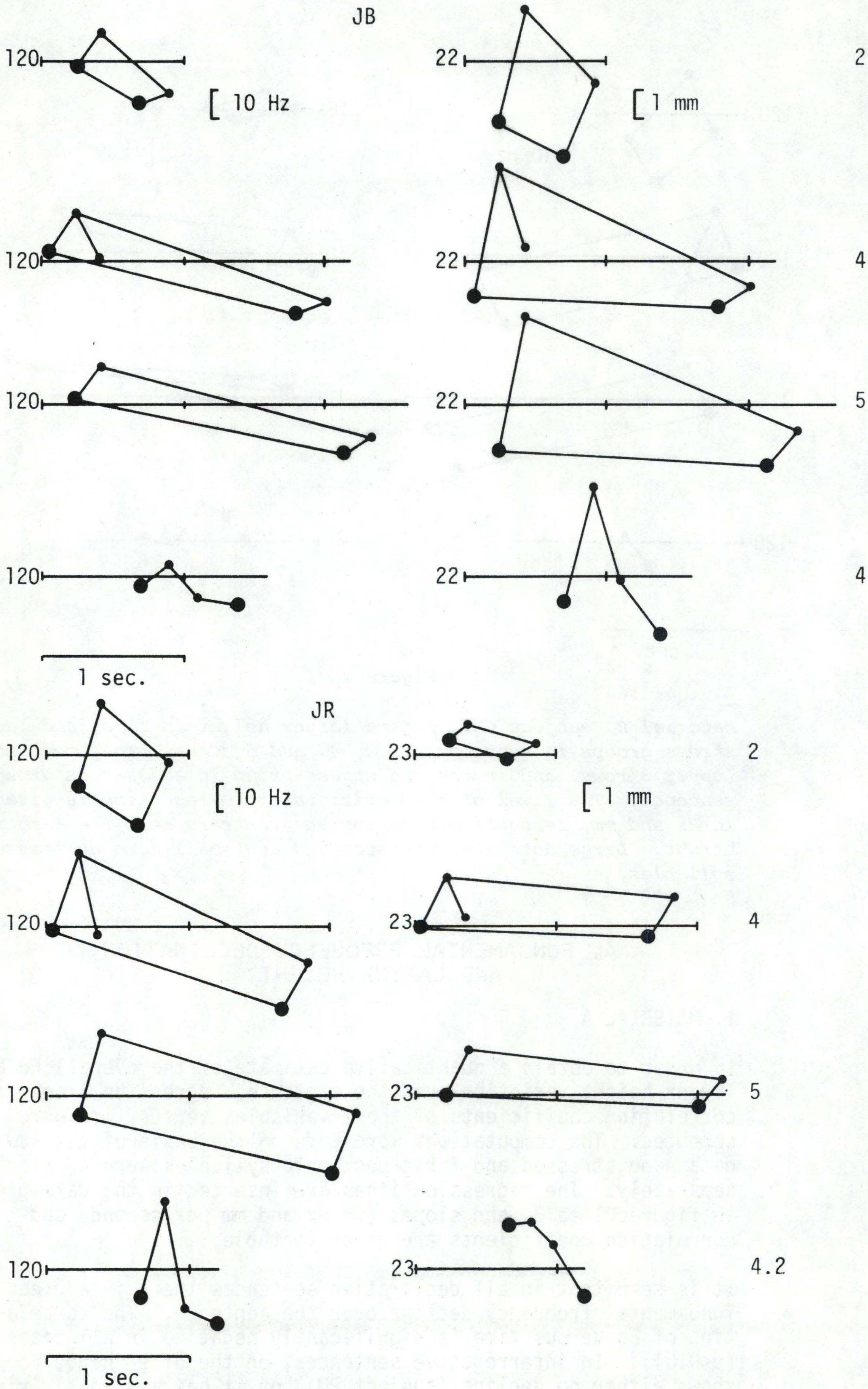


Figure 5

Material B, subjects JB and JR. See further legend to figure 4.

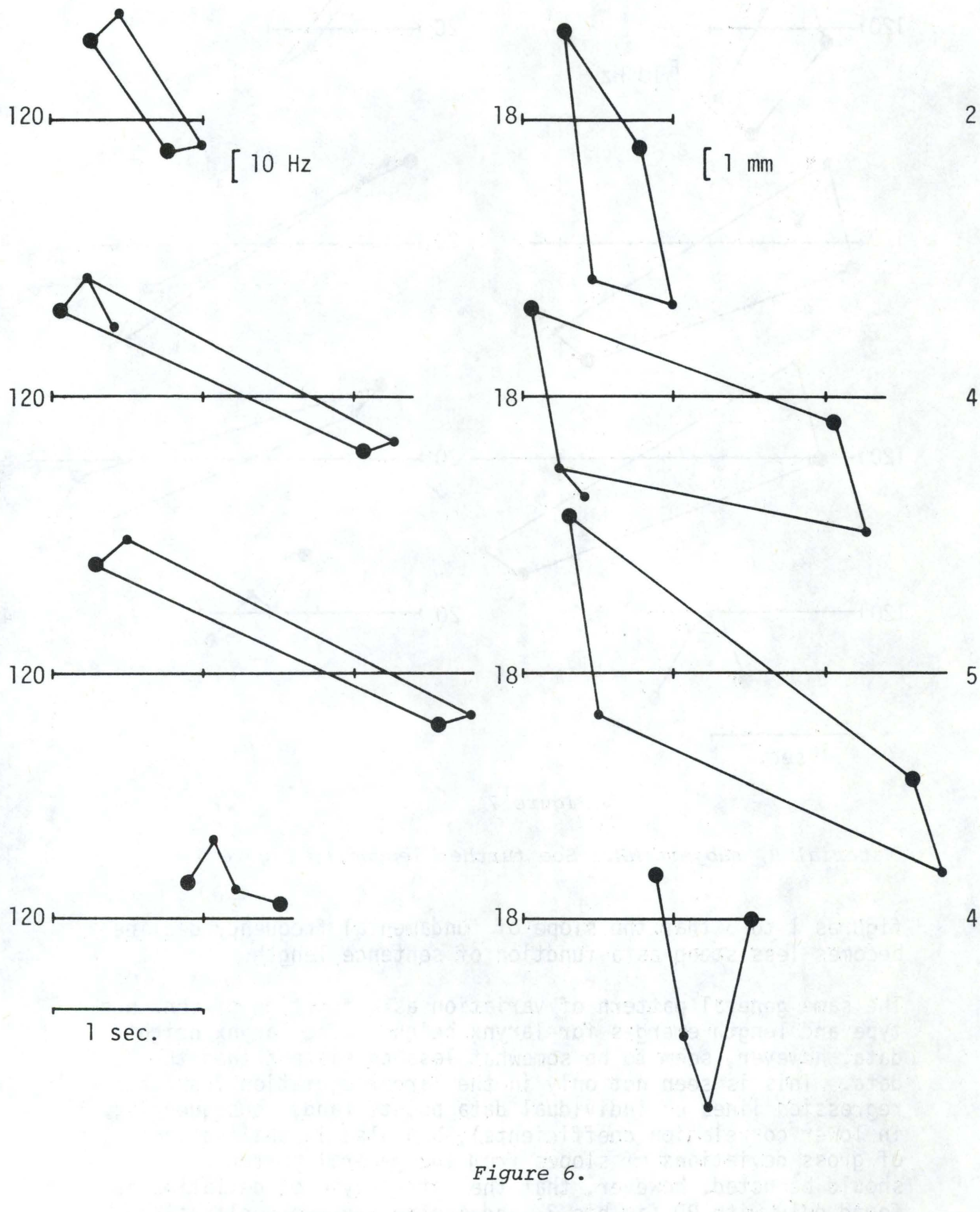


Figure 6.

Material B, subject ND. See further legend to figure 4.

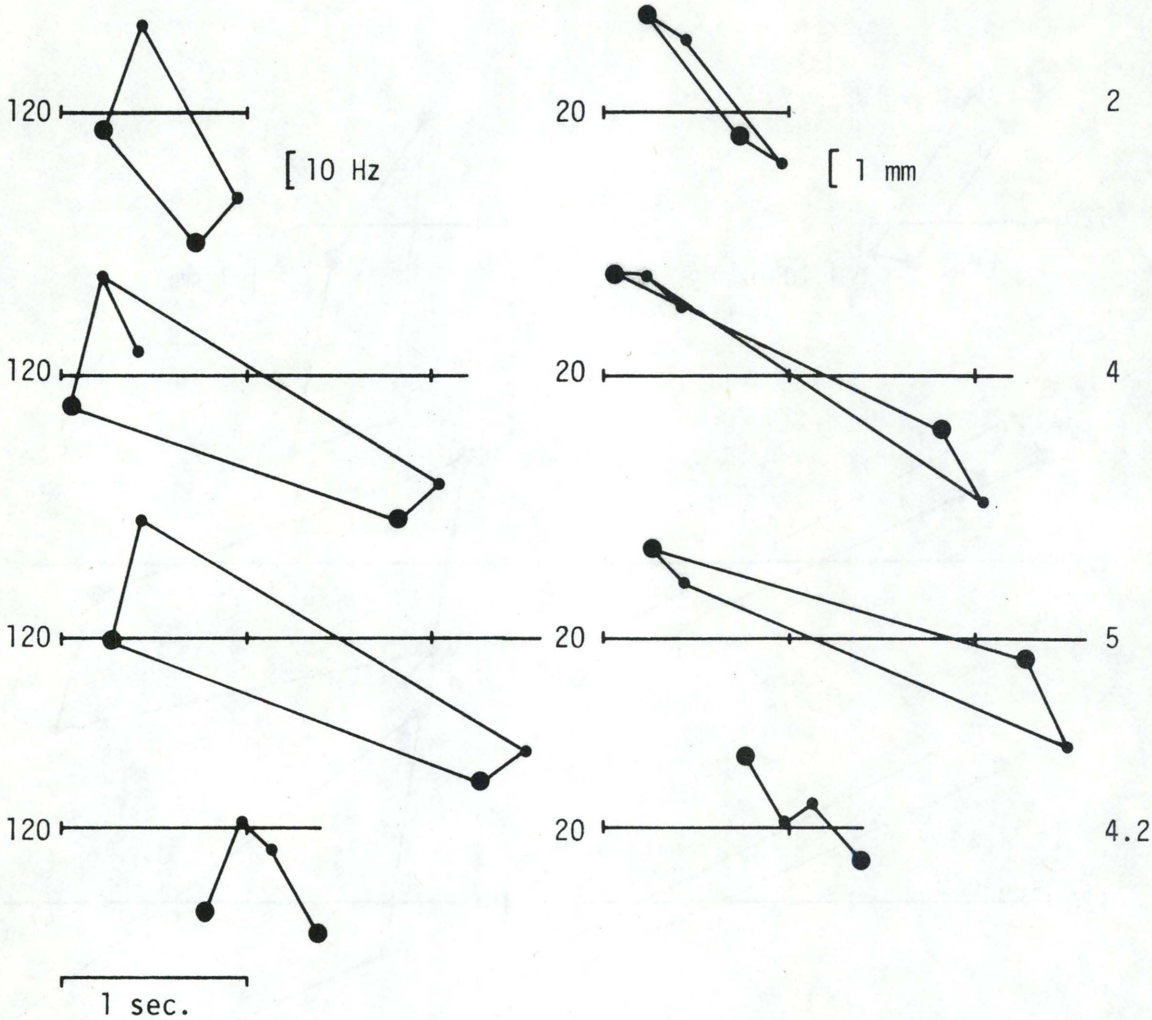


Figure 7

Material B, subject NR. See further legend to figure 4.

figures 1 to 3 that the slope of fundamental frequency decline becomes less steep as a function of sentence length.

The same general pattern of variation as a function of sentence type and length emerges for larynx height. The larynx height data, however, seem to be somewhat less consistent than the Fo data. This is seen not only in the larger deviation from the regression lines of individual data points (and, consequently, in lower correlation coefficients), but also in the instances of gross deviations of slopes from the general pattern. It should be noted, however, that the latter type of deviation is found only with PD (in his 3- and 4-stress-group declarative sentences). PD was also the less consistent speaker in the experiment reported in Reinholt Petersen 1984.

The results are summarized in figure 11, where Fo slopes have been plotted against larynx height slopes. The correlation is positive and statistically significant ($p<0.01$) for all three subjects (NR: $r=0.917$, PD: $r=0.856$, and PM: $r=0.926$). Thus, on the basis of material A, it seems justified to conclude that,

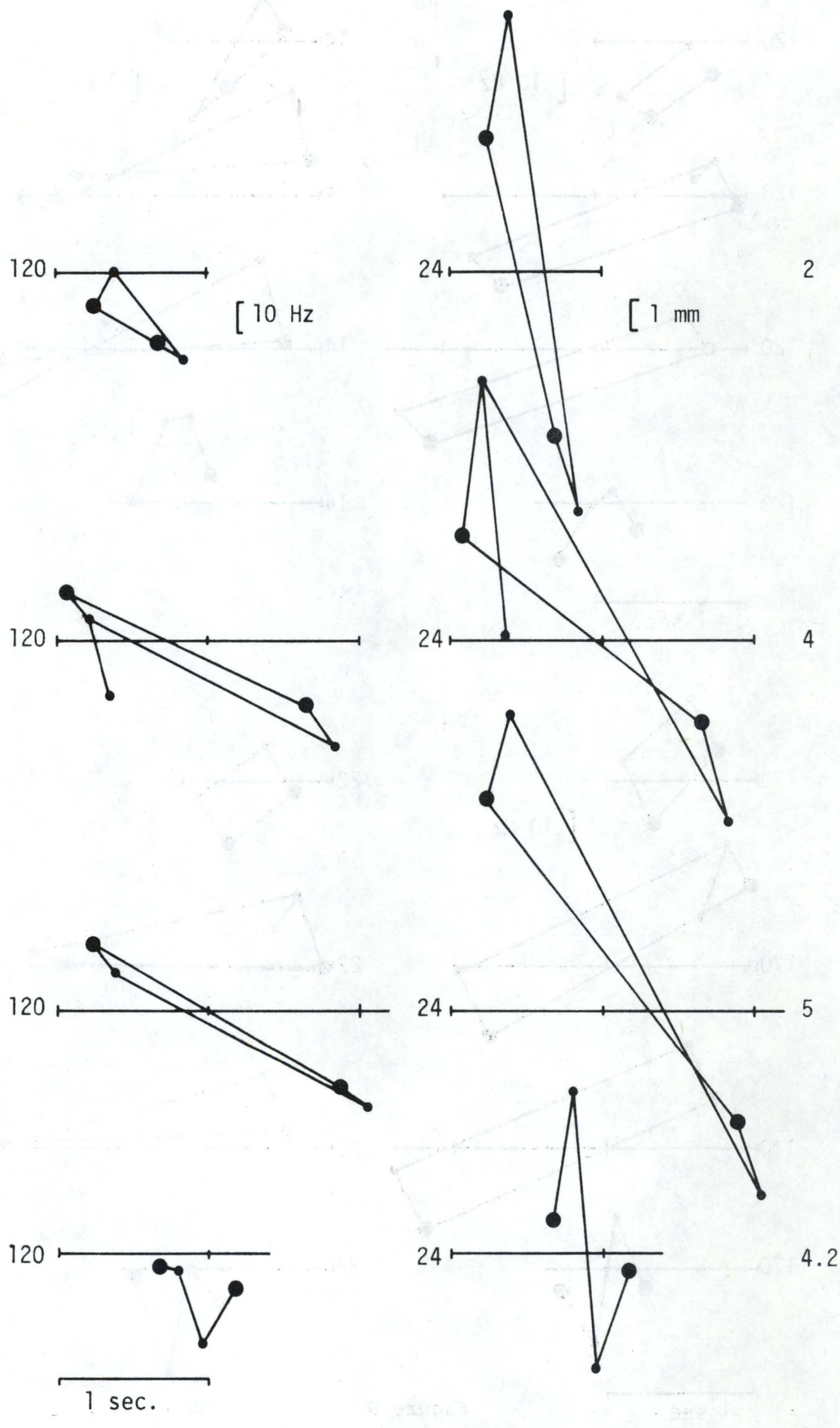


Figure 8

Material B, subject OB. See further legend to figure 4.

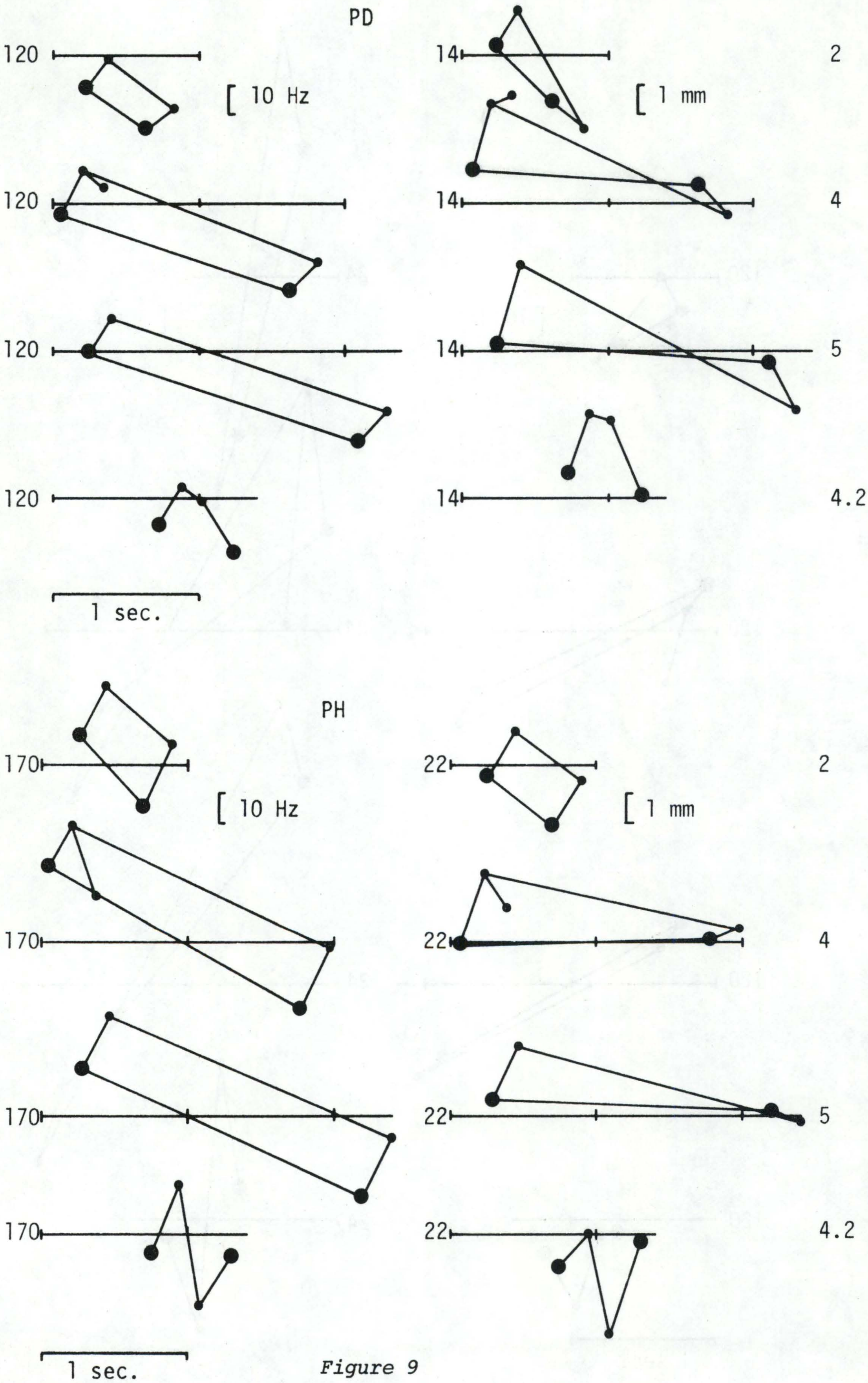


Figure 9

Material B, subjects PD and PH. See further legend to figure 4.

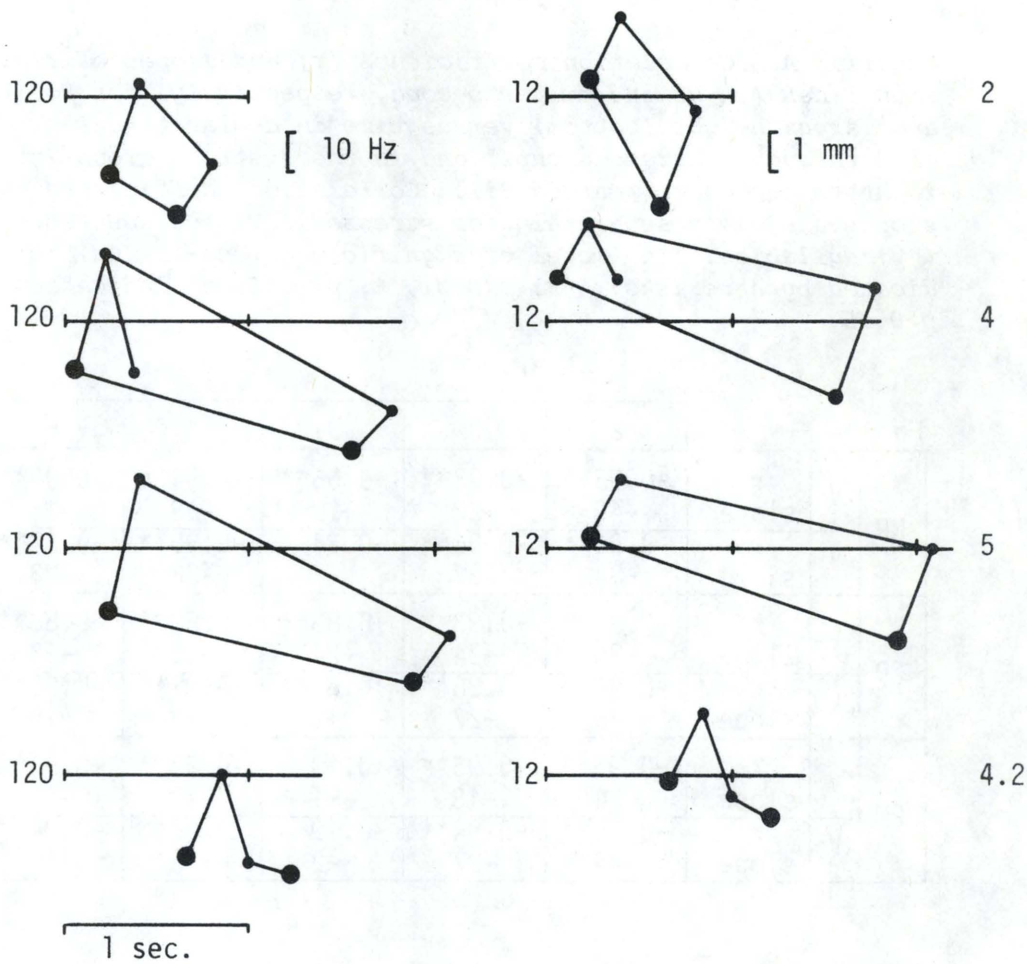


Figure 10

Material B, subject PM. See further legend to figure 4.

on the whole, varying sentence type and length influence fundamental frequency and larynx height in very much the same way.

2. MATERIAL B

In material B Fo and larynx height declination in declarative sentences could be observed by comparing the first and last stress group in sentences containing 2, 4, and 5 stress groups (cf. section III.A above). As is seen from table II Fo is always lower in the last than in the first stress group, this being true for stressed and first posttonic syllables alike. The differences were statistically significant ($p<0.01$) in all cases. Table II also gives the 'slopes' of declination. 'Slope' is here to be understood as the slope (in Hz or mm per second) of the line connecting data points of the first and the last stress group in the sentence, as is seen in figures 4 to 10.

TABLE I

*Material A. Correlation coefficients (r) and slopes of regression lines (in Hz and mm per second, respectively) for Fo (top) and larynx height (bottom) versus time in declarative sentences of 2 through 5 stress groups, and in the 3 stress group interrogative sentences (marked 3i). Correlation coefficients and slopes are given separately for stressed ('V) and unstressed ($_{\circ}$V) syllables. The level of significance (one-tailed) is indicated by asterisks: **: p<0.01, *: p<0.05, no indication: p>0.05.*

Fθ			2	3	3i	4	5
NR	'V	r	-0.93**	-0.92**	+0.55**	-0.94**	-0.90**
		slope	-49	-22	+10	-18	-15
	ₒV	r	-0.96**	-0.94**	+0.34	-0.95**	-0.92**
		slope	-84	-37	+7	-28	-23
PD	'V	r	-0.84**	-0.93**	+0.83**	-0.89**	-0.88**
		slope	-29	-24	+24	-16	-13
	ₒV	r	-0.96**	-0.86**	+0.64**	-0.91**	-0.92**
		slope	-43	-27	+9	-19	-18
PM	'V	r	-0.95**	-0.95**	+0.13	-0.92**	-0.93**
		slope	-34	-18	+1	-13	-12
	ₒV	r	-0.88**	-0.94**	-0.19	-0.96**	-0.94**
		slope	-44	-27	-2	-23	-17

LH			2	3	3i	4	5
NR	'V	r	-0.85**	-0.84**	+0.38	-0.79**	-0.79**
		slope	-5.7	-2.7	+0.7	-1.7	-1.5
	ₒV	r	-0.87**	-0.81**	+0.70**	-0.99**	-0.68**
		slope	-5.8	-3.2	+2.6	-1.8	-1.3
PD	'V	r	-0.63**	+0.42*	+0.85**	+0.09	-0.07
		slope	-4.0	+2.0	+4.8	+0.3	-0.2
	ₒV	r	-0.54*	-0.28	+0.80**	-0.44*	-0.45**
		slope	-2.9	-0.8	+3.3	-1.4	-1.4
PM	'V	r	-0.92**	-0.81**	+0.33	-0.83**	-0.63**
		slope	-9.6	-3.8	+1.1	-3.1	-1.7
	ₒV	r	-0.89**	-0.69**	+0.23	-0.69**	-0.43**
		slope	-10.3	-3.5	+0.6	-2.5	-1.2

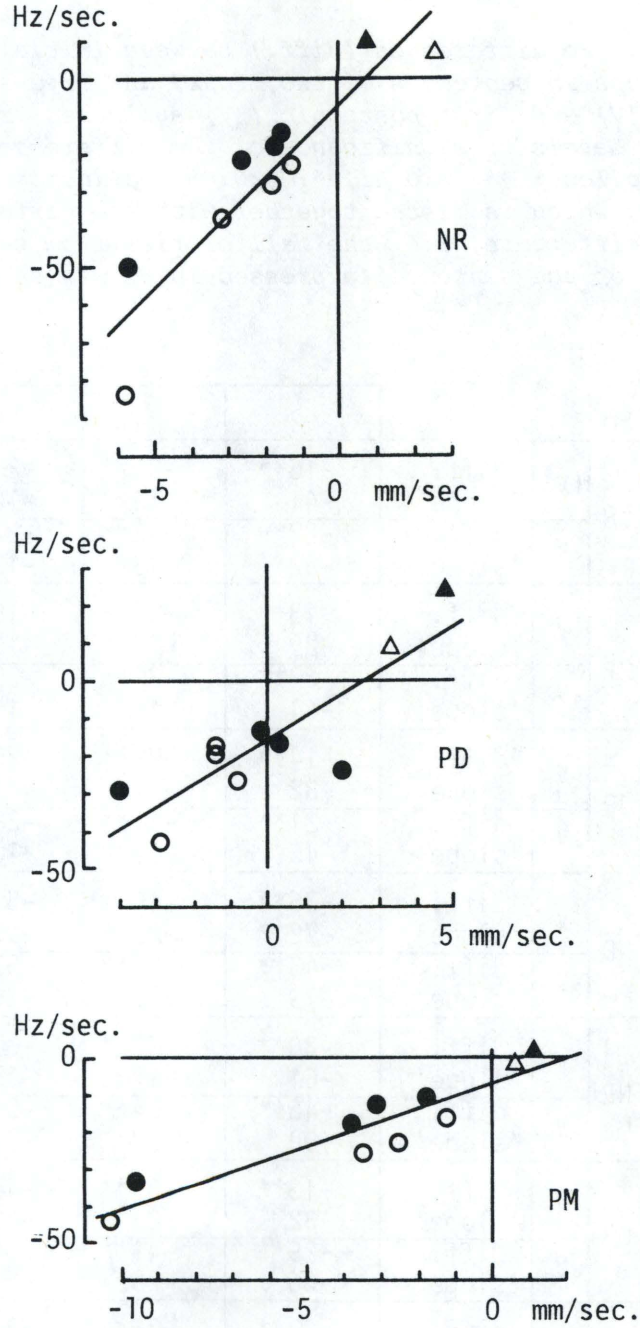


Figure 11

Material A. Overall Fo slopes (in Hz/sec.) versus overall larynx height slopes (in mm/sec.) in sentences of varying length and type. Circles indicate declarative, and triangles interrogative sentences. Filled signs indicate stressed, open signs unstressed syllables.

TABLE II

*Material B. Fo differences (diff.) between initial and final stress groups in sentences of two, four, and five stress groups. Stressed ('V) and first posttonic ($_{\circ}$V) syllables are listed separately. Levels of significance of the differences are indicated as follows: **: p<0.01, *: p<0.05, no indication: p>0.05. By 'slope', which is listed together with the differences, is meant the difference, i.e. the fall or rise from the beginning to the end of the sentence, expressed in Hz per second.*

Fo			2	4	5
CH	'V	diff.	-16**	-21**	-21**
		'slope'	-41	-13	-12
	$_{\circ}$V	diff.	-22**	-22**	-22**
		'slope'	-52	-13	-12
JB	'V	diff.	-13**	-23**	-19**
		'slope'	-29	-13	-10
	$_{\circ}$V	diff.	-22**	-32**	-25**
		'slope'	-44	-18	-13
JR	'V	diff.	-14**	-28**	-20**
		'slope'	-32	-17	-11
	$_{\circ}$V	diff.	-21**	-31**	-28**
		'slope'	-45	-19	-16
ND	'V	diff.	-36**	-46**	-53**
		'slope'	-72	-23	-23
	$_{\circ}$V	diff.	-43**	-53**	-57**
		'slope'	-79	-26	-25
NR	'V	diff.	-30**	-30**	-37**
		'slope'	-61	-17	-18
	$_{\circ}$V	diff.	-46**	-54**	-51**
		'slope'	-90	-30	-30
OB	'V	diff.	-13**	-36**	-48**
		'slope'	-32	-23	-29
	$_{\circ}$V	diff.	-29**	-41**	-44**
		'slope'	-66	-25	-26
PD	'V	diff.	-13**	-26**	-30**
		'slope'	-32	-17	-16
	$_{\circ}$V	diff.	-17**	-31**	-33**
		'slope'	-38	-19	-18
PH	'V	diff.	-25**	-49**	-43**
		'slope'	-57	-29	-22
	$_{\circ}$V	diff.	-21**	-41**	-41**
		'slope'	-44	-23	-21
PM	'V	diff.	-11**	-23**	-19**
		'slope'	-29	-15	-12
	$_{\circ}$V	diff.	-22**	-32**	-40**
		'slope'	-54	-21	-24

TABLE III

*Material B. Larynx height differences (diff.) between initial and final stress groups in sentences of two, four, and five stress groups. Stressed ('V) and first posttonic ($_{\circ}V$) syllables are listed separately. Levels of significance of the differences are indicated as follows: **: $p<0.01$, *: $p<0.05$, no indication: $p>0.05$. By 'slope', which is listed together with the differences, is meant the difference, i.e. the fall or rise from the beginning to the end of the sentence, expressed in mm per second.*

LH			2	4	5
CH	'V	diff.	-0.3	+1.4**	+0.6
		'slope'	-0.8	+0.8	+0.3
	$_{\circ}V$	diff.	-0.8*	+0.8**	+0.5
		'slope'	-1.9	+0.5	+0.3
JB	'V	diff.	-1.1*	-0.4	-0.5
		'slope'	-2.5	-0.2	-0.3
	$_{\circ}V$	diff.	-2.5**	-4.3**	-4.0**
		'slope'	-5.0	-2.4	-2.1
JR	'V	diff.	-0.6	-0.3	-0.4
		'slope'	-1.4	-0.2	-0.2
	$_{\circ}V$	diff.	-0.8	-0.7	-1.1**
		'slope'	-1.7	-0.4	-0.6
ND	'V	diff.	-3.6*	-3.6**	-5.7**
		'slope'	-7.2	-1.8	-2.5
	$_{\circ}V$	diff.	-0.8	-1.7*	-5.1**
		'slope'	-1.5	-0.8	-2.2
NR	'V	diff.	-3.4**	-4.2**	-3.5**
		'slope'	-6.9	-2.4	-1.7
	$_{\circ}V$	diff.	-3.3**	-6.0**	-4.4**
		'slope'	-6.4	-3.3	-2.1
OB	'V	diff.	-10.0**	-6.8**	-10.5**
		'slope'	-24.3	-4.3	-6.4
	$_{\circ}V$	diff.	-16.2**	-14.4**	-15.8**
		'slope'	-37.0	-8.9	-9.5
PD	'V	diff.	-2.0**	-0.5	-0.6
		'slope'	-5.0	-0.3	-0.3
	$_{\circ}V$	diff.	-4.0**	-3.7**	-4.9**
		'slope'	-9.0	-2.3	-2.6
PH	'V	diff.	-1.8**	+0.3	-0.4
		'slope'	-4.1	+0.2	-0.2
	$_{\circ}V$	diff.	-1.8**	-1.9**	-2.6**
		'slope'	-3.8	-1.1	-1.3
PM	'V	diff.	-3.4**	-3.1**	-2.8**
		'slope'	-9.1	-2.1	-1.7
	$_{\circ}V$	diff.	-2.5**	-1.7**	-1.8**
		'slope'	-6.1	-1.1	-1.1

The slopes are, of course, always negative and, like those of material A, they tend to decrease in steepness with increased sentence length. It is worth noting that the slopes for subjects NR, PD, and PM compare very well with the slopes observed for these subjects in 2-, 4- and 5-stress-group-sentences in material A, even if the 4-stress-group-sentence in the present material has a slightly different rhythmical structure than the 4-stress-group-sentences in material A and the 2- and 5-stress-group-sentences in both materials. Therefore it was considered justified to include that sentence type in the evaluation of the major tendencies in material B.

As was the case for material A, the larynx height data in material B are less consistent than the Fo data, see figures 4 to 10 and table III. There are a few cases where an Fo decline is accompanied by a larynx height rise, such as in CH's 4-stress-group-sentences, where the rise is statistically significant ($p<0.01$), and in his 5-stress-group-sentence, where it is non-significant. Likewise, a non-significant larynx rise is found in PH's 4-stress-group-sentence. In the majority of cases, however, a larynx height decline accompanies the fundamental frequency decline. This is true in 49 out of 54 cases (9 subjects x 2 stress categories x 3 sentence lengths), and out of the 49 cases 37 were statistically significant ($p<0.05$). In 5 cases (of which 2 were significant, $p<0.01$) there was a larynx rise over the sentence; 4 of these cases occurred with one speaker, viz. CH.

Finally, there is a strong tendency for the larynx height slope to be steeper in 2-stress-group-sentences than in 4- and 5-stress-group-sentences. Apart from the few instances of a positive slope, there is only one exception to this tendency, namely ND's first posttonic syllables, where the steepest slope is found in the 5-stress-group-sentences.

Thus, the data derived from material B can be taken to corroborate the conclusions arrived at in section III.A.1 above on the basis of material A, namely that fundamental frequency and larynx height show similar patterns of variation as a function of sentense type and length.

B. FUNDAMENTAL FREQUENCY AND LARYNX HEIGHT IN STRESSED AND UNSTRESSED SYLLABLES

1. MATERIAL A

Table IV gives differences in Fo and larynx height between stressed and first posttonic syllables in the test words in material A (see also figures 1 to 3). It is seen that the three subjects in all cases have significant ($p<0.01$) Fo rises between the two syllables. The rises are accompanied by larynx height rises for two of the subjects, PD and PM, whereas for the third subject, NR, the rises are accompanied by larynx lowering. Thus, there seems to be two distinctly different patterns of larynx movement corresponding to the same pattern of

TABLE IV

*Material A. Fo differences (top) and larynx height differences (bottom) between stressed and first posttonic syllables in first through fifth stress group (columns) in sentences of two through five stress groups (rows). The row marked 3i lists data for the three-stress-group interrogative sentence. Significance levels (one-tailed) are indicated as follows: **: p<0.01, *: p<0.05, no indication: p>0.05.*

Fo		1	2	3	4	5
NR	2	+28**	+8**			
	3	+30**	+19**	+11**		
	3i	+32**	+25**	+29**		
	4	+29**	+21**	+13**	+11**	
	5	+30**	+25**	+16**	+12**	+10**
PD	2	+18**	+11**			
	3	+18**	+13**	+9**		
	3i	+30**	+25**	+16**		
	4	+16**	+15**	+17**	+10**	
	5	+14**	+18**	+13**	+8**	+6**
PM	2	+22**	+17**			
	3	+28**	+20**	+17**		
	3i	+27**	+17**	+23**		
	4	+27**	+22**	+17**	+14**	
	5	+26**	+20**	+21**	+20**	+16**

LH		1	2	3	4	5
NR	2	-2.0**	-2.3**			
	3	-1.8**	-2.4**	-2.5**		
	3i	-1.8**	-0.9*	+0.4		
	4	-1.9**	-2.5**	0.0	-3.6**	
	5	-2.1**	-2.9**	-2.3**	-1.8**	-2.6**
PD	2	+2.0**	+2.3**			
	3	+3.6**	+1.6*	+0.8		
	3i	+4.6**	+2.8**	+3.3**		
	4	+3.3**	+2.5**	+1.8*	+0.6	
	5	+3.9**	+2.0**	+2.9**	+2.4**	+0.8
PM	2	+2.6**	+2.0**			
	3	+2.8**	+3.8**	+3.0**		
	3i	+2.5**	+2.9**	+2.1**		
	4	+2.4**	+2.6**	+3.8**	+2.9**	
	5	+3.0**	+2.4**	+3.6**	+6.0**	+2.3**

fundamental frequency movement. This is in agreement with the results reported in Reinholt Petersen 1984. The larynx movements - whether upward or downward - were statistically significant ($p<0.01$) in the vast majority of cases.

For Fo there is a tendency for the magnitude of the rise to become gradually smaller along the sentence. A similar tendency is seen for larynx height only in speaker PD, whereas in PM and NR there seems to be no correlation between the amount of Fo movement and the amount of larynx movement. This is illustrated in fig. 12, where Fo movements are plotted against larynx movements. The correlation coefficients were for NR 0.301 ($p>0.05$), for PD 0.661 ($p<0.01$), and for PM -0.131 ($p>0.05$).

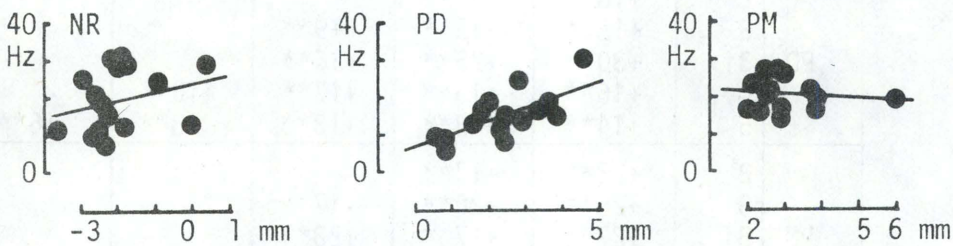


Figure 12

Material A. Fo movement versus vertical larynx movement between stressed and first posttonic syllables.

2. MATERIAL B

In material B it was possible not only to examine stressed and first posttonic syllables but also, in sentence 4.1/4 and 4.2, second posttonic syllables, and in sentence 4.2 the stressed vowel in the following stress group (see section III.4 above). Tables V and VI list the directions and magnitudes of Fo and larynx movements between these syllables (see also figures 4 to 9).

It is seen that all subjects but one (OB) have a fundamental frequency rise from stressed to first posttonic syllable, followed by a fall toward the relatively low level of the next stressed syllable. This pattern of Fo movement in the stress group conforms to the one described by Thorsen (e.g. Thorsen 1979) for Standard Copenhagen Danish. Subject OB, whose speech is appreciably influenced by his dialectal background (Northern Jutland), tends to show the opposite Fo pattern, viz. high stressed and lower unstressed syllables (cf. Thorsen 1982).

In figure 13 fundamental frequency movements have been plotted against larynx movements. Focusing first on the relation between rising fundamental frequency and larynx movement it is clear that, as in material A, the speakers can be divided into

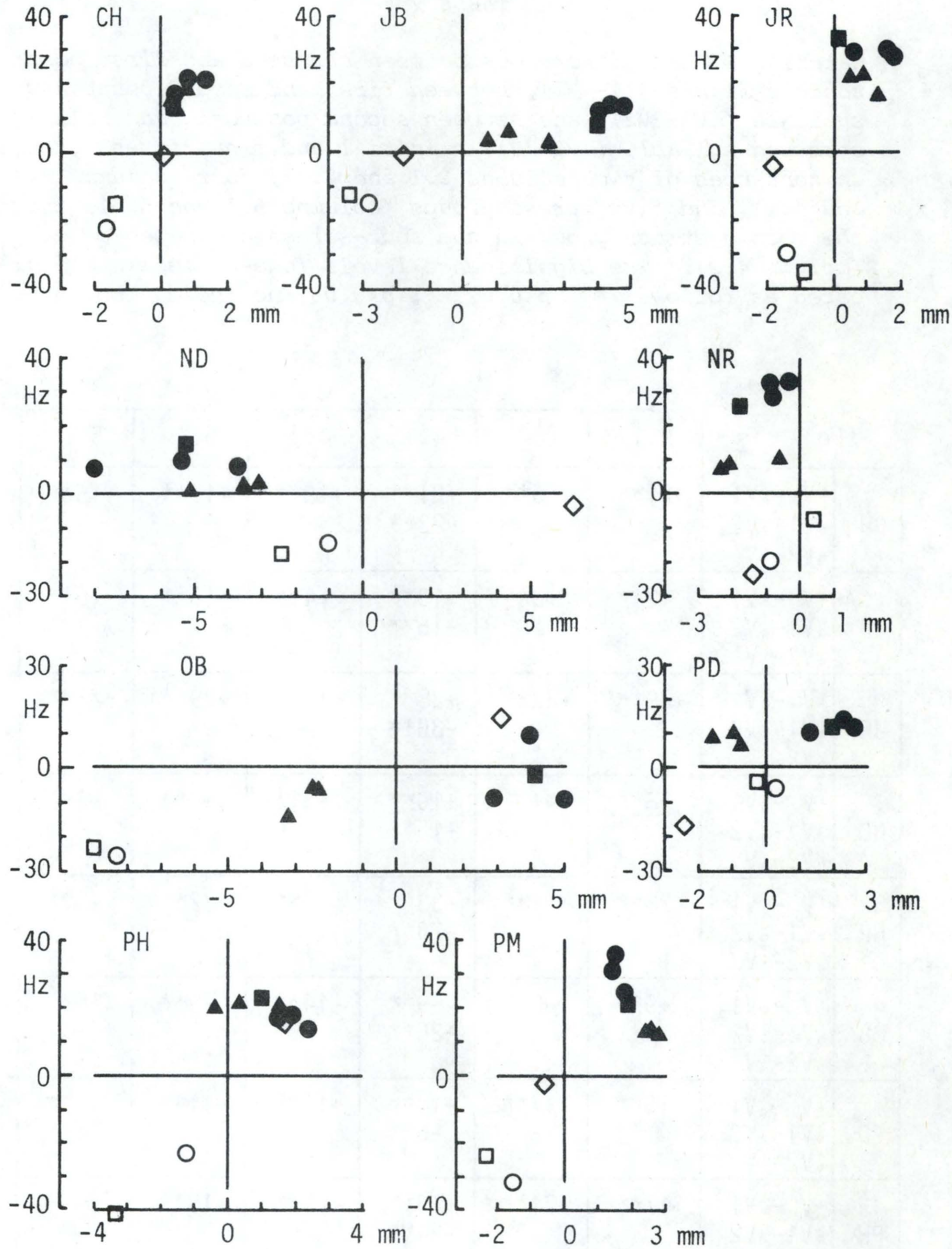


Figure 13

Material B. Fo movement versus larynx height movement between the following syllables:

between stressed and 1st posttonic, initial stress groups (●)
– – – – – *final* – – (▲)
– – – – – *2nd stress group sentence 4.2* (■)
– *1st* – *2nd* – *initial stress group sentence 4.1/4* (○)
– – – – – *2nd stress group sentence 4.2* (□)
– *2nd* – *stressed* – – – – – (◇)

TABLE V

Material B. Fo differences between stressed and first posttonic syllable ($'V-{}_\circ V1$)*, between first and second posttonic syllable* (${}_\circ V1-{}_\circ V2$)*, and between second posttonic and following stressed syllable* (${}_\circ V2-'V$) *in initial and final stress groups in sentences of two (columns 2.1 and 2.2), four (columns 4.1 and 4.4), and five stress groups (columns 5.1 and 5.5), and in the second stress group in the four-stress-group-sentences (column 4.2). The significance levels (one-tailed) are indicated as follows: **: p<0.01, *: p<0.05, no indication: p>0.05.*

Fo		2.1	2.2	4.1	4.4	5.1	5.5	4.2
CH	$'V-{}_\circ V1$	+21**	+15**	+21**	+19**	+17**	+15**	+12**
	${}_\circ V1-{}_\circ V2$			-22**				-15**
	${}_\circ V2-'V$							-1*
JB	$'V-{}_\circ V1$	+12**	+3**	+13**	+4**	+13**	+6**	+8**
	${}_\circ V1-{}_\circ V2$			-15**				-12**
	${}_\circ V2-'V$							-1
JR	$'V-{}_\circ V1$	+29**	+22**	+28**	+17**	+29**	+22**	+33**
	${}_\circ V1-{}_\circ V2$			-30**				-36**
	${}_\circ V2-'V$							-4**
ND	$'V-{}_\circ V1$	+8**	+1	+10**	+3**	+8**	+3*	+15**
	${}_\circ V1-{}_\circ V2$			-15**				-18**
	${}_\circ V2-'V$							-4
NR	$'V-{}_\circ V1$	+28**	+11**	+33**	+9**	+32**	+8**	+25**
	${}_\circ V1-{}_\circ V2$			-20**				-8**
	${}_\circ V2-'V$							-23**
OB	$'V-{}_\circ V1$	+9**	-6*	-9**	-14**	-9**	-6*	-2
	${}_\circ V1-{}_\circ V2$			-26**				-23**
	${}_\circ V2-'V$							+17**
PD	$'V-{}_\circ V1$	+10**	+7**	+14**	+10**	+11**	+9**	+12**
	${}_\circ V1-{}_\circ V2$			-6**				-4**
	${}_\circ V2-'V$							-17**
PH	$'V-{}_\circ V1$	+17**	+21**	+13**	+21**	+18**	+20**	+23**
	${}_\circ V1-{}_\circ V2$			-24**				-42**
	${}_\circ V2-'V$							+17**
PM	$'V-{}_\circ V1$	+25**	+14**	+31**	+12**	+35**	+13**	+21**
	${}_\circ V1-{}_\circ V2$			-32**				-23**
	${}_\circ V2-'V$							-3*

TABLE VI

*Material B. Larynx height differences between stressed and first posttonic syllable ($'V-{}_{\circ}V1$), between first and second posttonic syllable (${}_{\circ}V1-{}_{\circ}V2$), and between second posttonic and following stressed syllable (${}_{\circ}V2-'V$) in initial and final stress groups in sentences of two (columns 2.1 and 2.2), four (columns 4.1 and 4.4), and five stress groups (columns 5.1 and 5.5), and in the second stress group in the four-stress-group-sentences (column 4.2). The significance levels (one-tailed) are indicated as follows: **: $p<0.01$, *: $p<0.05$, no indication: $p>0.05$.*

LH		2.1	2.2	4.1	4.4	5.1	5.5	4.2
CH	$'V-{}_{\circ}V1$	+0.8**	+0.3	+1.4**	+0.8**	+0.5	+0.4	+0.4
	${}_{\circ}V1-{}_{\circ}V2$			-1.6**				-1.4**
	${}_{\circ}V2-'V$							+0.1
JB	$'V-{}_{\circ}V1$	+4.0**	+2.6**	+4.4**	+0.7*	+4.8**	+1.3*	+4.0**
	${}_{\circ}V1-{}_{\circ}V2$			-2.8**				-3.4**
	${}_{\circ}V2-'V$							-1.8*
JR	$'V-{}_{\circ}V1$	+0.6	+0.5	+1.8**	+1.4**	+1.6**	+0.9	+0.1
	${}_{\circ}V1-{}_{\circ}V2$			-1.4**				-0.9
	${}_{\circ}V2-'V$							-1.8**
ND	$'V-{}_{\circ}V1$	-8.0**	-5.1**	-5.3**	-3.5**	-3.7**	-3.1**	-5.2**
	${}_{\circ}V1-{}_{\circ}V2$			-1.0*				-2.4**
	${}_{\circ}V2-'V$							+6.3**
NR	$'V-{}_{\circ}V1$	-0.8	-0.6	-0.3	-2.1**	-0.9*	-2.3**	-1.8**
	${}_{\circ}V1-{}_{\circ}V2$			-0.8**				+0.4
	${}_{\circ}V2-'V$							-1.4*
OB	$'V-{}_{\circ}V1$	+4.0**	-2.5*	+5.0**	-3.2**	+2.9**	-2.4**	+4.1**
	${}_{\circ}V1-{}_{\circ}V2$			-8.3**				-9.0**
	${}_{\circ}V2-'V$							+3.1**
PD	$'V-{}_{\circ}V1$	+1.3**	-0.8	+2.3**	-1.0*	+2.6**	-1.6*	+2.0**
	${}_{\circ}V1-{}_{\circ}V2$			+0.3				-0.3
	${}_{\circ}V2-'V$							-2.5**
PH	$'V-{}_{\circ}V1$	+1.5**	+1.5**	+2.4**	+0.3	+1.9**	-0.4	+1.0
	${}_{\circ}V1-{}_{\circ}V2$			-1.2**				-3.4**
	${}_{\circ}V2-'V$							+1.7**
PM	$'V-{}_{\circ}V1$	+1.8*	+2.6**	+1.4**	+2.9**	+1.5*	+2.5**	+1.9**
	${}_{\circ}V1-{}_{\circ}V2$			-1.5**				-2.3**
	${}_{\circ}V2-'V$							-0.6

two groups: a smaller one comprising ND and NR, who consistently show downward larynx displacements, and a larger one consisting of the remaining speakers except PD. This group all have upward larynx movements in conjunction with Fo rises. There is one exception: in one case in speaker PH the Fo rise is associated with a (non-significant, $p<0.05$) fall of the larynx.

Speaker PD seems to have a distinction between sentence final and non-final stress groups; in the former the larynx is shifted downwards, in the latter upwards, even if Fo is rising in both cases. This pattern of variation deviates from the findings for PD in material A and in Reinholt Petersen 1984. There, Fo rises were always associated with larynx rises, whether in final or in non-final position. The larynx lowering in final stress groups observed for PD in the present material could perhaps be explained as an anticipation of the low level of the larynx which seems to be characteristic of the pauses between sentences (see section C below). If final stress groups are left out of consideration PD shows a pattern of larynx movement which is essentially congruent with the pattern of the larger group of speakers mentioned above. The question why PD's final stress groups behaved differently in the present material as compared to previous recordings still remains open, of course, and can hardly be addressed on the basis of the available data.

For the group of speakers having larynx rises accompanying Fo rises, it should be noted that only three of them (CH, JB, and PD) have positive correlations between the amount of Fo movement and the amount of larynx movement (CH: $r=0.333$, $p<0.01$, JB: $r=0.492$, $p<0.01$, PD: $r=0.302$, $p<0.05$). The remaining four speakers all have negative correlations (JR: $r=-0.064$, $p>0.05$, OB: $r=-0.160$, $p>0.05$, PH: $r=-0.099$, $p>0.05$, PM: $r=-0.560$, $p<0.01$), but only for PM the correlation can be shown to be statistically significant.

The patterns of larynx movement associated with fundamental frequency lowering seem to be less consistent with respect to their distribution among subjects. Four out of the nine subjects (ND, NR, OB, and PD) have larynx falls as well as larynx rises in such cases. The remaining five subjects always have downward larynx displacements in conjunction with a falling Fo. In the former group it should be noted, however, that the cases in which the greatest downward Fo movements are observed are also those in which the larynx is lowered. This point may, in fact, be generalized to the latter group mentioned above. Even if, in this group, Fo lowering is in all cases accompanied by larynx lowering, there is a tendency for the larynx lowering to be greater in the cases of more extreme Fo fall than in cases of a moderate one. The only exception to this general tendency is subject JR.

With respect to speaker OB it should be pointed out that there seems to be a systematic trend in the direction of larynx displacement associated with his more moderate Fo falls. In sentence final position the larynx is shifted downwards, in non-final position upwards. This distribution of larynx movements may be taken to indicate that the larynx lowering found in final position has the same explanation as was suggested for subject PD, namely an anticipation of the low level of the larynx during pauses.

C. LARYNX HEIGHT IN PAUSES

Figures 14 and 15 display average larynx height traces during the pause between the two identical 4-stress-group sentences. For the sake of comparison mean larynx height in all vowels in the two sentences is also plotted. As is seen there seems to be a clear tendency for the larynx to be lower in the pause than during speech. This tendency is very clear for speakers JR, ND, NR, OB, PD, PH, and PM, less so, perhaps, for CH and JB. For most of the subjects there is a steep fall of the larynx at the beginning of the pause, and a steep rise at or immediately before the beginning of the next sentence. For subjects OB and PD the larynx depression for the pause seems to start already before the end of the sentence.

Thus, the larynx seems to have two general levels, a relatively high speech level and a lower level associated with pauses. It is, of course, a question whether the level observed in pauses corresponds to some sort of relaxation level, since the pauses were relatively short, between one and two seconds, but the levels reached in the pauses between the two sentences compare very well with the levels observed in longer pauses, e.g. before the beginning of the reading of the material.

IV. DISCUSSION

A. FUNDAMENTAL FREQUENCY DECLINATION AND LARYNX HEIGHT

The results presented in section III.A above demonstrate that Fo and larynx height show very similar patterns of variation as a function of sentence length and type. This is in disagreement with what should be expected on the basis of theories of Fo declination which regard it as an automatic consequence of constraints in the speech production apparatus, or more specifically in the pulmonic system. Together with the observations on Fo behaviour referred to in section I, the results of the present study speak strongly in favour of the view that Fo declination is actively controlled by the speaker.

The results also suggest that it is the extrinsic rather than the intrinsic laryngeal muscles that are responsible for the gradual Fo decline over the sentence. Changes in the activity of the vocalis muscle could hardly give rise to changes in larynx height. The cricothyroid muscle could, however, be thought to influence the apparent larynx height. If the activity of that muscle causes the thyroid cartilage to rotate around the cricothyroid joint it would appear, as seen from

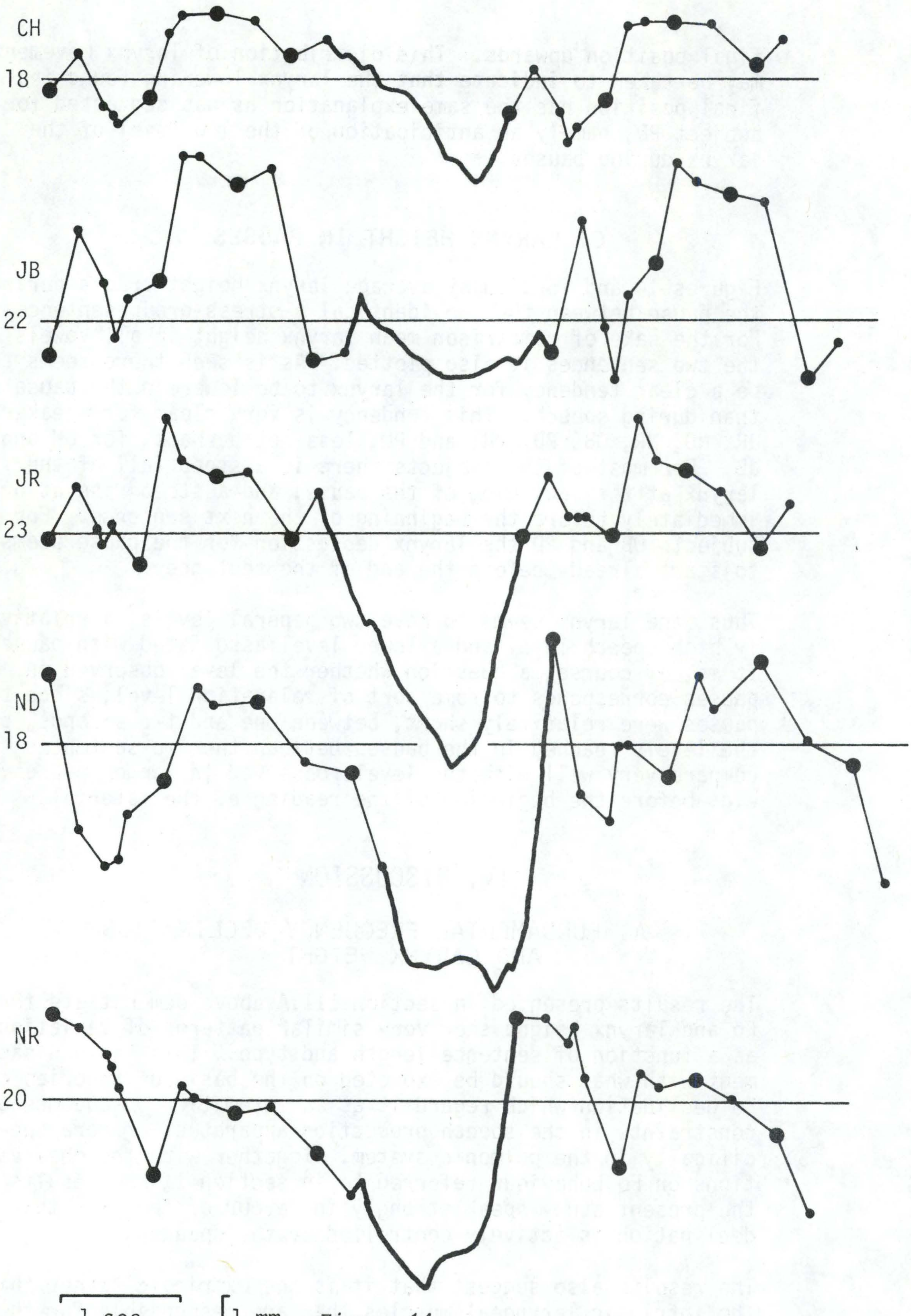


Figure 14

Material B. Subjects CH, JB, JR, ND, and NR. Average larynx height curves (time normalized for each subject) in the two identical sentences, 4.1/4 (thin lines) and in the pause between the sentences (heavy lines). Stressed syllables are indicated by large and unstressed syllables by small dots. The level of the horizontal reference line is given in mm, using an arbitrary zero.

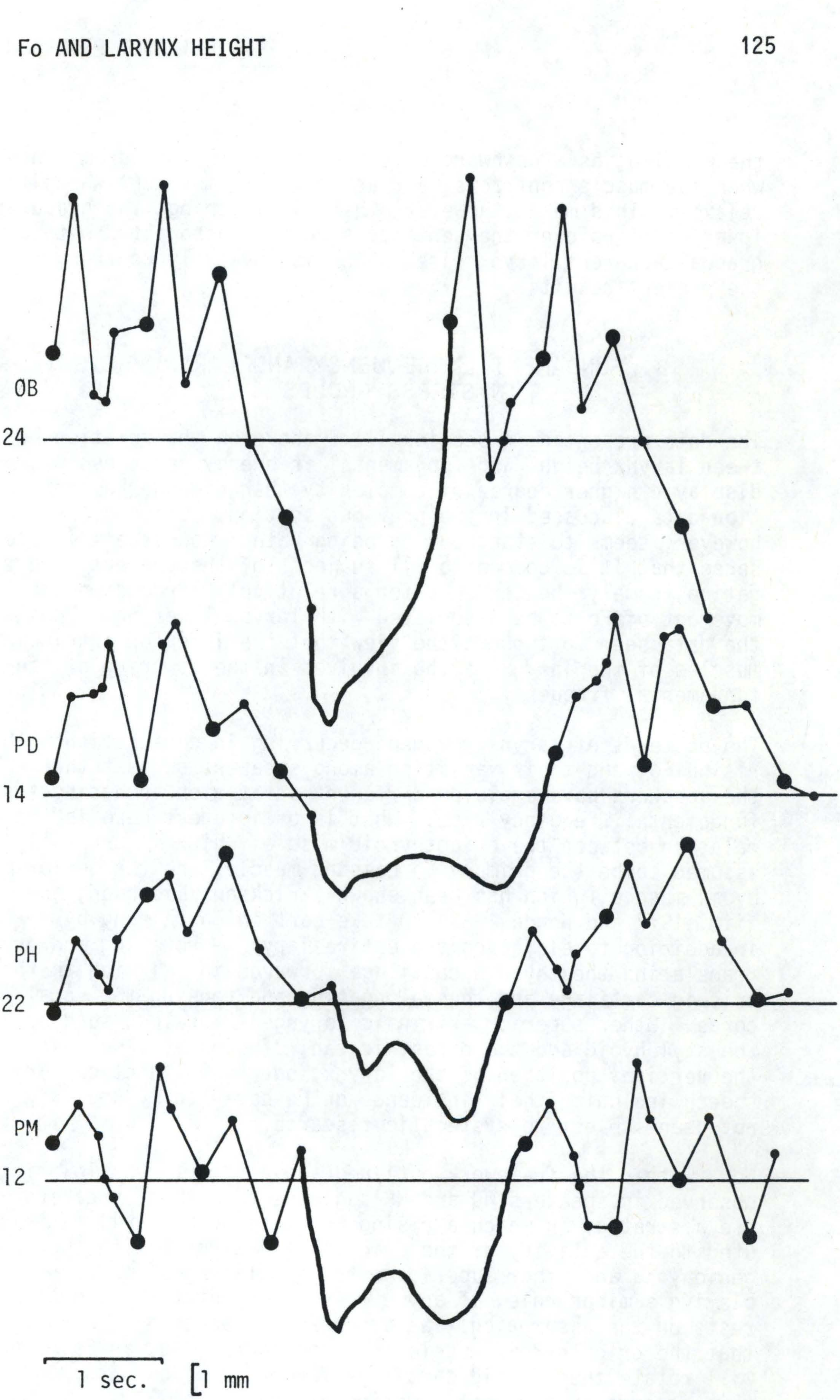


Figure 15

Material B. Subjects OB, PD, PH, and PM. See further legend to figure 14.

the outside, as a downward movement of the thyroid prominence when the muscle contracts, and as an upward movement when it relaxes. This means, however, that, if anything, the gradual lowering of Fo over the sentence should be associated with a gradual apparent larynx rise, which is obviously contrary to the present results.

B. FUNDAMENTAL FREQUENCY AND LARYNX HEIGHT IN STRESS GROUPS

The data presented in section III.B above on the relation between larynx height and fundamental frequency in stress groups display a higher degree of complexity than did the Fo declination data discussed in the previous section. One feature, however, seems to stand out as being rather consistent in the sense that it is common to all subjects of the present investigation, namely the tendency for a relatively large downward movement of Fo to be associated with larynx lowering. Thus, the data seem to support the view that the inferior extrinsic muscles of the larynx may be involved in the lowering of the fundamental frequency.

The patterns of larynx movement occurring in conjunction with rising Fo, and their variation among speakers suggest that the speakers have employed different strategies producing the fundamental frequency rise. What is of interest here is the relation between the cricothyroid muscle, which is generally assumed to be the primary Fo raising muscle, and the geniohyoid muscle, which has been shown (Erickson, Liberman, and Niimi 1977 and Honda 1983) to take part in Fo raising by - in addition to elevating the entire larynx - rotating and/or translating the thyroid cartilage forwards in relation to the cricoid cartilage and thus elongating and tensing the vocal cords. Other superior extrinsic laryngeal muscles, such as the stylohyoid and the digastric can, of course, also shift the vertical position of the larynx, but their function during speech including their influence on Fo has - to my knowledge - not been subject to systematic research.

Now, within the framework outlined above, the larynx lowering observed in speakers ND and NR suggests that these speakers use a strategy in which a rising Fo is to be primarily attributed to the activity of the cricothyroid muscle, with the geniohyoid and other superior extrinsic laryngeal muscles playing a minor role, if any. This interpretation of data rests on the assumption - as discussed in section IV.A above - that the cricothyroid muscle when contracting for an Fo rise will rotate the thyroid cartilage forward and downward, and thus produce an apparent lowering of the larynx.

The larynx elevation observed in the remaining speakers (CH, JB, JR, OB, PD, PH, and PM) suggests that these speakers employ a strategy for raising Fo which involves also the activity of the superior extrinsic laryngeal muscles. And further, it may be ventured that the differences observed within this group in regard to the correlation between larynx and Fo movement (cf. section III.B.1 above) may reflect different degrees of relative importance of intrinsic and extrinsic muscles in increasing the fundamental frequency. The positive correlation between larynx and Fo movement displayed by CH, JB, and PD may be the result of a relatively great and direct effect of the extrinsic muscles on Fo - the higher the Fo rise the more the extrinsic muscles will have to pull on the larynx. The opposite tendency is seen in speaker PM, who has a strong negative correlation between larynx rise and Fo rise. Here the extrinsic muscles can be assumed to pull upon the larynx by some force which is independent of the magnitude of the Fo rise to be produced. The cricothyroid will then be the muscle which directly controls the actual amount of Fo rise. The activity of that muscle will produce a downward (and forward) movement of the laryngeal prominence the magnitude of which will depend on the magnitude of the Fo rise, and this movement together with the upward displacement caused by the activity of the extrinsic muscles will, as observed from the outside, bring about an upward larynx movement which is inversely proportional to the Fo rise produced. The speakers JR, OB, and PH, who have non-significant (negative) correlations between larynx rise and Fo rise, may be assumed to represent degrees of relative activity of intrinsic and extrinsic laryngeal muscles which are intermediate between the patterns observed in CH, JB, and PD on one side, and PM on the other.

In addition to the differences in strategies between speakers discussed so far, different strategies for Fo control in stress groups can also be observed within the individual speaker, as is the case in sentence final position, where the deviating pattern presumably can be seen as an anticipation of the low level of the larynx generally observed in pauses.

The interpretation attempted here to account for the major trends in the data on Fo and larynx height in sentences and stress groups in terms of patterns of activity of extrinsic and intrinsic laryngeal muscles is, of course, highly speculative and needs further substantiation, preferably using EMG techniques, but to the extent that vertical larynx movements can be assumed to reflect muscular activity which may also influence Fo it seems justified to conclude from the data that while the overall Fo variation characterizing units of sentence size can be attributed to extrinsic laryngeal muscle activity alone, the more localized Fo variation of the stress group can only be accounted for by an intricate pattern of interaction between extrinsic and intrinsic laryngeal muscles.

REFERENCES

Atkinson, J.E. 1978: "Correlation analysis of the physiological factors concerning voice fundamental frequency", *J. Acoust. Soc. Am. 63*, p. 211-233

Atkinson, J.E. and Erickson, D. 1977: "The function of strap muscles in speech", *Status Rep. Speech Res., Haskins Labs. 49*, p. 97-102

Cohen, A., Collier, R., and 't Hart, J. 1982: "Declination: Construct or intrinsic feature of speech pitch?", *Phonetica 39*, p. 254-273

Collier, R. 1975: "Physiological correlates of intonation patterns", *J. Acoust. Soc. Am. 58*, p. 249-255

Erickson, D. and Atkinson, J.E. 1976: "The function of the strap muscles in speech", *Status Rep. Speech Res., Haskins Labs. 45/46*, p. 205-210

Erickson, D., Liberman, M., and Niimi, S. 1977: "The geniohyoid and the role of the strap muscles", *Status Rep. Speech Res., Haskins Labs. 49*, p. 103-110

Ewan, W.G. 1979: "Laryngeal behavior in speech", *Report of the Phonology Laboratory 3*, University of California, Berkeley, p. 1-93

Gandour, J. and Maddieson, I. 1976: "Measuring larynx movement in standard Thai using the cricothyrometer", *Working Papers in Phonetics, UCLA 33*, p. 160-190

Gelfer, C.E., Harris, K.S., Collier, R., and Baer, T. 1983: "Speculations on the control of fundamental frequency declination", *Status Rep. Speech Res., Haskins Labs. 76*, p. 51-63

Honda, K. 1983: "Relationship between pitch control and vowel articulation", *Status Rep. Speech Res., Haskins Labs. 73*, p. 269-282

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

Maeda, S. 1974: "A characterization of fundamental frequency contours of speech", *MIT Quart. Prog. Rep. 114*, p. 193-211

Maeda, S. 1979: "On the Fo control mechanisms of the larynx", *Séminaire Larynx & Parole, Institut de Phonétique Grenoble, 8-9 fév. 1979*, p. 245-257

Ohala, J.J. 1970: "Aspects of the control and production of speech", *Working Papers in Phonetics, UCLA, 15*

Ohala, J.J. 1973: "The physiology of tone", in: *Consonant Types & Tone (ed.: L.M. Hyman), Southern California Occasional Papers in Linguistics, No. 1.*

Ohala, J.J. 1978: "The production of tone", *Report of the Phonology Laboratory 2* (Berkeley), p. 63-117

Reinholt Petersen, N. 1984: "Global and local fundamental frequency variation and larynx height: Some preliminary observations", *Ann. Rep. Inst. Phon., Univ. Copenhagen 18*, p. 1-12

Sawashima, M., Kakita, Y., and Hiki, S. 1973: "Activity of the extrinsic laryngeal muscles in relation to Japanese word accent", *Ann. Bull. Res. Inst. Logopedics and Phoniatrics, Tokyo, 7*, p. 19-25

Shipp, T. 1975: "Vertical laryngeal position during continuous and discrete vocal frequency change", *J. Speech and Hearing Research 18*, p. 707-718

Sorensen, J.M. and Cooper, W.E. 1980: "Syntactic coding of fundamental frequency in speech production", *in: Perception and Production of Fluent Speech (ed.: Cole, R.A.)*, (Erlbaum Associates, Hillsdale, N.J.), p. 399-440

Thorsen, N. 1979: "Interpreting raw fundamental frequency tracings of Danish", *Phonetica 36*, p. 57-78

Thorsen, N. 1980a: "A study of the perception of sentence intonation - evidence from Danish", *J. Acoust. Soc. Am. 67/3*, p. 1014-1030

Thorsen, N. 1980b: "Intonation contours and stress group patterns of declarative sentences of varying length in ASC Danish", *Ann. Rep. Inst. Phon., Univ. Copenhagen 14*, p. 1-29

Thorsen, N. 1981: "Intonation contours and stress group patterns of declarative sentences of varying length in ASC Danish - supplementary data", *Ann. Rep. Inst. Phon., Univ. Copenhagen 15*, p. 13-47

Thorsen, N. 1982: "On the variability of Fo patterning and the function of Fo timing in languages where pitch cues stress", *Phonetica 39*, p. 302-316

Thorsen, N. 1983: "Standard Danish sentence intonation - phonetic data and their representation", *Folia Linguistica XVII/1-2*, p. 187-220

APPENDIX I

Material A. The following tables list means ($\overline{X}$), standard deviations (s), and number of observations (n)' for fundamental frequency (Fo) and larynx height (LH) for each of the three speakers NR, PD, and PM. Data are arranged according to sentence length (rows) and stress group number (columns). Within each column data for stressed ('V) and first posttonic ($_\circ$V) syllables are listed separately. The row marked 3i gives data for the three-stress-group interrogative sentence.

Fo			.1		.2		.3		.4		.5	
			'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V
NR	2	$\overline{X}$	106	134	82	90						
		s	6.453	9.150	2.167	4.274						
		n	8	8	8	8						
	3	$\overline{X}$	107	137	90	109	81	92				
		s	5.605	6.341	2.850	7.100	2.642	3.615				
		n	8	8	8	8	8	8				
	3i	$\overline{X}$	107	140	101	126	117	146				
		s	6.143	6.209	3.271	4.534	3.962	3.546				
		n	8	8	8	8	8	8				
	4	$\overline{X}$	111	140	93	114	88	101	78	88		
		s	3.998	4.912	1.959	5.120	2.212	5.120	2.360	2.605		
		n	8	8	8	8	8	8	7	7		
	5	$\overline{X}$	113	143	92	117	89	105	85	97	78	88
		s	6.296	5.460	4.033	7.760	2.435	4.464	1.982	3.703	3.237	4.796
		n	8	7	7	7	8	8	8	8	7	7

LH			.1		.2		.3		.4		.5	
			'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V	'V	$_\circ$V
NR	2	$\overline{X}$	18.8	16.8	16.0	13.8						
		s	1.165	1.035	0.535	0.886						
		n	8	8	8	8						
	3	$\overline{X}$	19.3	17.5	18.4	16.0	16.1	13.6				
		s	0.707	1.512	0.744	0.926	0.991	1.188				
		n	8	8	8	8	8	8				
	3i	$\overline{X}$	19.5	17.8	19.3	18.4	20.3	20.6				
		s	0.926	1.581	0.886	1.188	0.463	0.916				
		n	8	8	8	8	8	8				
	4	$\overline{X}$	19.3	17.4	18.8	16.3	17.4	17.4	16.6	13.0		
		s	0.886	1.408	0.886	1.035	0.916	1.061	0.787	1.506		
		n	8	8	8	8	8	8	7	7		
	5	$\overline{X}$	19.5	17.4	19.1	16.3	18.4	16.1	17.6	15.9	16.1	13.6
		s	1.195	1.188	0.690	0.488	0.744	0.354	0.744	1.356	0.900	1.133
		n	8	7	7	7	8	8	8	8	7	7

Fo			1		2		3		4		5	
			'V	∘V	'V	∘V	'V	∘V	'V	∘V	'V	∘V
PD	2	$\bar{X}$	117	135	104	114						
		s	5.418	2.748	3.464	3.927						
		n	8	8	8	8						
	3	$\bar{X}$	122	141	106	119	97	106				
		s	4.173	5.574	2.900	3.739	2.850	3.196				
		n	8	8	8	8	8	8				
	3i	$\bar{X}$	108	137	119	144	131	147				
		s	8.700	4.566	5.064	5.318	7.220	5.416				
		n	8	8	8	8	8	7				
	4	$\bar{X}$	123	139	107	122	103	121	97	107		
		s	6.413	4.713	2.375	4.051	2.062	6.952	2.748	5.148		
		n	8	8	8	8	4	4	8	8		
	5	$\bar{X}$	125	139	110	128	105	118	104	112	96	102
		s	3.295	8.734	2.619	4.121	3.059	4.743	4.998	2.964	2.866	3.643
		n	8	8	8	8	8	8	8	8	8	8

LH			1		2		3		4		5	
			'V	∘V	'V	∘V	'V	∘V	'V	∘V	'V	∘V
PD	2	$\bar{X}$	22.8	24.8	21.0	23.3						
		s	1.035	1.389	1.246	0.886						
		n	8	8	8	8						
	3	$\bar{X}$	20.9	24.5	22.6	24.3	22.8	23.5				
		s	1.808	1.195	2.326	1.282	1.389	1.512				
		n	8	8	8	8	8	8				
	3i	$\bar{X}$	19.4	24.0	22.9	25.6	24.3	27.6				
		s	1.302	0.926	0.835	1.302	1.035	0.976				
		n	8	8	8	8	7	7				
	4	$\bar{X}$	21.9	25.1	21.6	24.1	22.3	24.0	22.3	22.9		
		s	1.356	0.991	2.200	1.885	0.500	0.816	1.982	2.800		
		n	8	8	8	8	4	4	8	8		
	5	$\bar{X}$	20.5	24.4	23.0	25.0	21.8	24.6	21.9	24.3	21.9	22.6
		s	2.777	1.408	2.070	1.604	2.315	1.408	1.246	1.035	2.800	3.292
		n	8	8	8	8	8	8	8	8	8	8

Fo			1		2		3		4		5	
			'V	∘V	'V	∘V	'V	∘V	'V	∘V	'V	∘V
PM	2	$\bar{X}$	104	126	90	106						
		s	2.504	7.329	1.982	2.659						
		n	8	8	8	8						
	3	$\bar{X}$	105	133	97	117	88	105				
		s	3.682	3.682	2.138	2.659	1.356	2.828				
		n	8	8	8	8	8	8				
	3i	$\bar{X}$	107	134	113	130	109	132				
		s	3.454	4.504	2.726	2.816	3.871	5.249				
		n	8	8	8	8	8	8				
	4	$\bar{X}$	107	134	98	120	95	112	87	101		
		s	4.438	3.314	2.188	4.027	2.453	3.295	1.832	2.696		
		n	8	8	8	8	8	8	8	8		
	5	$\bar{X}$	110	136	101	121	96	117	92	112	86	102
		s	2.928	3.091	5.014	2.774	2.252	3.338	1.885	2.696	2.387	4.000
		n	8	8	8	8	8	8	8	8	8	8

LH			1		2		3		4		5	
			'V	∘V	'V	∘V	'V	∘V	'V	∘V	'V	∘V
PM	2	$\bar{X}$	8.1	10.8	4.1	6.1						
		s	1.126	1.035	0.354	1.356						
		n	8	8	8	8						
	3	$\bar{X}$	8.3	11.0	7.3	11.0	4.5	7.5				
		s	1.035	1.069	1.165	1.069	0.926	1.852				
		n	8	8	8	8	8	8				
	3i	$\bar{X}$	8.5	11.0	9.5	12.4	9.5	11.6				
		s	0.756	0.756	1.069	0.518	1.690	1.061				
		n	8	8	8	8	8	8				
	4	$\bar{X}$	8.3	10.6	7.5	10.1	6.4	10.1	3.9	6.8		
		s	0.866	0.916	1.414	0.991	1.188	1.553	0.354	1.165		
		n	8	8	8	8	8	8	8	8		
	5	$\bar{X}$	7.6	10.6	8.5	10.9	7.1	10.8	5.0	11.0	5.6	7.9
		s	0.916	1.685	0.926	0.835	1.640	1.165	0.926	1.690	1.598	1.808
		n	8	8	8	8	8	8	8	8	8	8

Material B. The following tables list means ($\overline{X}$), standard deviations (s), and number of observations (n) for fundamental frequency (Fo) and larynx height (LH) for the nine speakers CH, JB, JR, ND, NR, OB, PD, PH, and PM. The data are arranged according to sentence length and stress group number (rows) and position within the stress group (columns) ordered consecutively, i.e. stressed syllable ('V column 1), first posttonic (∘V column 2), second posttonic (∘V column 3), and stressed syllable in the following stress group ('V column 4).

Fo			'V	∘V	∘V	'V
CH	2.1	$\overline{X}$	113	134		
		s	4.809	7.689		
		n	8	8		
	2.2	$\overline{X}$	96	111		
		s	2.605	1.874		
		n	8	8		
	4.1	$\overline{X}$	116	136	114	
		s	7.145	14.511	6.742	
		n	16	16	16	
	4.4	$\overline{X}$	95	114		
		s	5.310	7.016		
		n	16	16		
	5.1	$\overline{X}$	115	132		
		s	6.739	10.575		
		n	8	8		
	5.5	$\overline{X}$	94	110		
		s	3.623	2.712		
		n	8	8		
	4.2	$\overline{X}$	111	122	107	106
		s	4.892	6.163	4.534	3.739
		n	8	8	8	8

LH			'V	∘V	∘V	'V
CH	2.1	$\overline{X}$	17.9	18.6		
		s	0.354	0.518		
		n	8	8		
	2.2	$\overline{X}$	17.6	17.9		
		s	0.518	0.835		
		n	8	8		
	4.1	$\overline{X}$	17.3	18.7	17.1	
		s	0.794	0.794	0.885	
		n	16	16	16	
	4.4	$\overline{X}$	18.7	19.4		
		s	0.479	0.727		
		n	16	16		
	5.1	$\overline{X}$	18.1	18.6		
		s	0.835	0.916		
		n	8	8		
	5.5	$\overline{X}$	18.8	19.1		
		s	0.707	0.354		
		n	8	8		
	4.2	$\overline{X}$	18.8	19.1	17.8	17.9
		s	0.463	0.835	0.463	0.641
		n	8	8	8	8

Fo			'V	∘V	∘V	'V
JB	2.1	$\overline{X}$	118	130		
		s	2.976	3.137		
		n	8	8		
	2.2	$\overline{X}$	105	108		
		s	1.309	2.816		
		n	8	8		
	4.1	$\overline{X}$	124	137	121	
		s	3.907	3.722	2.626	
		n	16	16	16	
	4.4	$\overline{X}$	101	106		
		s	2.330	4.147		
		n	16	16		
	5.1	$\overline{X}$	121	133		
		s	6.232	6.163		
		n	8	8		
	5.5	$\overline{X}$	102	109		
		s	2.588	4.658		
		n	8	8		
	4.2	$\overline{X}$	117	124	112	111
		s	2.976	4.899	1.642	3.137
		n	8	8	8	8

LH			'V	∘V	∘V	'V
JB	2.1	$\overline{X}$	19.8	23.8		
		s	1.035	1.669		
		n	8	8		
	2.2	$\overline{X}$	18.6	21.3		
		s	0.774	1.581		
		n	8	8		
	4.1	$\overline{X}$	20.8	25.3	22.5	
		s	1.276	1.571	1.592	
		n	16	16	16	
	4.4	$\overline{X}$	20.4	21.1		
		s	1.360	0.771		
		n	16	16		
	5.1	$\overline{X}$	20.4	25.1		
		s	1.923	1.727		
		n	8	8		
	5.5	$\overline{X}$	19.9	21.1		
		s	0.991	0.991		
		n	8	8		
	4.2	$\overline{X}$	21.1	25.1	21.8	20.0
		s	0.835	1.356	1.165	0.956
		n	8	8	8	8

Fo			'V	∘V	∘V	'V
JR	2.1	$\bar{X}$	109	138		
		s	5.148	7.846		
		n	8	8		
	2.2	$\bar{X}$	95	117		
		s	4.534	6.906		
		n	8	8		
	4.1	$\bar{X}$	118	146	117	
		s	5.899	7.949	6.407	
		n	16	16	16	
	4.4	$\bar{X}$	91	107		
		s	2.750	5.013		
		n	16	16		
	5.1	$\bar{X}$	113	142		
		s	4.342	6.266		
		n	8	8		
	5.5	$\bar{X}$	93	114		
		s	4.375	4.224		
		n	8	8		
	4.2	$\bar{X}$	110	143	106	101
		s	3.871	6.198	3.071	3.068
		n	8	8	8	8

LH			'V	∘V	∘V	'V
JR	2.1	$\bar{X}$	23.5	24.1		
		s	1.195	0.835		
		n	8	8		
	2.2	$\bar{X}$	22.9	23.4		
		s	0.835	1.188		
		n	8	8		
	4.1	$\bar{X}$	22.9	24.7	23.3	
		s	1.928	1.852	1.447	
		n	16	16	16	
	4.4	$\bar{X}$	22.6	24.0		
		s	1.204	0.966		
		n	16	16		
	5.1	$\bar{X}$	23.0	24.6		
		s	1.414	0.744		
		n	8	8		
	5.5	$\bar{X}$	22.6	23.5		
		s	1.061	0.926		
		n	8	8		
	4.2	$\bar{X}$	24.5	24.6	23.8	22.0
		s	1.069	1.506	1.165	0.756
		n	8	8	8	8

Fo			'V	∘V	∘V	'V
ND	2.1	$\overline{X}$	146	154		
		s	7.764	9.568		
		n	8	8		
	2.2	$\overline{X}$	110	111		
		s	5.793	6.341		
		n	8	8		
	4.1	$\overline{X}$	148	158	142	
		s	7.898	8.229	10.353	
		n	14	14	14	
	4.4	$\overline{X}$	102	105		
		s	4.266	4.334		
		n	16	16		
	5.1	$\overline{X}$	155	163		
		s	7.254	5.908		
		n	7	7		
	5.5	$\overline{X}$	103	106		
		s	3.105	5.139		
		n	8	8		
	4.2	$\overline{X}$	132	146	129	125
		s	7.833	11.250	9.418	8.812
		n	8	8	8	8

LH			'V	∘V	∘V	'V
ND	2.1	$\overline{X}$	20.8	12.8		
		s	4.097	2.503		
		n	8	8		
	2.2	$\overline{X}$	17.1	12.0		
		s	1.727	2.507		
		n	8	8		
	4.1	$\overline{X}$	20.9	15.7	14.7	
		s	2.119	2.277	2.440	
		n	15	15	15	
	4.4	$\overline{X}$	17.1	13.6		
		s	1.408	1.454		
		n	15	15		
	5.1	$\overline{X}$	23.1	16.6		
		s	4.811	2.828		
		n	7	7		
	5.5	$\overline{X}$	14.6	11.5		
		s	1.302	1.195		
		n	8	8		
	4.2	$\overline{X}$	19.3	14.1	11.9	18.0
		s	2.984	1.773	1.113	2.380
		n	7	7	7	7

Fo			'V	∘V	∘V	'V
NR	2.1	$\bar{X}$	115	143		
		s	7.434	3.742		
		n	8	8		
	2.2	$\bar{X}$	85	97		
		s	2.964	3.536		
		n	8	8		
	4.1	$\bar{X}$	112	146	126	
		s	5.095	7.881	5.196	
		n	16	16	16	
	4.4	$\bar{X}$	82	91		
		s	3.633	1.962		
		n	16	16		
	5.1	$\bar{X}$	119	151		
		s	2.900	3.852		
		n	8	8		
	5.5	$\bar{X}$	82	90		
		s	3.615	2.167		
		n	8	8		
	4.2	$\bar{X}$	97	121	114	92
		s	4.504	3.335	6.563	2.875
		n	8	8	8	8

LH			'V	∘V	∘V	'V
NR	2.1	$\bar{X}$	22.6	21.9		
		s	0.518	1.126		
		n	8	8		
	2.2	$\bar{X}$	19.3	18.6		
		s	0.886	1.302		
		n	8	8		
	4.1	$\bar{X}$	22.8	22.6	21.8	
		s	0.750	1.030	0.856	
		n	16	16	16	
	4.4	$\bar{X}$	18.6	16.6		
		s	1.628	1.210		
		n	16	16		
	5.1	$\bar{X}$	22.4	21.5		
		s	0.744	0.926		
		n	8	8		
	5.5	$\bar{X}$	19.4	17.1		
		s	1.302	0.991		
		n	8	8		
	4.2	$\bar{X}$	21.9	20.1	20.5	19.1
		s	0.835	1.246	0.926	1.246
		n	8	8	8	8

Fo			'V	∘V	∘V	'ˋV
OB	2.1	$\overline{X}$	109	120		
		s	4.243	4.435		
		n	4	4		
	2.2	$\overline{X}$	97	91		
		s	8.435	6.081		
		n	8	8		
	4.1	$\overline{X}$	135	127	102	
		s	7.220	5.357	4.241	
		n	16	16	16	
	4.4	$\overline{X}$	99	86		
		s	4.993	4.367		
		n	16	16		
	5.1	$\overline{X}$	141	132		
		s	12.487	7.704		
		n	8	8		
	5.5	$\overline{X}$	93	88		
		s	2.964	4.359		
		n	8	7		
	4.2	$\overline{X}$	116	114	91	109
		s	7.690	7.323	4.685	5.159
		n	7	7	7	7

LH			'V	∘V	∘V	'V
OB	2.1	$\overline{X}$	28.3	32.3		
		s	3.500	2.062		
		n	4	4		
	2.2	$\overline{X}$	18.6	16.1		
		s	1.923	2.696		
		n	8	8		
	4.1	$\overline{X}$	27.4	32.4	24.1	
		s	3.240	2.128	3.052	
		n	16	16	16	
	4.4	$\overline{X}$	21.3	18.1		
		s	2.236	2.670		
		n	16	16		
	5.1	$\overline{X}$	30.8	33.6		
		s	1.488	2.234		
		n	8	8		
	5.5	$\overline{X}$	20.3	17.9		
		s	0.707	2.167		
		n	8	8		
	4.2	$\overline{X}$	25.1	29.3	20.3	23.4
		s	1.959	1.832	0.707	1.768
		n	8	8	8	8

Fo			'V	˳V	˳V	'V
PD	2.1	$\overline{X}$	109	119		
		s	2.673	4.830		
		n	8	7		
	2.2	$\overline{X}$	95	102		
		s	3.357	3.182		
		n	8	8		
	4.1	$\overline{X}$	116	131	125	
		s	3.864	5.241	7.070	
		n	16	16	16	
	4.4	$\overline{X}$	90	100		
		s	2.828	3.071		
		n	16	16		
	5.1	$\overline{X}$	119	131		
		s	5.436	6.022		
		n	8	8		
	5.5	$\overline{X}$	89	99		
		s	2.532	5.153		
		n	8	8		
	4.2	$\overline{X}$	112	124	119	102
		s	3.295	8.464	5.502	3.412
		n	8	8	8	8

LH			'V	˳V	˳V	'V
PD	2.1	$\overline{X}$	14.3	15.5		
		s	0.707	1.069		
		n	8	8		
	2.2	$\overline{X}$	12.3	11.5		
		s	1.581	1.414		
		n	8	8		
	4.1	$\overline{X}$	15.1	17.3	17.6	
		s	0.927	0.874	1.408	
		n	16	16	16	
	4.4	$\overline{X}$	14.6	13.6		
		s	0.886	1.360		
		n	16	16		
	5.1	$\overline{X}$	14.3	16.9		
		s	0.886	1.458		
		n	8	8		
	5.5	$\overline{X}$	13.6	12.0		
		s	1.408	1.309		
		n	8	8		
	4.2	$\overline{X}$	14.9	16.9	16.6	14.1
		s	0.835	0.835	1.061	0.835
		n	8	8	8	8

Fo			'V	∘V	∘V	'V
		$\bar{X}$	180	197		
	2.1	s	7.200	4.259		
		n	8	8		
		$\bar{X}$	156	177		
	2.2	s	8.450	7.633		
		n	8	8		
		$\bar{X}$	196	209	185	
	4.1	s	5.690	7.101	7.693	
		n	14	14	14	
		$\bar{X}$	147	168		
PH	4.4	s	4.879	4.863		
		n	14	14		
		$\bar{X}$	186	204		
	5.1	s	10.650	8.271		
		n	8	8		
		$\bar{X}$	143	163		
	5.5	s	6.289	5.445		
		n	8	8		
		$\bar{X}$	164	187	146	163
	4.2	s	6.813	6.698	7.230	7.308
		n	8	8	8	8

LH			'V	∘V	∘V	'V
		$\bar{X}$	21.6	23.1		
	2.1	s	0.744	0.835		
		n	8	8		
		$\bar{X}$	19.9	21.4		
	2.2	s	0.835	1.061		
		n	8	8		
		$\bar{X}$	21.9	24.3	23.1	
	4.1	s	1.232	1.031	1.207	
		n	14	14	14	
		$\bar{X}$	22.1	22.4		
PH	4.4	s	0.864	1.089		
		n	14	14		
		$\bar{X}$	22.5	24.4		
	5.1	s	0.926	1.061		
		n	8	8		
		$\bar{X}$	22.1	21.8		
	5.5	s	0.354	1.282		
		n	8	8		
		$\bar{X}$	21.0	22.0	18.6	20.3
	4.2	s	1.512	1.414	0.916	1.669
		n	8	8	8	8

Fo			'V	∘V	∘V	'V
PM	2.1	$\bar{X}$	99	123		
		s	3.117	5.222		
		n	8	8		
	2.2	$\bar{X}$	88	102		
		s	4.373	1.852		
		n	8	8		
	4.1	$\bar{X}$	107	138	106	
		s	3.557	4.266	3.575	
		n	16	16	16	
	4.4	$\bar{X}$	85	96		
		s	2.366	2.128		
		n	16	16		
	5.1	$\bar{X}$	103	138		
		s	3.012	7.111		
		n	8	8		
	5.5	$\bar{X}$	84	97		
		s	1.506	3.068		
		n	8	8		
	4.2	$\bar{X}$	98	119	96	93
		s	2.357	5.548	3.583	2.357
		n	8	8	8	8

LH			'V	∘V	∘V	'V
PM	2.1	$\bar{X}$	12.4	14.1		
		s	1.996	0.991		
		n	8	8		
	2.2	$\bar{X}$	9.0	11.6		
		s	0.926	0.518		
		n	8	8		
	4.1	$\bar{X}$	13.2	14.6	13.1	
		s	1.169	1.087	0.885	
		n	16	16	16	
	4.4	$\bar{X}$	10.0	12.9		
		s	0.895	0.680		
		n	16	16		
	5.1	$\bar{X}$	12.3	13.8		
		s	1.035	1.282		
		n	8	8		
	5.5	$\bar{X}$	9.5	12.0		
		s	0.756	0.926		
		n	8	8		
	4.2	$\bar{X}$	11.8	13.6	11.4	10.8
		s	0.886	0.916	1.302	1.389
		n	8	8	8	8

ARIPUC 19, 1985, p. 143-168

EMG INVESTIGATION OF LABIAL ARTICULATIONS

ELIZABETH ULDALL

Results of electromyography with hooked-wire electrodes in one subject in a large number of articulations involving the lips are presented. Stops, pulmonic and glottalic, oral and nasal; fricatives, approximants; rounded vowels; some labials produced by assimilation in English are discussed.

I. INTRODUCTION

In 1974 I was hospitably included in a large programme of electromyography being conducted at the Institute of Phonetics at the University of Copenhagen. Dr. Hajime Hirose placed hooked-wire electrodes in my upper lip, in Orbicularis Oris Superior (OOS) at the vermilion border about 2 cm from the center, in the lower lip in Orbicularis Oris Inferior (OOI) similarly, and in Depressor Labii Inferioris (DLI) about 2 cm below the vermilion border and 3 cm away from the center.

The material consisted mainly of nonsense words with labial consonants between [i-i] or [u-u] (for implosives and ejectives also [a-a]) with stress on the second syllable, e.g. [i'bi], and a number of real words, all said in the frame "Say ... again". The list was read ten times. The sounds occurring in American English were pronounced in my normal pronunciation, i.e. [p] was aspirated and [b] was voiced in medial position, while sounds not part of the English sound system were pronounced on the basis of my general phonetic training (UC London).

The EMG signals and the audio signals were recorded on a professional DC tape recorder. Mingograms containing raw and integrated EMG signals, duplex oscillograms and intensity curves were produced for inspection and for placing the line-up point for averaging. The EMG signals, particularly those of OOI, contained a number of (relatively weak) artefacts.

Filtering experiments showed that almost all artefacts could be removed by means of HP filtering at 390 Hz without distorting the relevant signal, which remained relatively unchanged up to HP 1500, although it was somewhat weakened at higher frequencies (cf. Rischel and Hutters 1980, p. 306-308).

After integration (25 ms) and HP filtering (390 Hz) the EMG signals were fed into a PDP-8 computer together with an intensity curve and an F_0 curve. After control of the print-outs of the sampled signals, the EMG curves were averaged. The Figures are somewhat reduced copies of these print-outs. The standard deviation in percent of the mean is shown below the EMG curves. In most cases the boundary between the first vowel (of the nonsense word) and the consonant was chosen as the line-up point for the averaging. This boundary was identified by the abrupt decrease in intensity.

The indication of the "text" below the individual curves in the figures is not in phonetic transcription but in an arbitrary orthography for the use of the computer, e.g. IPI= [iphi], IPHI= [iɸi]. The phonetic transcription which is given at the bottom of the pages is in IPA, with the additional diacritic [̢] below a symbol (β̢ ɹ̢) to indicate approximants.

II. COMPARISON OF [p b m p' ɓ]

All show some activity in DLI at the implosion. In the type V'CV this activity is, however, always less than the activity found at the release (see Figures 1, 2, 4, 5). In "capable" [keipəbəl] and "probable" [pɹɑbəbel] DLI has more activity at the implosion than at the release of medial, unstressed [p] and [b], respectively (see Figures 13, 14).

Activity in DLI at the point of closure was also found by Fischer-Jørgensen and Hirose (1974) for three out of five subjects, for one of them only before unrounded vowels. Hadding et al. (1969) say, "Hirano reports that contrary to his expectation from reading the literature, DLI may be involved in the lip-closing gesture" and "Later findings indicate that some speakers normally activate the depressor when articulating labial stops".

Action in OOS and OOI begins about 15 cs before the implosion for [p b m]. The duration of activity for these is presumably related to the speaker's language, American English. Action for [p'] and [ɓ] is rather short in the [a-a] context.

Relative amounts of activity (see Figures 1 and 2):

before i		before u	
OOS	m>b>p	OOS	m>b>p
OOI	m>p>b	OOI	m>b>p
DLI		DLI	
at implosion	p>m>b	at implosion	b>m>p
at release	m>b>p	at release	m>b>p

[m] clearly has higher activity than [p], and this is significant in OOS, and for OOI in the [u-u] context, and probably also in the other cases. With a few exceptions (OOI in the [i-i] context and DLI at the implosion) [m] has more activity than [b] and [b] more than [p], but these differences are rarely significant. These judgments are based partly on an inspection of the overlappings of the ten individual curves, and partly on the differences in millivolts compared to the standard deviation. The differences were measured in millimeters and converted to millivolts; the scale was not large enough for precise measurements.

One can conclude that in this speaker [m] always has relatively more activity than [b] and [p] (and the difference from [p] is almost always significant) and [b] tends to have higher activity than [p]. This is evidence against the assumption that the strength of the activity in the lips depends on the intra-oral pressure; [p] should require more effort to keep the lips closed against the higher air pressure behind the closure.

Öhman et al. (1966): "...the voiceless stop...has a higher intraoral pressure than the voiced cognate. This higher pressure must be counteracted by a more energetic implosion gesture. At the release, however, the separation of the lips is aided by the intraoral pressure so that in this phase of the consonant the muscular activity does not have to be as energetic in the voiceless as in the voiced case."

The results in the present material agree with the conclusions of Fischer-Jørgensen and Hirose (1974) and Tatham and Morton (1973) that the activity in the closing muscles is in no way related to the intra-oral air pressure. Malécot and Richman's observation (1974) is of interest here: "We now consider force of articulation to be based on a synesthetic interpretation of intra-buccal air-pressure."

The relations between [p] and [b] are of interest for the fortis-lenis question. Perrin and Sharf (1970) (quoted in Tatham and Morton 1973) "...obtained results indicating greater EMG peak amplitude for [p] than for [b]." Most investigators have found no consistent difference.

Lysaught et al. (1961): "The electromyographic measures showed no significant difference in timing or amount of muscle activity at the lips between /p/ and /b/."

Harris et al. (1965): "[There is a] slight average tendency for "tense" sounds to be produced more forcefully than "lax", but this tendency is present only in some subjects and when large numbers of responses are averaged." "Differences exist on the average, but they are small at best and non-existent for some subjects." "The differences observed [between /p/ and /b/] ...could not serve as a basis for a workable phonemic distinction based on muscular tension or laxness."

Fromkin (1966): "Since the articulation of /p/ does not in any way show consistently greater muscular activity of the orbicularis oris muscle than the articulation of /b/, it can be concluded that a feature other than the tense-lax feature must differentiate these two phonemes in American English."

Lubker and Parris (1970): "The labial gesture for the phone types /p/ and /b/ is essentially monotypic, requiring no more forceful labial contact or EMG activity for one than for the other."

Tatham and Morton (1973): "[The] difference between the peak amplitude of EMG from Orbicularis Oris associated with [p] and [b] is statistically insignificant at the 5% level."

Gatehouse and Kelman (1976): "[There is] no reliably consistent discrimination between /p/ and /b/."

Studies which distinguish between activity for implosion and for release have found no difference, or slightly greater activity for [b] at the release: Harris et al. (1965): "The difference between /p/ and /b/ is not significant for the lip-opening gesture."

Öhman (1968): "...the release command is usually stronger in /b/ than in /p/."

Fischer-Jørgensen and Hirose (1974): "As for the closing muscles, ...EFJ has a clear tendency to more activity for b. In ten of twelve pairs b has a higher maximum of activity in OOS, and in nine of twelve pairs in OOI. All exceptions are in unstressed position... Only three pairs in stressed position have a significant difference (5% level), but the number of pairs showing the same difference cannot be accidental. Moreover, the activity for b is normally of a slightly longer duration, which fits well with the longer closure time. But none of the other subjects has a clear difference between p and b. The number of individual means having higher activity for p or for b is about equal, and the differences are small." "On the whole, it must be said that the maxima are rather variable, but four of the six subjects have a clear tendency to stronger activity for b than for p at the release.

On all consonants with labial closure, OOS shows more activity than OOI. This might, of course, be due to accidents of the placement of the electrodes, but a comparison with the rounded vowels shows that in vowel rounding there is practically no difference between OOS and OOI. Thus OOS is relatively more important for closures.

This is also apparent from [u-u] items: OOS is always higher at the implosion of the stop than for the rounding of the preceding vowel, while for OOI there is hardly any difference. The following vowel [u] has lower activity than the consonant for both OOS and OOI, but this may be because the activity is high in the consonant, and keeping the rounding after a closure does not require so much activity. Cf. Leanderson et al. (1971), who found that "the degree of motor activity in a muscle used for the production of a certain speech sound was strongly dependent on how much that muscle was activated during the preceding sound."

[upu], [ubu] and [uwu] (Figures 2 and 10) show three peaks in OOS and OOI, the peak for the consonant being the highest in OOS, and the peak for the second [u] (though this is the stressed one) being considerably lower for both muscles than the first one.

If the sequences [ˈ--pəb--] of "capable" and [ˈ--bəb--] of "probable" (unstressed) are compared, we find that OOS and OOI have lower peaks for the second labial (Figures 13 and 14).

In [ip'i] and [iɓi], [p'] has higher peaks in all three muscles: in [ap'a] and [aɓa] the same relation holds, except that [p'] = [ɓ] in DLI (Figures 3 and 4).

I had thought that DLI might show more action in getting from lip-closure to an open vowel [a] than in getting to a close vowel [i], but the amount of action is the same in DLI in [ip'i] and [ap'a], and in [iɓi] and [aɓa]. Fischer-Jørgensen and Hirose (1974) found that "most of the subjects have a tendency to stronger activity of DLI before a than before i."

There is a great deal more activity in DLI before unrounded vowels, where DLI is involved in opening and spreading the lips, than before rounded vowels (the lips remain rounded during the closure) (cp. Figures 1 and 2). Fischer-Jørgensen and Hirose (1974) point out that in their Danish subjects DLI has very little activity (or none at all) for the opening of stops before [u], except for one subject out of five.

III. COMPARISON OF STOPS, FRICATIVES AND APPROXIMANTS

One thinks of fricatives as requiring a "nicer" adjustment of the articulators to arrange the position where friction will be produced, and hold it.

In Hardcastle (1976) we find a clear statement of this view: "In general terms, then, fricatives can be said to require more delicate neuro-muscular control than stops" ... "The categories stop and fricative for example can be differentiated with reference to the type of muscular activity employed in their articulation. For the production of a stop, a ballistic type of muscular contraction is necessary, i.e. one involving primarily protagonist muscles operating relatively independently of antagonist muscles." ... "A fricative, on the other hand, requires a far more delicate balance of protagonist and antagonist muscles to create the specific stricture required to maintain the turbulent flow of air necessary for its production.

In the sets [p ɸ] and [b β β̞] (Figures 5 and 6) there is less action in OOS and OOI for the fricatives than for the stops, and less still for the approximant [β̞]. It looks as though it is the activity for the distance over which the lips move that is being recorded, rather than tension. There is also much less action in the antagonist DLI for the fricatives and approximant than for the stops. The stops seem to have more controlled movement of protagonist and antagonist than the fricatives, but other muscles may also be involved, and the problem is very complicated.

The fricatives which have their main narrowing at the lips, [ɸ β], have more activity in OOS than in OOI, like the stops (Figures 5 and 6). In those fricatives which have their main place of articulation elsewhere but have rounding additionally [ʃ dʒ ɹ̝] (Figures 7, 8, 9), we find that between [i-i] OOI is appreciably higher than OOS, just the opposite of the case of the mainly labial sounds. In the [u-u] context, these fricatives often have a valley in OOS for the consonant, but not in OOI; that is, OOS is less active for consonant- than for vowel-rounding. In the [u-u] context, the approximant [w] has a peak in OOS in the consonant compared to the vowels (Figure 10), but this is not the case in [uɸu] (Figure 11), [uβu] and [uβ̞u]. The labio-dental fricatives, [f] (Figure 12) and [v] have, of course, more activity in the lower lip.

There are also differences in timing between the various types of consonant, which can be seen most clearly in the [i-i] context (see Figures 1 and 3). For consonants with bilabial closure the peak in OOS and OOI is approximately at the moment of implosion. The activity starts about 16-20 cs earlier, rather slowly at the beginning. Only [iɓi] starts later (OOS 7 cs, OOI 12 cs). The fall is abrupt and the zero line is generally reached about 8-9 cs after the peak.

Where lip activity is a secondary articulation (see Figures 7, 8, 9, [i-i] context), the activity starts earlier, 23-34 cs before the line-up point (which is the beginning of the abrupt fall in the intensity curve) and the peak is also much earlier, 10-20 cs before the line-up point. The rise and fall are rather symmetrical.

[iwi] (Figure 10) looks very much like the plosives, but in [iɸi], [iβi], [iβ̞i] (Figures 5 and 6) OOS, but not OOI, starts later, about 12-15 cs before the line-up point. In OOI the activity starts 18-20 cs before the line-up point, as for the plosives, but the peak is 8-9 cs earlier, and the fall continues until 10-20 cs after the line-up point. These consonants therefore have a more rapid rise than fall in the closing muscles.

IV. SOME LABIALS PRODUCED BY ASSIMILATION

In English (and Danish) [ə] is frequently represented by a syllabic consonant of the same place of articulation as a following consonant, particularly when the possible [ə] comes between two consonants of the same place of articulation, e.g. "back again", "all alone", "on a national basis", "rather than": [ˈbækg̩gɛn] [ˈɔl̩ˈloʊn] [ɒn̩ˈnæʃnəl ˈbeɪsɪs] [ˈrɑðð̩ən].

I was curious to know whether such syllabic consonants would be articulated with a separate burst of muscular energy. I looked at

ˈkeɪpəbəl ˈkeɪpb̩bəl

ˈprɑbəbəl ˈprɑbb̩bəl

and also cases with [m] and [v]. In the assimilated cases there are peaks in OOS and OOI for the implosion of the long labial articulations, and then a continuous very slight action until the release of the closure, but nothing that could be described as separate peaks for the syllabic consonants representing [ə] (see Figures 13 and 14).

V. VOWELS

Of the series <u>who'd</u>, <u>hood</u>, <u>hoed</u>, <u>hawed</u> [hɔɪd] and <u>how'd</u>, only <u>who'd</u> [hu:d], [hɔɪd] and <u>how'd</u> [haud] are presented (Figures 15, 16, 17). In all the rounded vowels the amount of activity is very similar in OOS and OOI. Fischer-Jørgensen and Hirose (1974), however, found that in their material "the dominating muscle for rounding seems to be OOI", and three out of four subjects hardly used OOS in rounding.

In the monophthongs, action in OOS and OOI decreases throughout the vowels and stops at the closure for [d]. For [oɷ] it increases in OOS and decreases in OOI. The increase is presumably because [oɷ] is a "closing" diphthong and the rounding increases as the vowel becomes closer. We transcribe this end of the diphthong with the symbol [ɷ], but this does not mean that it necessarily has the quality or the rounding of the monophthong [ɷ], but merely that the movement is towards a closer vowel. I can suggest no reason why the activity should <u>decrease</u> in OOI.

ACKNOWLEDGEMENTS

Eli Fischer-Jørgensen read an earlier draft of this paper and made many helpful suggestions for additions and the formulation of statements. She drew my attention to various points that should be discussed further. I am very grateful to her. Preben Dømler of the Institute staff was so kind as to give me a lot of help with the use of the computer in averaging the results of the ten repetitions.

REFERENCES

Fischer-Jørgensen, E. and Hirose, H. 1974: "A preliminary electromyographic study of labial and laryngeal muscles in Danish stop consonant production", *Status Rep. Speech Res., Haskins Labs. 39/40*, p. 231-254

Fromkin, V. 1965: "Some phonetic specifications of linguistic units: an electromyographic investigation", *Working Papers in Phonetics, UCLA 3*

Fromkin, V. 1966: "Neuromuscular specification of linguistic units", *Language and Speech 9*, p. 170-199

Gatehouse, S. and Kelman, A.W. 1976: "Comparison of the electromyographic activity of the muscle Orbicularis Oris using different electrode configurations", *Folia Phoniatrica 28, 1*, p. 52-62

Hadding, K., Hirano, M. and Smith, T. 1969: "Electromyographic study of lip activity in Swedish CV:C and CVC: syllables", *Working Papers 1, Phonetics Laboratory Lund University*, p. 1-8

Hadding, K., Hirose, H. and Harris, K. 1976: "Facial muscle activity in the production of Swedish vowels: an electromyographic study", *J. Phonetics 4*, p. 197-209

Hardcastle, W. 1976: *Physiology of Speech Production*, (Academic Press: Harcourt Brace Jovanovich)

Harris, K., Lysaught, G.F. and Schvey, M.M. 1965: "Some aspects of the production of oral and nasal labial stops", *Language and Speech 8*, p. 135-147

Leanderson, R., Persson, A. and Öhman, S. 1971: "Electromyographic studies of facial muscle activity in speech", *Acta Oto-Laryngologica 72*, p. 361-369

Lubker, J.F. and Parris, P. 1970: "Simultaneous measurements of intraoral pressure, force of labial contact, and labial electromyographic activity during the production of the stop consonant cognates /p/ and /b/", *J. Acoust. Soc. Am. 47*, p. 625-634

Lysaught, G., Rosov, R. and Harris, K. 1961: "Electromyography as a speech research technique with an application to labial stops", *J. Acoust. Soc. Am. 33*, p. 842

Malécot, A. and Richman, M. 1974: "A labiographic study of bilabial stops in English and French", *Studia Linguistica 38*, p. 100-108

Öhman, S. 1968: "Peripheral motor commands in labial articulation", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Institute of Techn. Stockholm 4/1967*, p. 30-63

Öhman, S., Leanderson, R. and Persson, A. 1965: "Electromyographic studies of facial muscles during speech", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn. Stockholm 3/1965*, p. 1-11

Öhman, S., Leanderson, R. and Persson, A. 1966: "EMG studies of facial muscle activity in speech II", *Speech Transm. Lab., Quart. Prog. and Status Rep., Royal Inst. Techn. Stockholm 1/1966*, p. 1-4

Perrin and Sharf 1970: "Differential electromyographic responses of bilabial stop consonants", *Speech and Hearing Science Research Reports, Univ. of Michigan 1* (quoted in Tatham and Morton 1973)

Rischel, J. and Hutters, B. 1980: "Filtering of EMG signals", *Ann. Rep. Inst. Phon., Univ. Copenhagen 14*, p. 285-316

Tatham, M. and Morton, K. 1969: "Some electromyography data towards a model of speech production", *Essex Occasional Papers, Language Centre 1*, p. 1-59

Tatham, M. and Morton, K. 1973: "Electromyographic and intraoral air-pressure studies of bilabial stops", *Language and Speech 16*, p. 336-350.

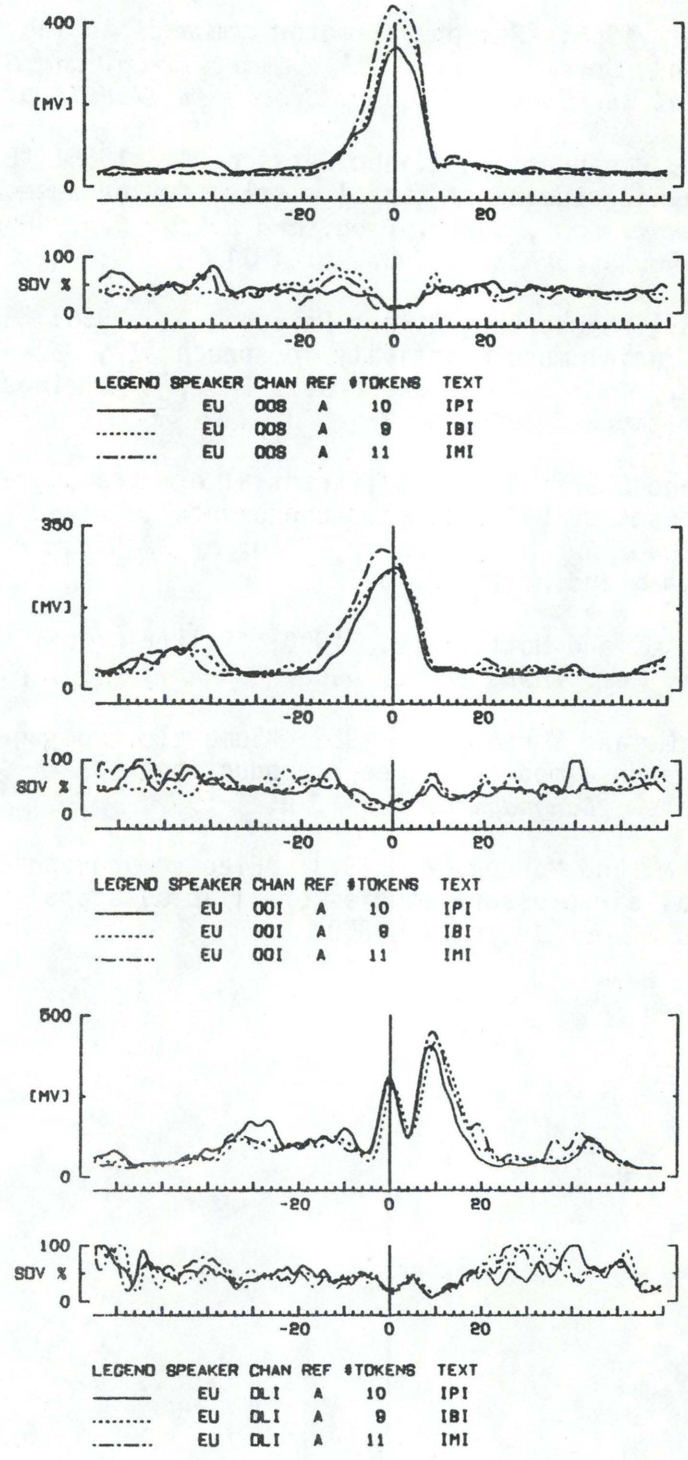


Figure 1

[iphi] ———, [ibi] ·····, [imi] -·-·-·-·

Average curves.

CHAN: *channel = muscle (OOS, OOI, DLI)*

REF: *line-up point = start of intervocalic consonant (valid for Figures 1-7 and 9-14).*

TEXT: *not phonetic transcription but an orthography for the use of the computer.*

SDV %: *standard deviation in percent of mean.*

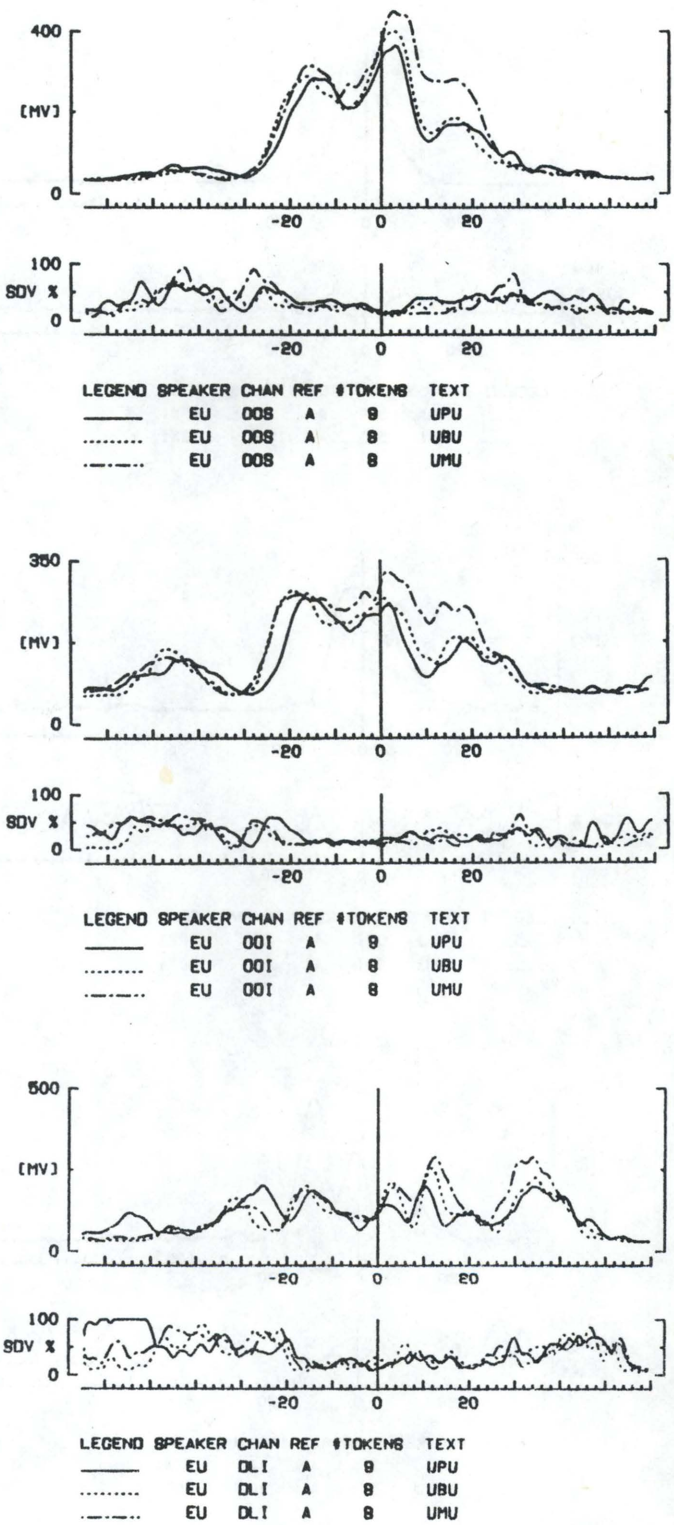


Figure 2

[uphu] ——— , [ubu] ····· , [umu] -·-·-·-

(see further legend to Figure 1)

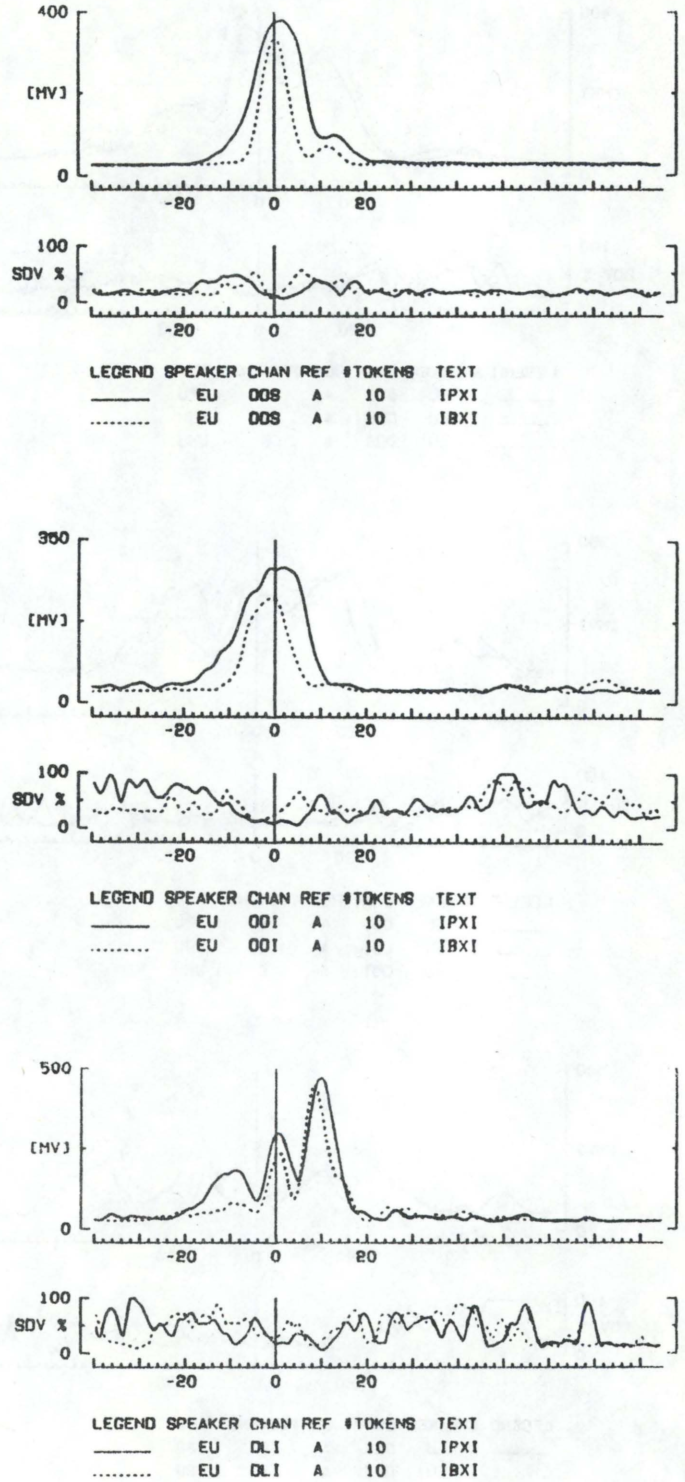


Figure 3

[ipˀi] ——— , [iɓi] .....

(see further legend to Figure 1)

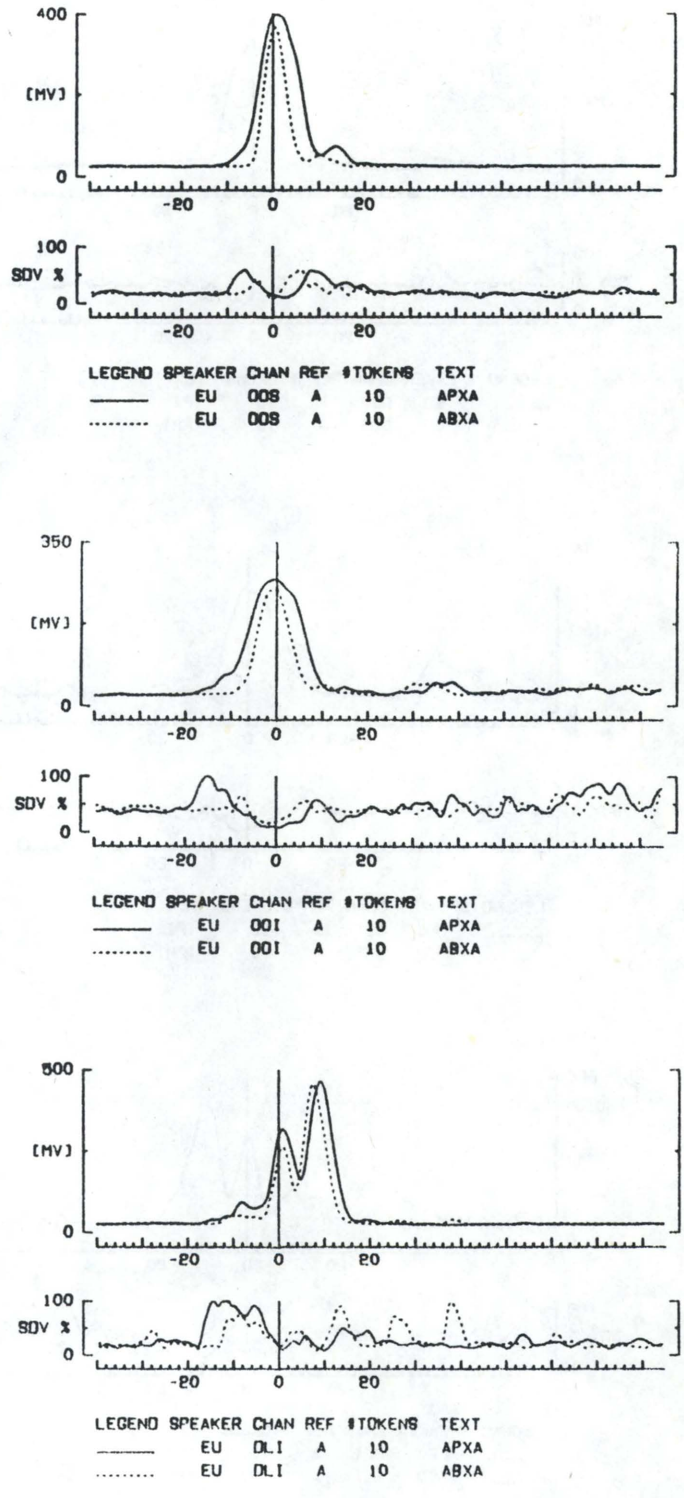


Figure 4

[ap'a] ——— , [aɓa] .....

(see further legend to Figure 1)

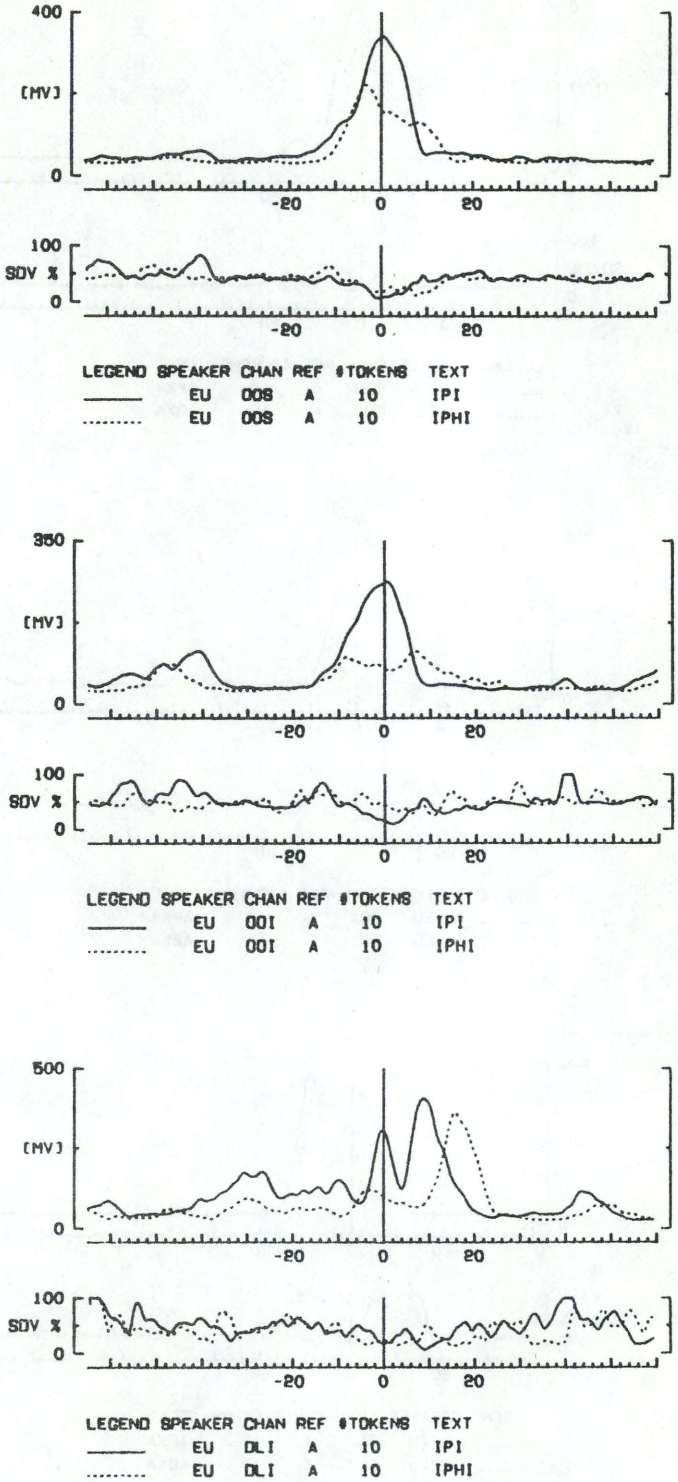


Figure 5

[iphi] ——— , [iɸi] ·····

(see further legend to Figure 1)

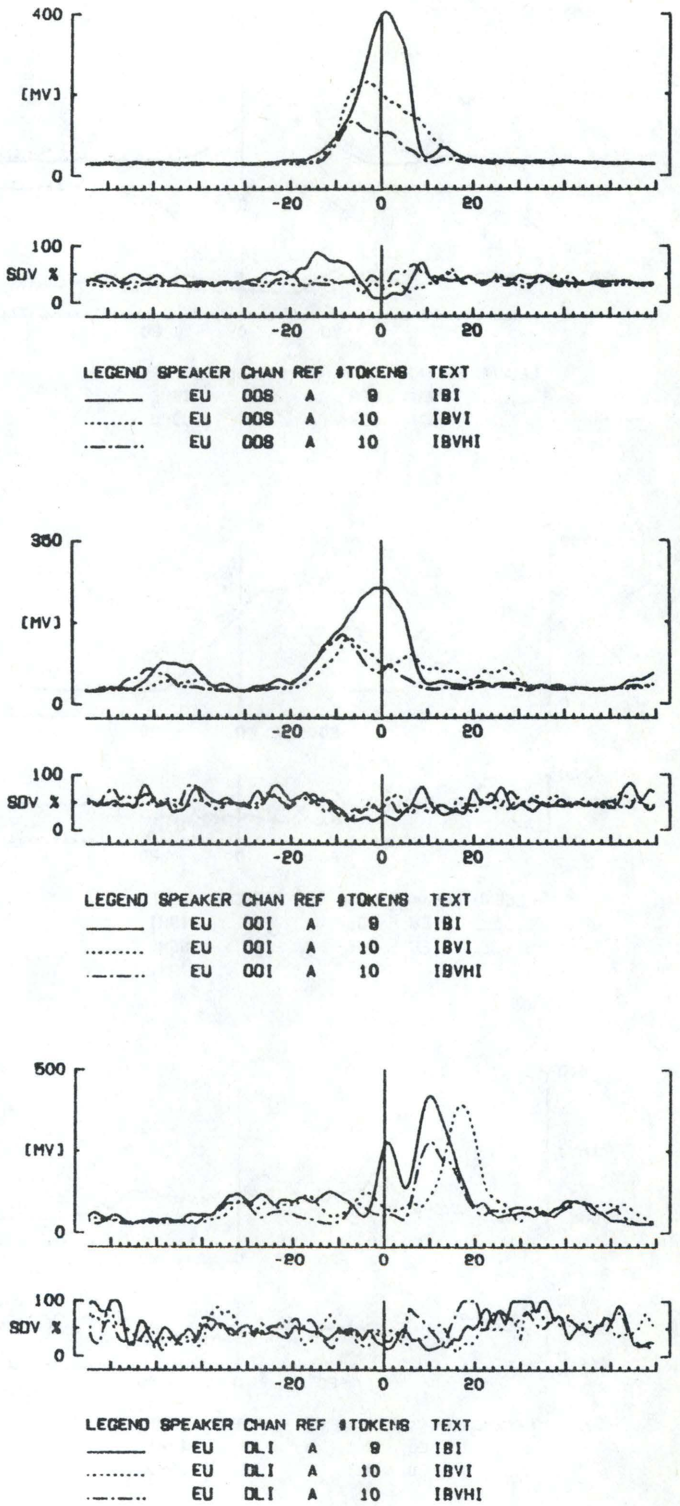


Figure 6

[ibi] ——— , [iβi] ····· , [iβ̞i] -------

(see further legend to Figure 1)

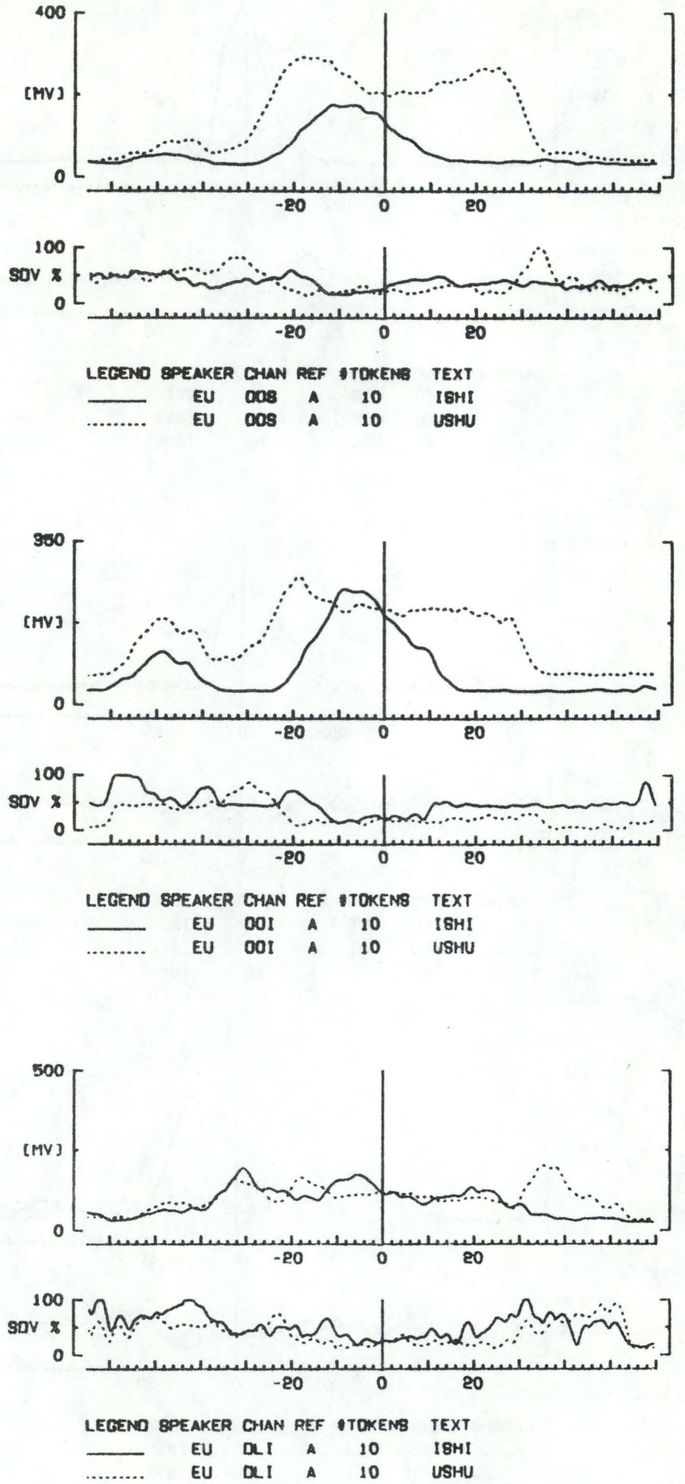


Figure 7

[iʃi] ——— , [uʃu] ·····

(see further legend to Figure 1)

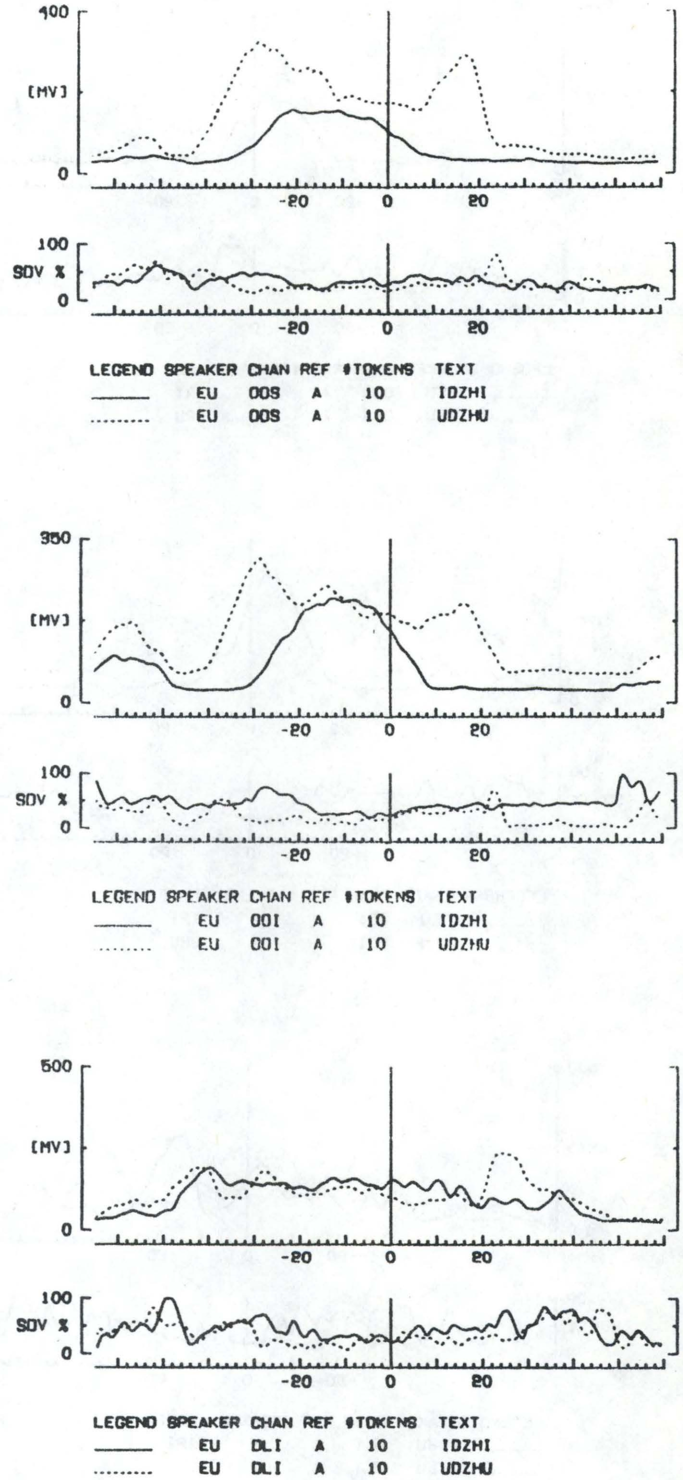


Figure 8

[idʒi] ——— , [udʒu] .....

line-up point: release of [d]
(see further legend to Figure 1)

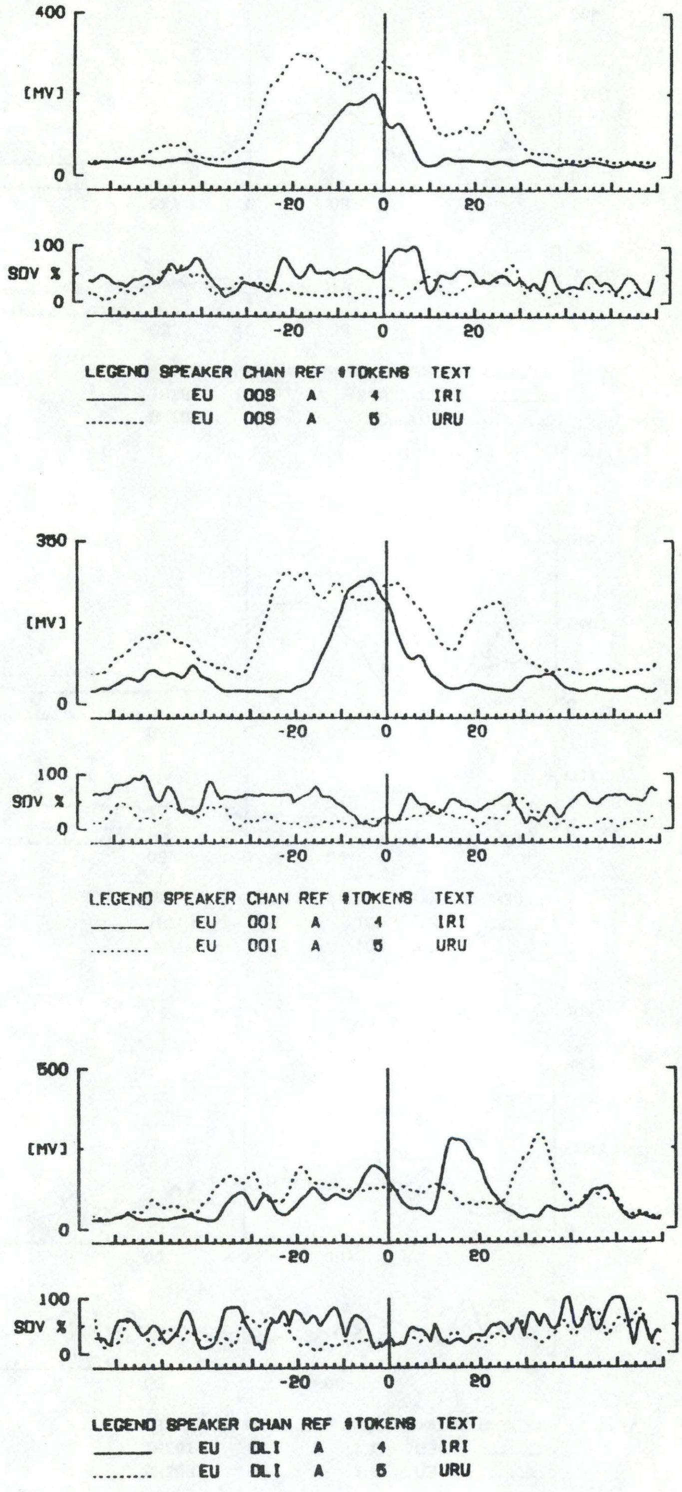


Figure 9

[iɽi] ——— , [uɽu] ·····

(see further legend to Figure 1)

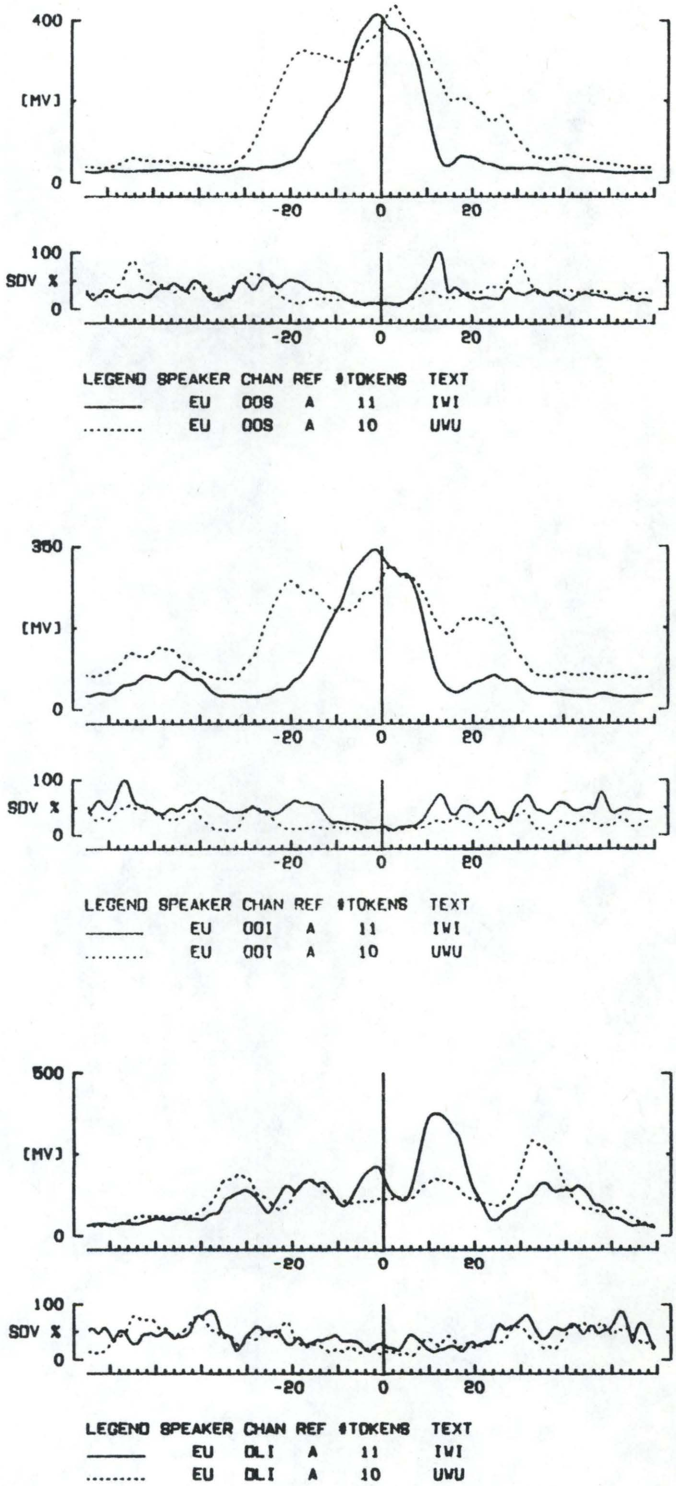


Figure 10

[iwi] ——— , [uwu] .....

(see further legend to Figure 1)

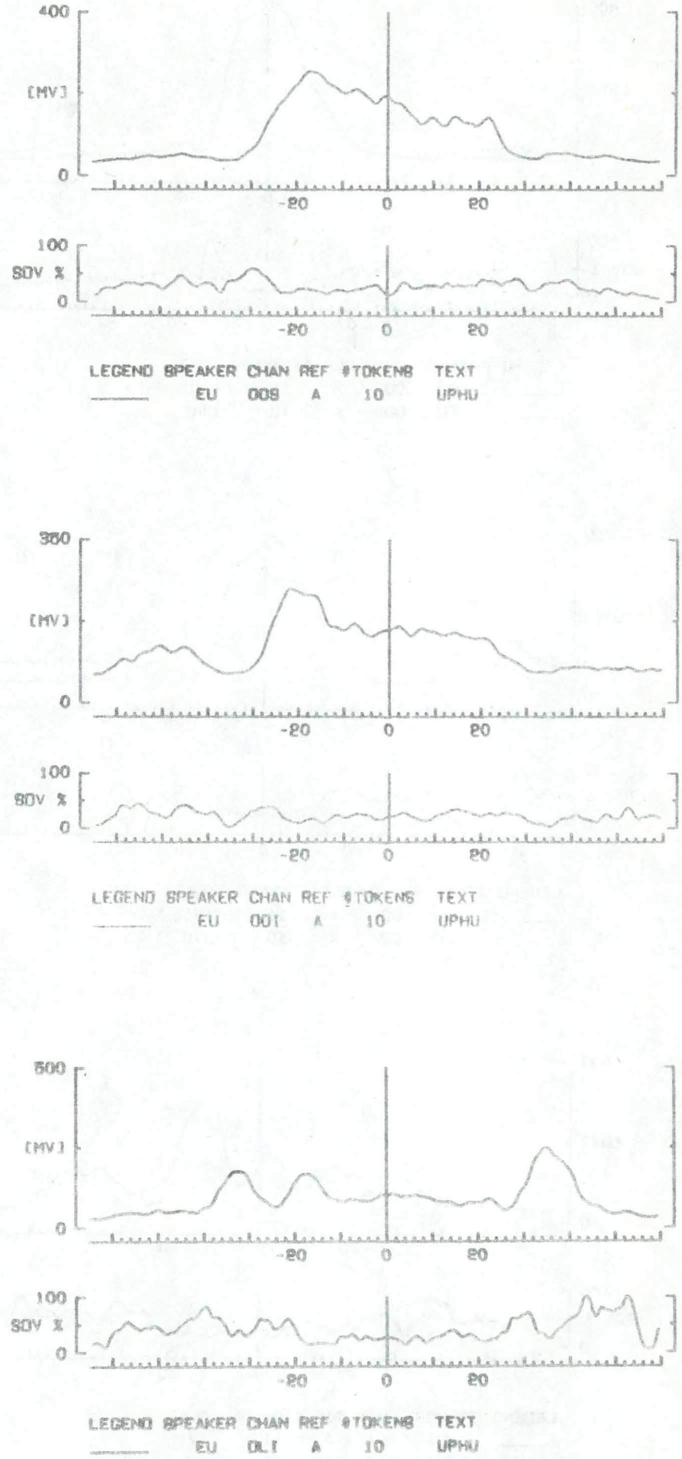


Figure 11

[uɸu]

(see further legend to Figure 1)

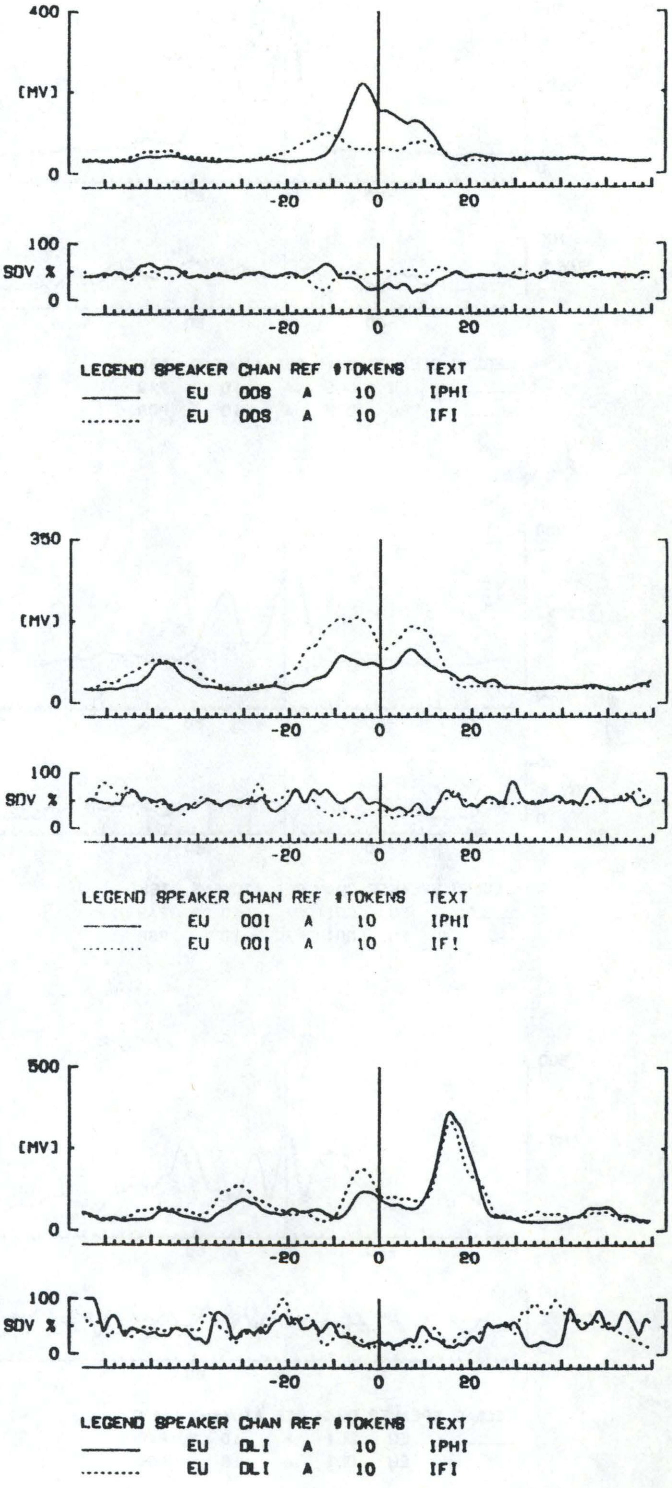


Figure 12

[iΦi] ——— , [ifi] ·····

(see further legend to Figure 1)

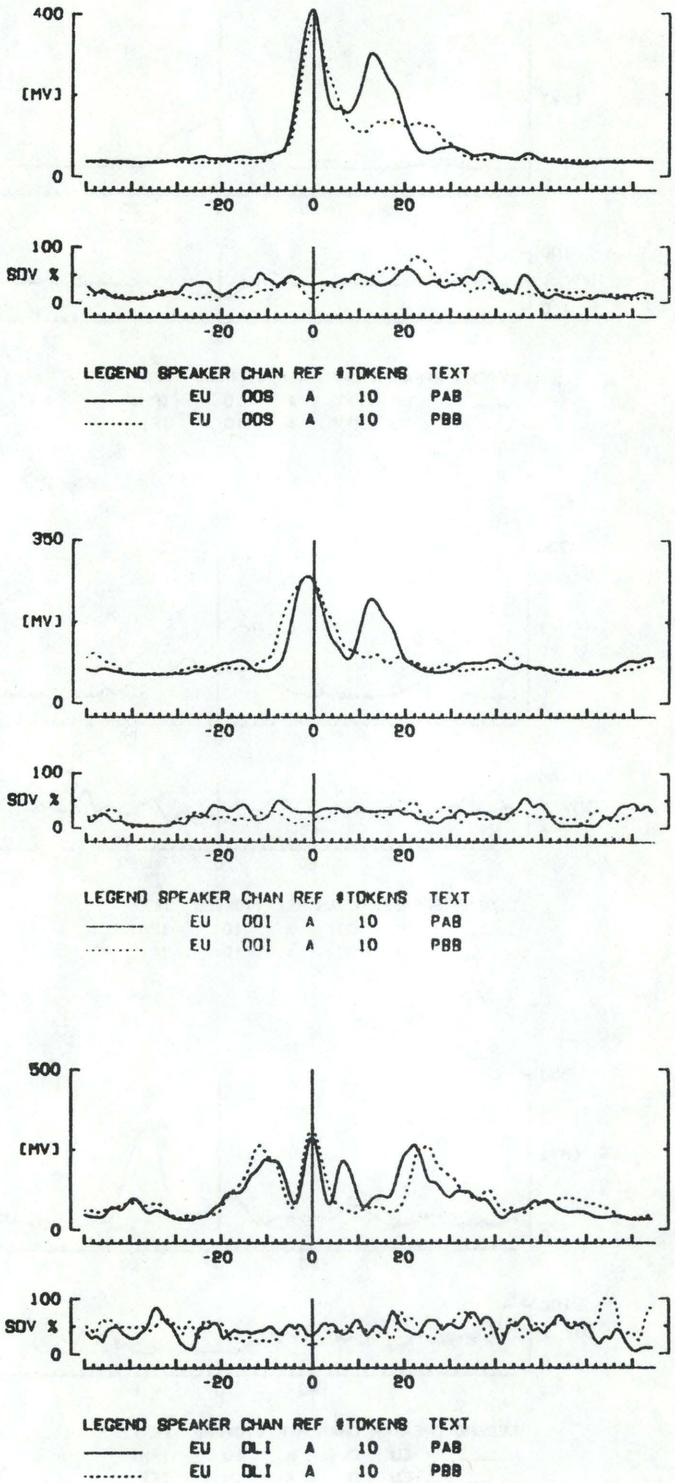


Figure 13

[keipəbəl] ——— , [keipb̥bəl] ·····

(see further legend to Figure 1)

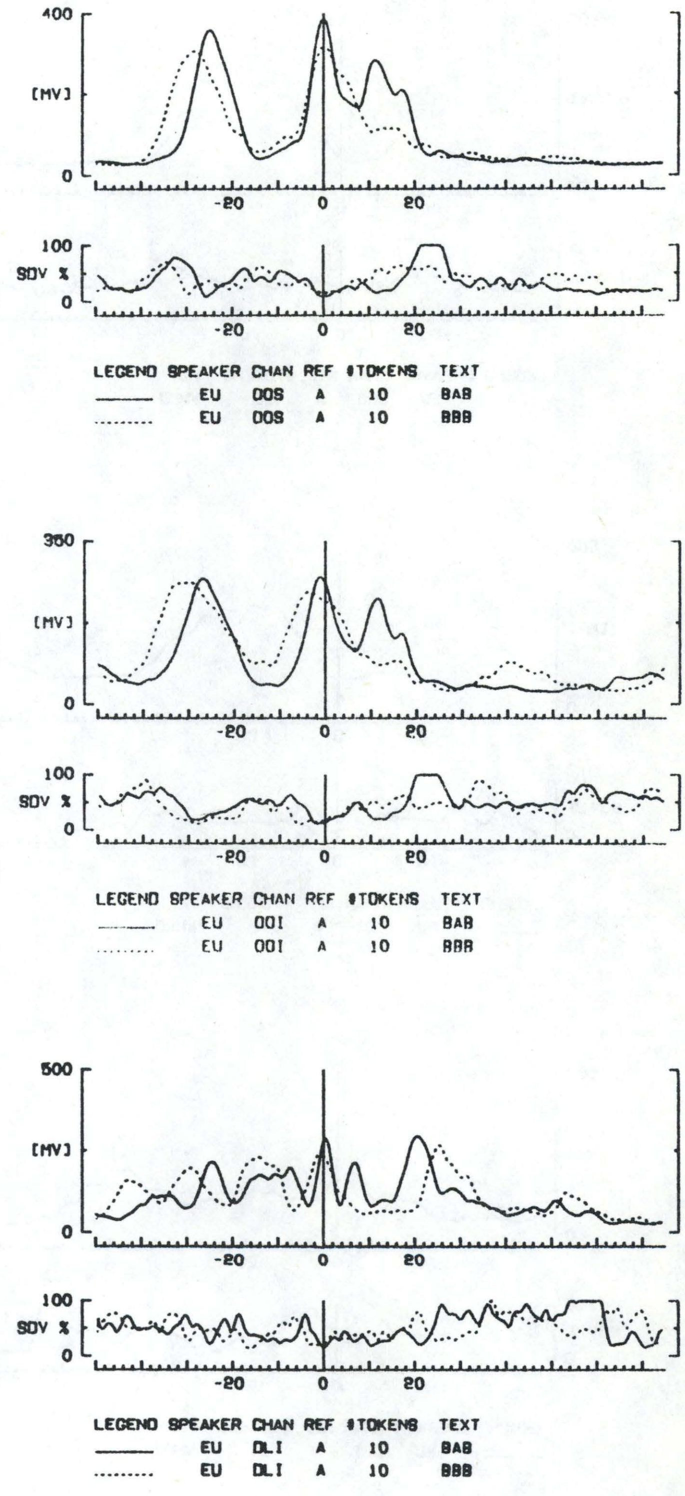


Figure 14

[prɑbəbəl] ——— , [prɑbb̩l] ·····

(see further legend to Figure 1)

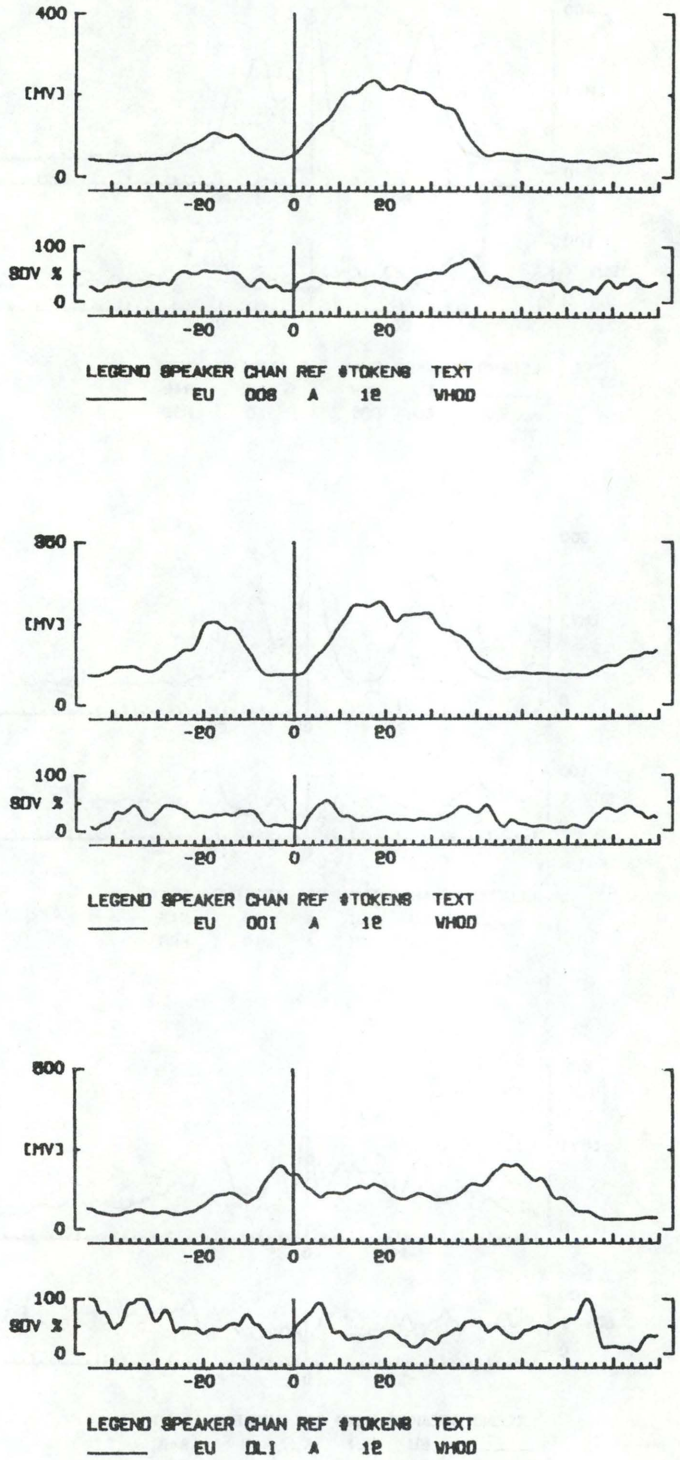


Figure 15

[hu:d]

line-up point boundary between [s] *and* [ei] *in the frame (see further legend to Figure 1)*

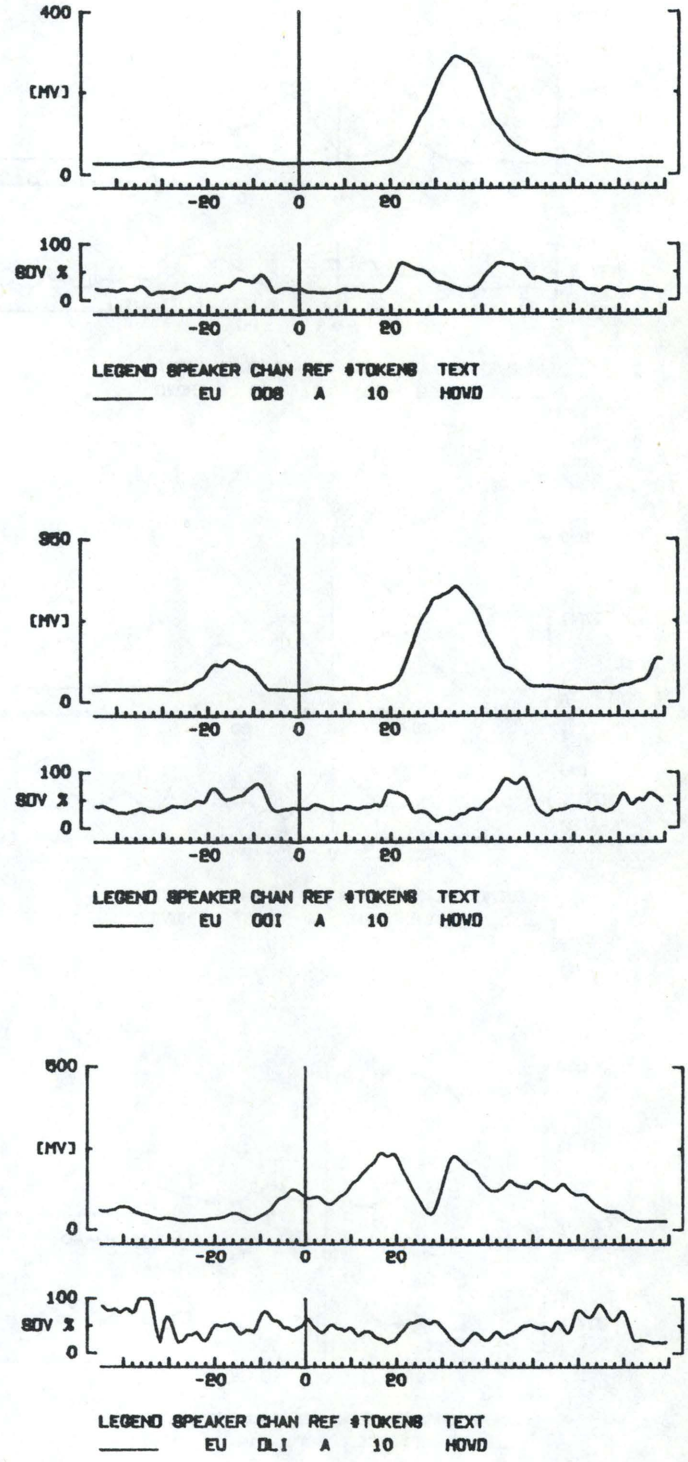


Figure 16

[haud]

line-up point as in Figure 15 (see further legend to Figure 1)

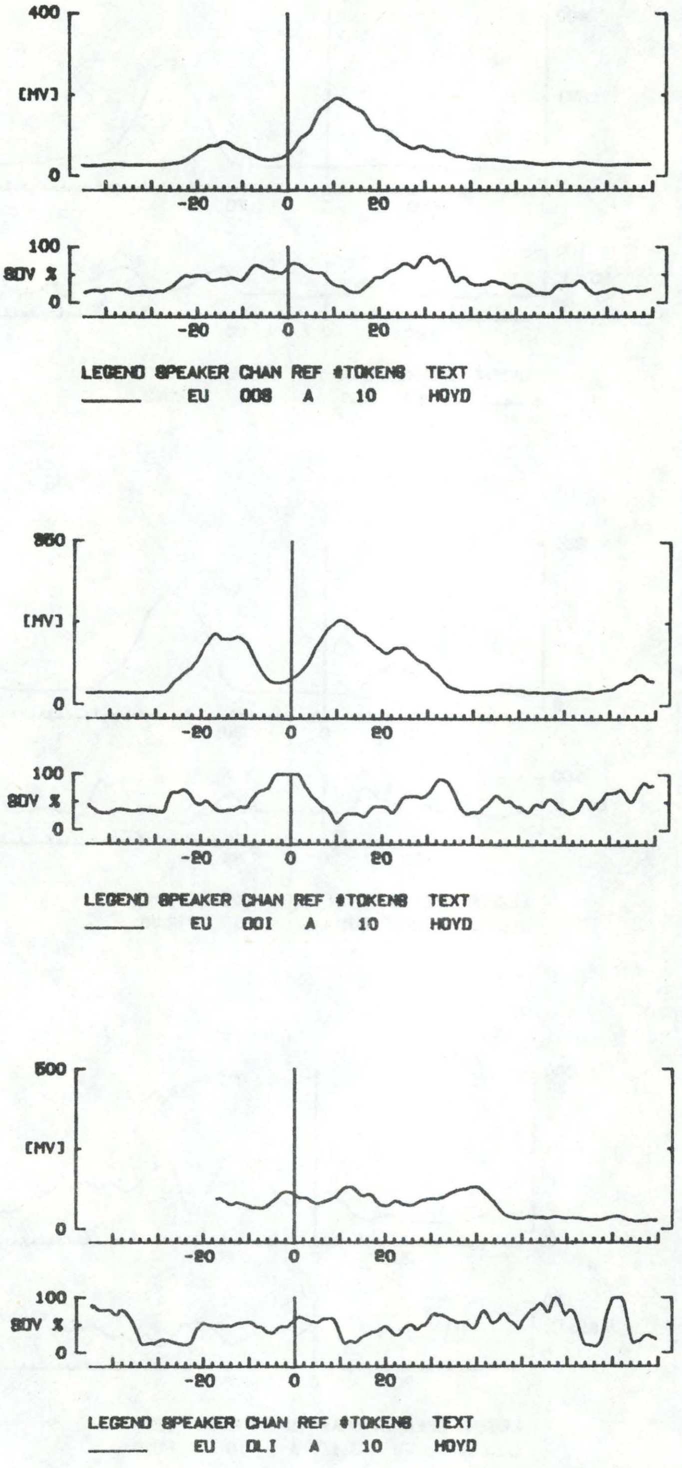


Figure 17

[hɔid]

line-up point as in Figure 15 (see further legend to Figure 1)

INSTITUTE OF PHONETICS

JULY 1, 1983 - DECEMBER 31, 1984

I. PERSONNEL OF THE INSTITUTE

PROFESSOR:
Jørgen Rischel, dr.phil.

ASSOCIATE PROFESSORS:
Børge Frøkjær-Jensen, cand.mag. (seconded to the Audiologopedic Centre)
Peter Holtse, cand.phil. (on leave from April 1, 1984)
Birgit Hutters, cand.mag.
Niels Reinholt Petersen, cand.phil.
Nina Grønnum Thorsen, lic.phil.
Oluf Thorsen, cand.mag.

RESEARCH FELLOWS:
Peter Molbæk Hansen, cand.mag.
Peter Holtse, cand.phil. (from April 1, 1984)

TEACHING ASSISTANTS:
Peter Molbæk Hansen, cand.mag.
Jørn Westengaard Holm, cand.mag.

ENGINEERS:
Otto Bertelsen, M.Sc.
Preben Dømler, B.Sc.

TECHNICIAN:
Svend-Erik Lystlund

SECRETARY:
Else Parkmann

GUEST RESEARCHERS:
Eli Fischer-Jørgensen, dr.phil.h.c.
Kiros Fre Woldu, Uppsala University (January - July 1984)
Elizabeth Uldall, Ph.d. (spring 1984)

II. PUBLICATIONS BY STAFF MEMBERS AND GUESTS

Eli Fischer-Jørgensen: "Vowel features and their explanatory power in phonology", Abstracts of the Xth Int. Congr. Phon. Sc. Symposium 5, Utrecht 1983, p. 259-265

Eli Fischer-Jørgensen: "The acoustic manifestation of Danish stress with particular reference to the reduction of stress in compounds", Studies for Antonie Cohen, Sound Structures (eds.: v.d. Broecke, van Heuven, and Zonneveld), Dordrecht (Foris) 1983, p. 81-105

Eli Fischer-Jørgensen: "Vokalforlængelse i danske tostavelsesord", From Sounds to Words, Essays in Honor of Claes-Christian Elert (eds.: Dahlstedt, Hansson, Hedquist, and Lindblom), Acta Universitatis Umensis, Umeå Studies in the Humanities 60, 1983, p. 87-106

Eli Fischer-Jørgensen: "Some aspects of the 'phonetic sciences', past and present", Opening Address, Xth Int. Congr. Phon. Sc., Utrecht 1983, p. 285-291

Eli Fischer-Jørgensen: "Hans Jørgen Uldall", Dansk Biografisk Leksikon 15, 1984, p. 134

Eli Fischer-Jørgensen and Birgit Hutters: "Aspirated stop consonants before low vowels. A problem of delimitation - its causes and consequences", Topics in Linguistic Phonetics in Honour of E.T. Uldall (eds.: J.A.W. Higgs and R. Thelwall), Occasional Papers in Linguistics and Language Learning 9, The New University of Ulster, 1984, p. 159-194

Jørgen Rischel: "On unit accentuation in Danish - and the distinction between deep and surface phonology", Folia Linguistica XVII, 1983, p. 51-97

Jørgen Rischel: "The abstractness paradox in Hjelmslevian linguistics", Proc. XIIIth Int. Congr. Ling., 1983, p. 884-887

Jørgen Rischel: "Om den fonematiske og morfofonematiske funksjonen til dei såkalla ordtonane i norsk" & "Morphemic tones and word tone in Eastern Norwegian", both reprinted (the former in translation by Katrin Lunde) in Prosodi/Prosody, ed. Jahr & Lorentz, 1983, p. 256-265 & 266-276

Jørgen Rischel: "Devoicing or strengthening of long obstruents in Greenlandic", Riepmočála, Festschrift Knut Bergsland, ed. Brendemoen, Hovdhaugen, and Magga, 1984, p. 122-135

Jørgen Rischel: "Mlabri: skovmenneskene i Nordthailand og deres sprog", Carlsbergfondet/Frederiksborgmuseet/Ny Carlsbergfondet, Årsskrift 1984, p. 19-23

Nina Thorsen: "Standard Danish sentence intonation - Phonetic data and their representation", Folia Linguistica 17, 1983, p. 187-220

Nina Thorsen: "F_0 timing in Danish word perception", Phonetica 41, 1984, p. 17-30

Nina Thorsen: "Variability and invariance in Danish stress group patterns", Phonetica 41, 1984, p. 88-102

Nina Thorsen: "The tonal manifestation of Danish words containing assimilated or elided schwa", in Nordic Prosody III (eds.: Claes-Christian Elert, Iréne Johansson, and Eva Strangert), Acta Universitatis Umensis, 1984, p. 215-230

Nina Thorsen and Oluf Thorsen: Fonetik for Sprogstuderende, 3rd revised edition, 6th printing, Copenhagen 1984. 170 pp.

III. GUEST LECTURES AND SEMINARS

September 21, 1983 - Dr. Sidney Wood, Lund University: "Vowel production".

October 5, 1983 - Professor John Clark (Macquarie University): "The problem of perceptual equivalence in synthesized speech".

November 30, 1983 - Dr. Robert Bannert (Lund University): "En model för tysk intonation" ("A model of German intonation").

December 7, 1983 - Dr. Gösta Bruce (Lund University): "Den svenska rytmens fonologi och fonetik" ("Phonology and phonetics of rhythm in Swedish").

January 18, 1984 - Professor Jørgen Rischel: A survey lecture on autosegmental phonology.

February 15, 1984 - Dr. Olle Engstrand (Uppsala University): "Spektrala och artikulatoriska korrelat till tense, lax och stress i svenska vokaler" ("Spectral and articulatory correlates to tense, lax, and stress in Swedish vowels").

February 17, 1984 - Kiros Fre Woldu: "The production and perception of ejective consonants".

March 14, 1984 - Dr. Jerzy Rubach (Warszawa): "The problem of abstractness in cyclic phonology".

March 28, 1984 - Dr. Francisco Lacerda (Stockholm): "How are CV and VC syllables represented in the auditory system?".

May 9, 1984 - Dr. Eva Strangert (Umeå University): "Rytmiska mönster i talad svenska" ("Rhythmical patterns in spoken Swedish").

May 30, 1984 - Dr. Karin Holmgren (Stockholm University): "Rapport från projektet 'Beskrivning av barns tidiga vokalisation" ("Report on the project 'Description of young children's vocalization").

September 26, 1984 - Dr. John Ryalls (Max Planck Institute, Nijmegen): "An acoustic investigation of vowel production in aphasia".

IV. PARTICIPATION IN CONGRESSES ETC.

Birgit Hutters, Nina Thorsen, and Oluf Thorsen participated in the Tenth International Congress of Phonetic Sciences, Utrecht, August 1-8, 1983.

Birgit Hutters participated in a symposium at Universität Kiel, April 5-6, 1984, and gave a paper: "Vocal fold adjustments in unvoiced obstruents - with special reference to aspiration and devoicing".

Jørgen Rischel participated in the International Conference on Thai Studies in Bangkok, 22-24 August 1984, and gave a paper on "Achievements and challenges in Thai phonetics".

Jørgen Rischel acted as a consultant on instrumental phonetics to the Departments of Linguistics of two universities in Thailand (Chulalongkorn, Mahidol) in December 1984.

Nina Thorsen participated in the Fifth International Phonology Meeting, Eisenstadt, Austria, June 25-28, 1984, and gave a paper: "Intonation and text in Standard Danish - with special reference to the abstract representation of intonation".

Nina Thorsen gave the following guest lectures: "Textual prosody", at the Cambridge Linguistic Society, Cambridge University, January 26, 1984; "Fo timing in Danish word perception", at Sussex University, January 31, and at Oxford University, February 1, 1984; "Text and intonation in Standard Danish" at Universität Kiel, April 6, 1984; and "Sentence intonation and its abstract representation - a discussion of two different models" at Stockholm University, November 22, 1984.

V. INSTRUMENTAL EQUIPMENT OF THE LABORATORY

The following is a list of instruments that have been purchased or built during the period July 1, 1983 - December 31, 1984.

INSTRUMENTATION FOR SPEECH ANALYSIS

1 pressure transducer, Gaeltec, type 8T-2.
1 Sonagraph, Kay Elemetrics, type 7800.

INSTRUMENTATION FOR SPEECH SYNTHESIS

1 microprocessor-controlled speech synthesizer, type PD.

TAPE RECORDERS

2 semi-professional recorders, Revox, type B77
1 cassette recorder, Tandberg, type TAT 771
1 cassette recorder, SONY, type TC D5M.

LOUDSPEAKERS

4 headphones, Sennheiser, type HD 430.

EQUIPMENT FOR EDP

1 CTR terminal, Televideo, type 925
1 CTR terminal, Pericom, type 7800.

GENERAL PURPOSE ELECTRONIC INSTRUMENTATION

1 frequency counter, Philips, type PM 6667
1 oscillator, Bang & Olufsen, type TG 8
1 V 24 tester, Trend, type BOB 3.

MICROPHONES

1 microphone, Sennheiser, type MKE 2002