ARIPUC VOL. 20

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1986

ARIPUC VOL. 20

ISSN 0589-6681

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

96, Njalsgade
DK 2300 Copenhagen

COPENHAGEN 1986

Reproafdelingen
Det humanistiske Fakultet
Københavns Universitet

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

CONTENTS

ARIPUC 20, 1986, p. 1-33

CONSONANTS IN SPONTANEOUS SPEECH OF FIVE DANISH CHILDREN WITH OPERATED CLEFT PALATE - AN IMPRESSIONISTIC ANALYSIS

BIRGIT HUTTERS
AND
KIRSTEN BRØNDSTED*

This paper examines consonants in spontaneous speech of five 4-5 year old Danish children with persisting speech problems after primary cleft palate surgery. Based on phonetic transcriptions from video-recordings the phonetic and phonological behaviour is described for each child and related to the target norm. The findings are discussed in the light of strategies in cleft palate speech. Furthermore, some aspects of universal and language dependent characteristics in cleft palate speech are treated.

I. INTRODUCTION

Both from clinical experience and from the literature, "cleft palate speech" appears to be clearly different from other forms of pathological speech (see, e.g., Bzoch et al. 1984, Edwards 1980, Morley 1970), even though the consequences of congenital cleft palate for communication are so varied, that Morris (1979) states that "it is difficult if not downright impossible, to describe meaningfully "cleft palate speech"" (p. 193).

Congenital cleft palate may, even when surgically repaired, impair the velopharyngeal function, but also the function of the tongue tip and blade which - for reasons as yet only partly explainable - may have reduced sensibility (Edwards 1980) and dyscoordinated mobility (Fletcher 1978, Shelton 1979).

---

*) Institute of Speech Disorders, Hellerup.

Dental and occlusal anatomical abnormalities also represent hazards to speech production for the cleft palate population, but in spite of comprehensive research on this subject, no clear conclusions as to cause and effect have been drawn (Starr 1979).

Fluctuating conductive hearing-impairment should also be mentioned as a probable cause of deviating sound production. Finally, it has been shown in several studies that congenital cleft palate may be accompanied by delayed language development (see, e.g., Fox et al. 1978, Clifford and Clifford 1979, Pannbacker 1975). The characteristics are mainly expressive in character and may include all aspects of language.

To avoid misunderstandings in the following, it should be stressed that in describing the consequences of cleft palate for communication, we are talking about conditions after primary surgical repair, that is after the closure of the cleft. Most of the children have no communication problems after primary repair, but some children go on having more or less serious speech and language disorders.

II. BACKGROUND AND PURPOSE OF THE STUDY

The main purpose of this study is to describe how speech characteristics arising from the cleft palate interfere with consonants in spontaneous speech of some Danish children. The phonetic and phonological consequences of this interference will be described.

The existing studies of speech sound production in children with cleft palate are based on tests, not on spontaneous speech - with Lynch et al. (1983) as an exception. Sound production in spontaneous speech shows greater variation than in speech elicited from tests, and because great variation in sound production is sometimes characteristic for individuals with cleft palate, spontaneous speech seems more suited to reveal the true extent of this variation.

Further, most of the existing studies of cleft palate speech are frequency studies, for instance such that deal with the frequency of correct or deviant pronunciation of a given phoneme, or the frequency of certain categories of deviations (see, e.g., Moll 1968, Fletcher 1978, Van Demark 1974). There are very few phonetic descriptions of cleft palate speech, although this is a necessary basis for describing how a cleft palate speaker deviates phonetically and phonologically from a target norm. Such descriptions are important for the following reasons: Firstly, the phonetic and phonological competence of the speaker play an important part in the communicative problems caused by the cleft palate. Secondly, it is possible to classify the speech in categories that are based on well-defined phonetic and phonological criteria.[1] Van Erp (1984) performed a phonetic transcription of Dutch cleft palate speakers, but

their deviant sound production is, also in this study, presented in terms of non-phonetic error categories. Finally, the vast majority of studies of cleft palate speech are based on American English speaking individuals. But one cannot implicitly assume that what is true about cleft palate speech in speakers of one language - for example American English - will necessarily be true about cleft palate speakers of other languages.

For these various reasons we found it relevant to look at the phonetic and phonological consequences that arise from cleft palate in the spontaneous speech of some Danish children.

III. METHOD

A. SUBJECTS AND EXPERIMENTAL SET-UP

The East Danish population of individuals with cleft palate consists of approximately 35 births per year, not including submucous clefts of the palate, or cleft lip only. About two thirds of these come from the greater Copenhagen area.

For this study we needed children of the same age, with normal intelligence, normal hearing, with no accompanying defects, but with speech defects, including some with glottal stop articulation. Also, for practical reasons they should live in the Copenhagen area. All these preconditions - along with parents who did not want to participate - reduced the number of potential children considerably. It should be added that two of the children previously had had periods of medically well-treated middle ear problems, but their hearing was normal at the time of recording. Due to the small number of subjects it was not possible to differentiate as to sex and type of cleft. Clinical information about the five children with repaired cleft palate are given in figure 1. Notice that two of the children have a normal velopharyngeal function.

A video film of about 30 minutes was made of each child in play with a parent, sitting at a small felt-covered table, so that at least the front of the child's, and 3/4 of the parent's face and upper torso could be viewed on the film. The toys were supplied both by the parents and the experimenters, but had to be "quiet" toys, to get a good sound recording for the phonetic analysis. All the children chose long periods of drawing and talking about this activity.

B. PHONETIC TRANSCRIPTION

The verbal communication between parent and child was transcribed orthographically. Furthermore, a narrow segmental transcription of the child's speech was performed in the IPA system supplemented by some ad hoc symbols.[2] The transcription was undertaken by a trained phonetician (BH) and discussed with

Child	Age	Sex	Cleft type	Age at primary palatal repair	IQ*	Hearing up to present age	Development of language	Velopharyngeal funct.	Hypernasality
I	5.0	M	lip/palate	22 months	normal	normal	delayed	sufficient	none
II	4.10	M	palate	22 "	normal	normal	normal	insufficient	mild
III	4.9	M	lip/palate	22 "	normal	well-treated fluctuating	normal	insufficient	severe
IV	4.2	F	palate	22 "	normal	normal	normal	insufficient	moderate
V	5.2	M	palate	22 "	normal	well-treated fluctuating	slightly delayed	sufficient	none

*) *Tested with Leiter, a non-verbal intelligence test, by the psychologists C. Lowe Nielsen and W. Cohn, two of the authors of the appendix to the present paper.*

Figure 1

Clinical information about the five Danish children included in the study and who have speech problems after primary cleft palate surgery.

a speech pathologist (KB) who, unlike the phonetician, has specialized in cleft palate. Since segmental transcription is extremely time consuming, we had to confine the number of transcribers and transcriptions to just one, in spite of the fact that transcriptions made by more transcribers show inter- as well as intra-transcriber variation (Vieregge 1985a). However, since the transcription was based on video recordings including visual as well as auditory information, the transcription task was somewhat facilitated. Theoretical and practical aspects of transcription are thoroughly treated in Vieregge (1985b).

Based on the phonetic transcription the speech of each child is analyzed with regard to the interference of the speech characteristics from cleft palate with normal Danish. Two types of interference are considered, namely interference with the sound realization of the phonemes, and with the number of phonemes. Interference with realization causes pronunciation to deviate more or less from the target norm, which does not necessarily damage intelligibility seriously, if the phonemes are still identifiable. Contrarily, interference reducing the number of phonemes - which occurs when it is not possible to distinguish between two or more phonemes - is in principle always damaging to intelligibility, as it may cause words to be pronounced identically that are distinguished in the norm language.

C. SPEECH MATERIAL

The sounds analyzed are limited to consonants. The Danish consonant system and its phonetic realization in syllable initial position preceding a full vowel are shown at the top of figure 3, while syllable final position is seen in figure 4. It is apparent that the initial system is the more differentiated one. However, final /*v j r*/ are excluded since they are realized as true semivowels forming part of diphthongs - and /*r*/ may even disappear in this position. Notice that in Danish both of the stop series are unvoiced. They are separated by a difference mainly in aspiration, so that /*ptk*/ are aspirated, /*bdg*/ unaspirated, but they only contrast in (absolute) initial position in syllables containing a full vowel. However, the main difference between /*t*/ and /*d*/ is one of affrication rather than aspiration. In other positions with stops only the unaspirated series /*bdg*/ occur. Notice also that the Danish *r*-sound is [ʁ], i.e. pronounced with the posterior part of the tongue.

In speakers who suffer from velopharyngeal insufficiency it is primarily the obstruents (= pressure consonants) that are distorted, due to the fact that the intra-oral pressure demanded for production of these sounds cannot be established. Also, these sounds may be disturbed by nasal friction. Contrarily, the oral sonorant sounds are usually normally produced, except for hypernasality, which is less disturbing to the communication. Nevertheless, only the vowels are excluded from the ana-

lysis besides final /v j r/, since consonants classified as sonorants may be affected by other distortions than nasalization.

Consonants are described in the following positions: 1. Single consonant in syllable initial position preceding a full vowel (ex.: *penge, Jakob*). 2. Single consonant in word final position (ex.: *Jakob*) (i.e. in syllable final position before word boundary), and 3. two-consonant clusters in syllable initial position preceding a full vowel (ex.: *sparke, puslespil*). The type of clusters in the present material includes the following combinations: *s* + unaspirated stop[3] (ex.: *stor*), stop + voiced consonant (ex.: *pjat, blå*), fricative + voiced consonant (ex.: *flaske*). Note that not all possible Danish combinations of consonants occur in the speech material. Some utterances or single words are left out because of uncertainty of their semantic content and/or because the pronunciation was impossible to transcribe meaningfully. You have to know both what is being said and how it is said to be able to compare the pronunciation with the target norm.

Table I

The table shows how many intended single consonants and consonant clusters are included in the description of the five cleft palate children. For comparison also a normal child (N) is included. CV = single consonant in syllable initial position preceding a full vowel, VC+ = single consonant in word final position, and CCV = two-consonant cluster in syllable initial position preceding a full vowel.

CHILD	CV	VC+	CCV
I	530	100	38
II	329	92	31
III	305	157	50
IV	421	127	51
V	809	216	45
N	693	222	61

In table I is shown how many intended consonants and consonant clusters are included in the description of each child. No study has been published about the frequency of occurrence of different sounds used in Danish spontaneous speech - neither for children nor for adults. To get an impression of any possible difference from the norm in cleft palate children in this regard, we video-taped a normal five year old girl in the same

play situation as the cleft palate children. It will be seen that there is a good deal of variation among the children. Child V's output is by far the largest, even though only 3/4 of his utterances have been included. N's output comes second, but taking into account that nothing had to be left out due to unintelligibility, it seems safe to conclude that the normal child does not differ from the five cleft palate children.

To the left in figure 2 is shown the frequency of a given intended consonant or consonant cluster in percent of the total number of intended consonants, averaged over the five children with cleft palate. To the right is shown the same data for child N. Some consonants and consonant clusters seem to appear relatively more frequently than others, but no fundamental difference is found in the frequency of the various consonants and consonant clusters between N and the children with cleft palate. This difference in frequency is partly due to the fact that some specific sounds appear in particles and inflections which are frequent in spontaneous speech. If the consonant frequency is calculated from appearance in different words, the frequency of some consonants and clusters is reduced (especially initial /d j/, final /n ɑ:ɳ/ and initial /s+k/). The number of words with initial consonant appears to be sufficient to give a comprehensive description of the variation in pronunciation for each consonant. The occurrence of words with intended final consonants and initial consonant clusters, however, may for certain consonants and consonant clusters be limited to one word - or none. Half an hour's spontaneous speech at play with a parent, therefore, is not enough to get a sample in which the number of all possible consonants and clusters is sufficient. On the other hand, the normal speech sample from child N suggests that the less frequent single consonants and consonant clusters are the same for normal as for cleft palate children of the same age.

As the phonetic/phonological conditions in normal 4-5 year old children are equal to those of adults as far as the consonant material analysed in this study is concerned (excepting a very few consonant clusters), the phonetic/phonological conditions in the five cleft palate children will be compared to normal adult language.

IV. RESULTS

In figures 3 and 4 - below the normal consonant system and its realization - is shown the realization(s) most frequently used by each child for a given phoneme, called the primary variant(s). For each child is also shown how often the primary variant(s) is/are used in relation to other variants.

Encircling of the phonetic symbol indicates that the primary variant(s) is/are in accordance with the norm. In figure 5 is shown in outline the realizations of the consonant clusters for each child.

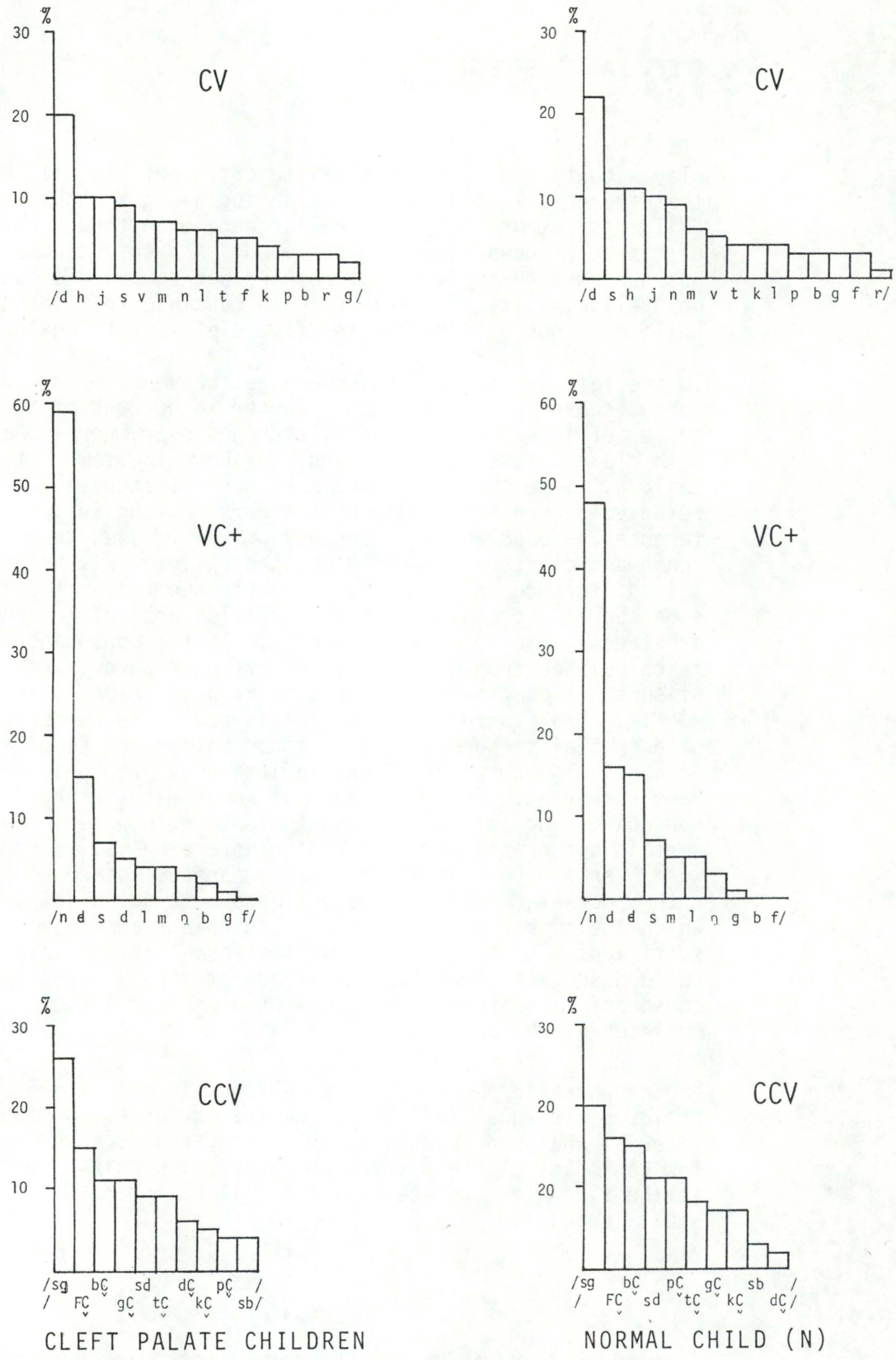


Figure 2

The figure shows the frequency of a given intended consonant or consonant cluster in percent of the total number of intended consonants. To the left data averaged over the five cleft palate children are indicated, and to the right the data for a normal child (N) is shown. The consonants and consonant clusters are shown with decreasing frequency. CV = single consonant in syllable initial position preceding a full vowel, VC+ = single consonant in word final position, and CCV = two-consonant cluster in syllable initial position preceding a full vowel. C̬ = voiced consonant, F = fricative (≠ /h/).

THE DANISH CONSONANT SYSTEM AND ITS REALIZATION -

IN SYLLABLE INITIAL POSITION PRECEDING A FULL VOWEL

PHONEMES	/ p	t	k	b	d	g	f	s	h	v	l	r	j	m	n /
REALIZATION	[$\underset{\circ}{b}^{h}$	$\underset{\circ}{d}^{s}$	$\underset{\circ}{g}^{h}$	$\underset{\circ}{b}$	$\underset{\circ}{d}$	$\underset{\circ}{g}$	f	s	h	v	l	ʁ	j	m	n]
Child I															
Primary variant(s)	[(b$^{h}_{\circ}$)	?	?	(b$_{\circ}$)	?	?	?	?/h	(h)	(v)	(l)	h	(j)	(m)	(n)]
Frequency (%)	*94*	*90*	*75*	*82*	*82*	*80*	*95*	*50/ 43*	*100*	*78*	*91*	*75*	*89*	*100*	*57*
Child II															
Primary variant(s)	[*F*	*F*	?/?^{h}	?$\underset{\smile}{m}$	?$\underset{\smile}{n}$	?	(f)	$\dot{s}/\dot{\underset{\circ}{n}}$	(h)	$\tilde{v}$	$\tilde{l}$	$\tilde{ʁ}$	$\tilde{j}$	(m)	(n)]
Frequency (%)	*71*	*45*	*44/ 31*	*100*	*68*	*56*	*75*	*24/ 22*	*91*	*74*	*75*	*50*	*97*	*100*	*100*
Child III															
Primary variant(s)	[(b$^{h}_{\circ}$)	$\dot{n}$/(d$^{h}_{\circ}$)	$\eta_{\circ}$	(b$_{\circ}$/m)	n/(d$_{\circ}$)	n	(f/$\dot{f}$)	$\dot{s}$	(h)	$\tilde{v}$	$\tilde{l}$	$\tilde{ʁ}$	$\tilde{j}$	(m)	(n)]
Frequency (%)	*43*	*30/ 25*	*36*	*46/ 31*	*56/ 21*	*44*	*44/ 32*	*38*	*100*	*76*	*70*	*75*	*100*	*100*	*100*
Child IV															
Primary variant(s)	[m	h	h	m	n/?	h/?	h	h/?	(h)	$\tilde{v}$	$\tilde{l}$	$\tilde{ʁ}$	$\tilde{j}$	(m)	(n)]
Frequency (%)	*71*	*56*	*67*	*91*	*60/ 14*	*27/ 27*	*86*	*71/ 29*	*89*	*58*	*83*	*43*	*99*	*94*	*93*
Child V															
Primary variant(s)	[(b$^{h}_{\circ}$)	(d$^{h}_{\circ}$)	h	(b$_{\circ}$)	(d$_{\circ}$)	d$_{\circ}$	(f)	(s)	(h)	(v)	(l)	(ʁ)	(j)	(m)	(n)]
Frequency (%)	*71*	*100*	*68*	*100*	*97*	*94*	*98*	*99*	*100*	*98*	*98*	*100*	*95*	*98*	*93*

Figure 3

At the top of the figure is shown the Danish consonant system and its realization for consonants in syllable initial position preceding a full vowel. Below, the primary realizations - primary variants - used by each child are indicated. Also, the frequency of occurrence for each variant is shown (in %). Encircling of the phonetic symbol indicates that the primary variant is in accordance with the target norm. [$^{\bullet}$] *means 'nasal friction'. As to child II, no primary variant can be decided for /pt/; instead the frequency of variants produced with any kind of friction is indicated, symbolized by* [*F*]. *As to child III, /k/ and /g/ are in no case correctly produced, but the deviating realizations can be explained by assimilation. With child IV the interdental realization of /n d l/ is included in the frequency of occurrence.*

Child I

1. Single consonant.

INITIAL: Child I's realization of stops and fricatives is dominated by glottal stops [ʔ]. Only the bilabial stops are - with a few exceptions - pronounced according to the norm, and never as [ʔ]. Notice also that /*f*/ is realized as [ʔ], whereas /*s*/ is realized almost equally frequently as [ʔ] and as [h]. /*h*/ is correctly produced, but [h] is also the frequent pronunciation of /*r*/. The remaining voiced consonants are usually pronounced according to the norm. Moreover, Child I in nearly all cases adds an [h] in front of word initial vowels, a position where /*h*/ occurs in the normal language, so that this "addition" of [h] may cause confusion of words.

FINAL: In this position child I does not use glottal stop. He shows a clear tendency to deletion of the nasal consonants /*n*/ and /ŋ/, and of unvoiced consonants, while the phonetic realization of the remaining three phonemes is normal in most cases.

2. Initial clusters.
With a few exceptions all clusters are reduced to single consonants (≈100%). The most frequent pronunciation is [ʔ] (about 60%), /*b*/ + sonorant consonant being an evident exception as it is always realized as either [b̥ʰ] or [b̥]. It cannot be decided, however, if these manifestations can be generalized as valid for bilabial stop + sonorant consonant, since /*p*/ + sonorant consonant does not occur in the sample. But it is clear that his realization of stop + sonorant consonant cannot simply be described as identical with the realization of the stop as single consonant, because /*b*/ as single consonant is realized solely as [b̥].

It should be added, too, that child I has a number of assimilations of word initial consonants, both progressive and regressive (ex.: *vm[m]armere, nl[l]emlig, larmer lm[m]idt, læse nl[l]ogen*). Also, the way in which child I treats the Danish stød is worth a comment. Danish stød is considered a prosodic phenomenon with the syllable as its domain: if the vowel is long, the stød will appear at the end of the vowel; if the vowel is short, it appears on the following voiced consonant. In child I's sample 67 words should be produced with stød, but he only uses stød in 10% of these words. In most of the cases, where stød should have occurred on the vowel, it is omitted while keeping the vowel long (ex.: *sol* [so:'l]→[ʔo:l], *træ* [d̥ʁɛ:']→[ʔɛ:]). In the remaining cases with stød on the vowel, child I uses a short vowel. In the words occurring in the sample, where stød should have been on the voiced consonant, he sometimes lengthens the preceding vowel (ex.: *hånd* [hʌn']→[hʌ:n]), but more often the stød is just omitted.

As can be seen from the diagram in figure 1, child I's language development is delayed (in morphology and syntax). The assimilations, the problem with stød, and the addition of word initial [h] is probably also due to the language retardation.

THE DANISH CONSONANT SYSTEM AND ITS REALIZATION -

IN SYLLABLE FINAL POSITION

PHONEMES	/ b	d	g	f	s	l	ɖ	m	n	ŋ /
REALIZATION	[b̥	d̥	g̊	f	s	l	ɑ	m	n	ŋ]
Child I										
Primary variant(s)	[	0			0	(l)	(ɑ)	(m)	(n)/0	0]
Frequency (%)		÷ 100	÷	÷	100	89	90	67	45/35	80
Child II										
Primary variant(s)	[ʔm̥		0		s	l̃	ɑ̃	(m)	(n)	]
Frequency (%)	50	÷	50	÷	50	100	100	100	95	÷
Child III										
Primary variant(s)	[(b̥)	0/(d̥)		f̥	s/s̥*	l̃	ɑ̃	(m)	(n)	(ŋ)]
Frequency (%)	67	43/29	÷	100	33/33	100	100	100	93	100
Child IV										
Primary variant(s)	[b*	0			h/0*	l̃	ɑ̃	(m)	(n)	(ŋ)]
Frequency (%)	100	100	÷	÷	67/33	100	97	86	76	100
Child V										
Primary variant(s)	[	(d̥)/0			(s)	(l)(ɑ)/0		(m)	(n)	(ŋ)]
Frequency (%)	÷	50/42	÷	÷	73	100	65/35	75	88	83

Figure 4

At the top of the figure is shown the Danish consonant system and its realization for consonants in syllable final position. Final /v j r/ are excluded since they are realized as true semivowels forming part of diphthongs - and /r/ may even disappear. Below, the primary realizations - primary variants - used by each child are indicated. Also, the frequency of occurrence is shown (in %). Notice that only consonants in syllable final position before word boundary are included. Encircling of the phonetic symbol indicates that the primary variant is in accordance with the target norm. As to child IV, the interdental realization of /l n/ is included in the frequency of occurrence. 0 in the row of phonetic symbols means 'omission of consonant', while ÷ in the row of frequency of occurrence indicates that the intended sound occurs only once or not at all with a given child. [•] means 'nasal friction' and [*] 'weak articulation'. The symbol ɖ is used for the 'soft d', which is a voiced approximant with a slight raising of the front and back of the tongue.

Also, as child I's velopharyngeal function is normal, his comprehensive and persisting use of glottal stop may be reinforced by his language delay, even though a status of the total East Danish cleft palate population born 1970-1978 (N = 293), shows no significant relationship between language delay and the use of glottal stop. In any case, the consequence of his dominant use of [ʔ] and [h], and of the frequent omission of final consonants is that both the phonetic realization and the number of phonemes deviate considerably from the norm.

Child II

1. Single consonant.

INITIAL: Child II also uses glottal stop, mainly for the stop consonants. The glottal stops, however, in most cases are articulated in combination with the correct supraglottal place of articulation. Only the velar stops are most frequently realized as a simple glottal stop. Also, child II is characteristic in that he shows a large variation in his realization of stops, except for /*b*/, which is always realized as [ʔm̆]. For /*p t*/ the variation is so great that the occurrence of the different variants is too small to decide which is the primary variant for these stops. But there is a clear tendency to maintain a distinctive difference between aspirated and unaspirated stops, friction of some kind being frequent in the manifestation of /*ptk*/.

/*s*/ also has many variants, but there are two primary variants, namely the nasal fricative [n̥̈] and *s* with audible nasal emission, that is with nasal friction [s̈][5]. The remaining /*s*/-variants generally have a weaker and more diffuse friction noise, but the place of articulation for all of them - including the primary variants - is dental/alveolar. /*h*/ and /*f*/ and the voiced consonants are mostly correct, except that the oral voiced consonants are nasalized.

FINAL: In child II is also found omission of final stop consonant - but the occurrence of words with final consonants is very limited. It is noticeable, though, that final /*b*/ may be pronounced as a combination of bilabial and glottal stop, just as bilabial stops in initial position. /*s*/ is realized either as a 'weak s', that is as a weak and diffuse s-friction [s̻][5], or with evident nasal friction [s̈]. The voiced consonants are pronounced normally, except for nasalization of the oral consonants.

2. Initial clusters.
There is a clear tendency for child II to realize clusters of two consonants as a combination of two consonants (67%), often in the way they are realized as single consonants - and thus with great variation (ex.: *blir* [ʔm̆ɪ̈], *smadre* [n̥̈m]). /*s+n*/, though, is always reduced to [n̥̈], which can be described as omission of /*n*/, since [n̥̈] is a frequent realization of /*s*/ (ex.: *snakke* [n̥̈]). This is possibly not to be explained as a simple omission but as a process in which /*n*/ is assimilated

Child I

Main pattern:	/CC/ > [C]	(100%)
example:	/CC/ > [ʔ]	(60%)
Systematic exception:	/bC/ > [b̥ʰ] or [b̥]	

Child II

Main pattern:	/CC/ > [CC]	(67%)
example:	/bl/ > [ʔm̰l̰]	
Systematic exception:	/sn/ > [ṅ̥]	

Child III

Main pattern:	/CC/ > [CC]	(70%)
example:	/tr/ > [ṅ̥ʁ̰]	
Systematic exception:	/sb/ > [ṁ̥]	

Child IV

Main pattern:		
if C_1 = fricative	/C_1C/ > [C]	(90%)
example:	/C_1C/ > [h]	(88%)
if C_1 = stop	/C_1C/ > [C]	(65%) (≠ [h] (73%))
example:	/gr/ > [ʁ̞]	

Child V

Main pattern:	/CC/ > [C] or [CC]	depending on cluster type
examples:	/sg/ > [s]	
	/sd/ > [d̥]	
	/bl/ > [bl̥]	
	/sn/ > [sn]	

Figure 5

The figure shows in outline the realization of two-consonant clusters in syllable initial position preceding a full vowel. C = consonant, [˙] = nasal friction.

to /s/ with regard to voicing and friction - that is, the two consonants become identical ([n̥n]→[n̥n̥]) after which one of them is omitted. Moreover, assimilation of place of articulation can explain that /s+g/ as a cluster is often realized as [n̥ʔn] because /d/ as a single consonant - but never /g/ - is realized as [ʔn] (ex.: *skal* [n̥ʔn]).

Thus, child II's pronunciation of consonants deviates considerably from the target norm, while the decrease in number of phonemes is limited. His language development is normal, so that his consonant problems appear to be a direct consequence of his cleft palate, that is, of his insufficient velopharyngeal function.

Child III

1. Single consonant.
Glottal stop does not occur in child III (except three times sporadically).

INITIAL: Child III's realization of stops is often correct, least frequently for the velar stops, as they are nearly always assimilated to the surrounding consonants in place of articulation and/or nasality. The realizations are furthermore characterized by the use of voiced nasal consonants and unvoiced nasal consonants with nasal friction (i.e. nasal fricative), but other variants occur as well. Even though several of the variants deviate from the norm, there is a clear tendency in child III to try to maintain a distinctive difference between aspirated and unaspirated stops, as the voiced nasal consonants are the dominant realization of the unaspirated stops, while the unvoiced nasal consonants are dominant for the aspirated ones. The pronunciation of /f/ is often correct, but /f/ accompanied by nasal friction also occurs ([f̊]). The primary variant for /s/ is *s* + nasal friction alias [s̊]. There is, however, great variation in the pronunciation of the fricatives /f/ and /s/, although the place of articulation is nearly always correct. The voiced consonants are usually correctly produced - except for nasalization of the oral consonants. It should be added that child III shows examples of correct manifestation of all consonants (except /s/) in spite of evident hypernasality.

FINAL: For child III omission of final consonants is limited to /d/, as /b/ is never omitted (no occurrence of words with final /g/). /b/ and /d/ - when not omitted - are often correctly produced. /f/ is most frequently produced with accompanying nasal friction, [f], whereas /s/ is realized either with accompanying nasal friction, [s̊], or as weak, diffuse *s*-friction alias [s]. The voiced consonants are pronounced according to the target norm, except for nasalization of the oral consonants.

2. Initial clusters.
Child III mostly realizes clusters of two consonants as a cluster (about 70%), often consisting of consonants the way they are realized as single consonants (*trapperne* [n̥ʁ], *fjernsyn* [f̊j]). Omission of consonants is most frequent in /s+b/,

most often pronounced as [m̥]. This can be explained as a process in which /s/ - often as a single consonant realized as [n̥] - is assimilated in place of articulation, after which /b/ is omitted.

Finally, both progressive and regressive assimilation occurs in child III's speech, most frequently "fronting" of the velars accompanied by assimilation of nasality (*g*[n]*år* *d*[n]*e*, *hende* *g*[n]*år*).

As the correct pronunciation of consonants in child III's speech are quite often among his variants, and as omission of final consonants also in clusters is limited, his pronunciation of consonants and the number of phonemes are found to be only moderately deviating from the norm - judged by the primary variants.

Child IV

1. Single consonant

INITIAL: Child IV's realization of stops and fricatives is dominated by the nasal consonants [m] for /p/ and /b/, and [n] for /d/, and by [h] for the remaining stops and for the fricatives. Glottal stop is also relatively frequent for /d/, /g/ and /s/. [h] and [ʔ] occur sporadically for nearly all the voiced consonants, and even /h/ may be realized as [ʔ]. Voiced consonants are usually correctly produced, though, except for nasalization of the oral consonants. It should be mentioned, too, that child IV shows a rather large variation in her realization of alveolar and velar stops.

FINAL: In child IV's sample there are not many occurrences either with intended final stop or fricative. But in the very limited material /d/ is always omitted, whereas /b/ is realized as a very 'weak *b*', actually only a very loose constriction (/g/ occurs only once - correctly produced). /s/ is also realized as a very weak and diffuse friction, [h̥]. /n/ too may be omitted, but otherwise the remaining voiced final consonants are usually correct, except for nasalization of the oral consonants.

2. Initial clusters.
If the first consonant is a fricative, child IV mostly reduces the cluster to single consonant (90%), but only /s+g/ occurs more than a few times. The realization is usually [h], but [ʔ] also occurs. As /s/ and /g/ as single consonants are both realized as [h] and [ʔ], it cannot be decided if the realization of /s+g/ should be described as omission of /s/ or of /g/. An argument in favour of the stop being omitted would be that /s+d/ is realized as [h] - in the few occurrences found - and that /d/ as a single consonant seldom is realized as [h]. If the first consonant is a stop, the frequency of reduction to a single consonant decreases (65%). Only /b/ and /d/ + a sonorant consonant occur more than a few times. With /b/ the realization shows great variation, but mostly it is pronounced either as single [m] or followed by a sonorant consonant (ex.: *bruge* [m] or [mʁ̃]). With /g/ the cluster is most frequently

reduced to a single consonant. In all but one case the sonorant is /r/, mostly realized as a retracted uvular unvoiced fricative [ʁ̞]. As the production of single /r/ in the target norm may be very close to an unvoiced uvular fricative, child IV's pronunciation of /g+r/ is not as deviating as it may appear.

It should be mentioned that child IV speaks with a very low intensity, and that her speech is frequently interrupted by superficial inspiration. Her comprehensive use of nasal consonants, [h] and - to a lesser degree - [ʔ], means that both pronunciation and number of phonemes deviate considerably from the norm. Also, her [n], which is the realization of /n/ and frequently also of /d/, is in many cases produced with interdental articulation. This also applies to /l/. Child IV's use of glottal stop is noticeable, as it occurs also sporadically for sonorant consonants and /h/. This will be discussed later.

As she shows no language delay, her consonant problems may be considered due to her cleft palate.

Child V

1. Single consonant.

INITIAL: In child V's sample only few, sporadic deviances from the norm occur, except for the velars /k/ and /g/, which are realized as [h] and [d̥], respectively. As his velopharyngeal function is normal, [h] for /k/ may be a reminiscence of an earlier insufficiency. His "fronting" of /g/, however, may be seen as a symptom of his slight language delay, and it is possible that also the [h] for /k/ is related to the language delay.

FINAL: The pronunciation is usually correct, but omissions do occur for nearly all types of consonants.

2. Initial clusters.
Child V's clusters are realized both as single consonants and as clusters, depending on cluster type. However, his treatment of clusters is common for children with developmental language delay as illustrated by the following examples: *grusgrav* [ʁ], *skal* [s], *stor* [d̥], *svært* [f]. The reduction in the last example can be described as a process where /v/ is assimilated to /s/ with regard to voicing, and then /s/ is omitted.

Also, child V uses many assimilations in his speech. Assimilations over a distance are regressive, while contact assimilations are mainly progressive (*k*[f]*om til at* $\overleftarrow{\text{få}}$, *d*[b̥]$\overleftarrow{\text{em}}$, *fars* $\overrightarrow{\text{k}}$[s]*ammerater*).

It may be concluded that in child V's speech the interference with both realization and number of phonemes is very limited. Furthermore, it is supposed that his slight consonant problems are due to his mild language delay rather than to a reminiscence from his velopharyngeal insufficiency.

V. DISCUSSION

The results presented above show that the five children differ in their phonetic as well as in their phonological behaviour. In this section the results are considered in a more general aspect concerning strategies in cleft palate speech. Also, some aspects on universal versus language specific characteristics in cleft palate speech are treated.

A. STRATEGIES IN CLEFT PALATE SPEECH

Speech characteristics associated with cleft palate relate primarily to the insufficient velopharyngeal function, but also dyscoordinated mobility due to reduced sensibility in the anterior part of the tongue must be considered.

Velopharyngeal insufficiency results in abnormal pronunciation of different origins. If the abnormal pronunciation is a consequence of insufficient velopharyngeal valve function only, with the glottal and supraglottal articulations being normal, the resulting abnormal pronunciation of the speech is unavoidable and thus passive in nature. This may be considered as the "laisser faire" strategy or the passive strategy in cleft palate speech. Contrarily, if the speaker attempts to reduce unavoidable consequences of his or her handicap, the abnormal pronunciation is due to an active behaviour. There are two techniques of reduction, and these may be considered as the active strategies in cleft palate speech. Below is described how the physiological characteristics related to each strategy affect various types of sounds. Next, the five children are discussed according to this description. It should be pointed out, that in the present description the speech characteristics are considered from a causal point of view rather than a symptomatic one dealing with the consequences for the sound production, as is the case in most other descriptions.

The consequences of the passive strategy depend on the vocal tract configuration, i.e. on the type of speech sound to which it is applied. In the description of these consequences the velopharyngeal insufficiency is supposed to prevent any considerable increase in the intraoral pressure, and the glottal and supraglottal articulations are as in the target sound except for the velopharyngeal behaviour.

The oral non-pressure sounds (also including vowels) are voiced and thus realized as the nasalized cognates of the intended sounds. If intended voiced fricatives are realized as voiced, not only are they nasalized, but the conditions for friction are also reduced, resulting in a more or less non-fricative realization. In this context Ohala (1975) should be mentioned because he questions the existence of friction noise in speech sounds categorized as nasal voiced fricatives. As to the stops intended to be voiced, the passive strategy results in nasalization; in other words, these stops are realized as (voiced) nasal consonants.

A special category comprises unvoiced stops produced with a state of the vocal folds which results in voicing when airflow passes between them. This is the case in Danish (unvoiced) unaspirated stops. Since a blockage of the airflow is impossible in case of a velopharyngeal leakage, the Danish unaspirated stops will be voiced with the passive strategy. The only means to devoice these stops are to abduct the vocal folds - or to make a glottal closure - and then the strategy is no longer purely passive. Thus, the consequence of the velopharyngeal insufficiency is that the Danish unaspirated stops are realized as nasal consonants with the passive strategy. Voicing and devoicing in Danish stops - which is a more complicated matter than appears from the present description - is discussed in Hutters (1985).

Unvoiced sounds produced with a high rate of airflow presuppose abduction of the vocal folds. This category comprises fricatives and aspirated stops. The velopharyngeal leakage results in nasal emission of air, but due to the high rate of airflow they are produced with audible nasal emission of air, i.e. with nasal friction. Thus, the unvoiced fricatives are produced as a combination of oral and nasal friction (or, in case of turbulence, with nasal snort), one of them being the more dominating one. The aspirated stops will be realized as unvoiced nasal consonants with more or less nasal friction (or snort).

Thus, a characteristic feature of the passive strategy is nasalization and nasal friction. In the present description it is supposed that any considerable increase in intraoral pressure is prevented by the leakage of air through the velopharynx. However, in a speaker who sticks to the passive strategy, the resistence to the air may in fact be sufficient for production of consonants that are more similar to the intended pressure consonants than appears from the present description. The better pressure conditions - for a given vocal tract configuration - may be due not only to a less open velopharyngeal passage but also to the conditions in the nasopharynx and in the nasal airway. These factors, of course, also influence the degree of nasalization.

There are two active strategies that cleft palate speakers may utilize in order to reduce the unavoidable consequences of the organic handicap. One strategy is to compensate for some or all of the pressure consonants normally produced at, or in front of the velopharyngeal valve, by making constrictions appropriate to consonants produced posterior to the valve. Thus, the place of articulation is changed to an area in the vocal tract where obstruction to the airflow is possible in spite of the velopharyngeal leakage. Thus, by definition, the compensatory strategy only applies to pressure consonants. The other types of sounds are normally realized by means of the passive strategy. The compensatory sounds are primarily glottal stops and pharyngeal fricatives. However, Trost (1981), Lawrence and Phillips (1975) and Phillips and Kent (1984) also find pharyngeal stops in individuals with cleft palate. This finding is

interesting, since with normal speakers it is assumed that it is not possible to establish a closure between the posterior pharyngeal wall and the tongue root, which explains why pharyngeal stops have not been registered as speech sounds.

The other active strategy is to mask or camouflage the unavoidable consequences of the organic handicap by means of weak articulation resulting in weak sound production. This strategy may apply to all types of sounds. Included in this strategy of weak sound production is the use of [h] for obstruents, since [h] is a weak sound. Thus, the use of this sound should not be considered a consequence of the other active strategy, where the place of articulation is posterior to the velopharyngeal valve in order to produce an effective blocking or constriction of the air, even though [h] is normally classified as a glottal fricative. Also, the pharyngeal and nasal alea constriction may be considered part of the active camouflage strategy, as it is an attempt to reduce nasal emission of air. Paradoxically, the nasal alea constriction may result in a more audible nasal emission.

A common realization of /s/ and other sibilant fricatives is a "nasal fricative", i.e. an unvoiced nasal consonant with friction. There may be two explanatory reasons for this realization. One is that the speaker wants to produce a more audible friction noise. If so, the strategy is of the compensatory active type. However, in this case the change in articulation is better described as a change in manner of articulation, that is an oral versus a nasal fricative, rather than in place of articulation. Another reason could be a reduced sensibility and dyscoordinated mobility of the tongue tip and blade which may result in articulation with the front of the tongue rather than with the tongue tip and blade. Since the production of sibilants is very demanding on tongue articulation, the speaker with tongue problems may "choose" the much simpler closing tongue gesture required for the nasal fricative. Notice that a nasal fricative may also be a consequence of the passive strategy, but in this case it is an abnormal production of dental/alveolar stops.

Further, a common consequence of the reduced sensibility and dyscoordinated mobility of the tongue tip and blade, resulting in articulation with the front of the tongue, is that dental/alveolar consonants are produced as palatal or prepalatal consonants. This tendency to articulation with the front of the tongue may also explain the paradox that the same consonants may be produced with interdental articulation - provided that the interdental articulation is considered a consequence of the front of the tongue articulating against the teeth and the alveolar ridge.

Only children who suffer from a persisting velopharyngeal insufficiency need to "choose" a strategy for their speech production. Thus, only the children II, III, and IV answer to this description. However, they differ as to "choice" of strategy, which appears from the realization of the pressure consonants.

Both child II and child IV use active strategies. As to child II, glottal stop is used for the stops - frequently articulated in combination with the correct supraglottal articulation - while /s/ is produced either as a nasal fricative or as an *s* with nasal friction. Thus, child II's realization of the pressure consonants is dominated by the active compensatory strategy. Child IV's realization of these consonants is dominated by [h] and weak sound production, which are characteristics of the active camouflage strategy. However, the pronunciation of some of her stops is dominated by nasal consonants produced with the correct place of articulation. This feature of nasal consonants may be explained by the overall weak articulation including the abduction of the vocal folds. That is, as explained above, if a laryngeal devoicing mechanism is not introduced simultaneously with the supraglottal closure, the intended stop consonant is realized as a voiced nasal consonant due to the velopharyngeal leakage. Of course, the production of nasal consonants instead of oral stops may also result from the passive strategy. However, with the purely passive strategy stops which in normal speech are articulated with a considerable abduction of the vocal folds should not be realized as voiced nasal consonants. Thus, if passive, her /p/ should be realized as [m̥] rather than [m].

Also, it is surprising that glottal stop occurs spontaneously in her speech in spite of her camouflage strategy. However, based on a more detailed analysis of her use of glottal stop, it appears that this sound only occurs in word initial position, i.e. at a word boundary, no matter whether the word in the target norm has a consonant (or consonants) in this position, or whether it starts with a vowel.[6] In the latter case it functions obviously as a 'hard' attack, that is as a non-phonemic sound, whereas the former case may be interpreted as a 'hard' attack with omission of the word initial consonant rather than as a phonemic compensatory glottal stop. One argument in favour of the non-phonemic interpretation of her glottal stops is that if phonemic, it must be accepted that they are used for non-pressure as well as for pressure consonants. Thus, her use of glottal stop does not seem to be compensatory in nature. Also, it should be recalled that child IV shows interdental articulation, probably due to a vacillating use of the tip and blade of the tongue as described above.

In child III's speech the fricatives are realized with nasal friction, while the aspirated and unaspirated stops are realized as unvoiced nasal consonants and voiced nasal consonants, respectively. Thus, child III's pressure consonants are dominated by the passive realization - when they are not correctly produced.

The children I and V have normal velopharyngeal function. Nevertheless, child I's speech is dominated by glottal stops instead of normal pressure consonants. Thus, what was former an active strategy, intended to compensate for a velopharyngeal insufficiency, has now become a habit.

In the speech of child V, as stated above, the very few abnormal consonant realizations are probably due to his slight language delay rather than to a former velopharyngeal insufficiency. Thus, a description according to strategies is irrelevant for child V.

Child I exemplifies the well-known fact that glottal stop - and other compensatory sounds - is a very ingrained habit even after the velopharyngeal function is no longer insufficient. Henningsson and Isberg (1985) have demonstrated that in speakers with good or moderate velopharyngeal activity and with persistent use of glottal stops, the velopharyngeal activity is considerably reduced during pronunciation of these glottal stops. The explanation seems to be that since sounds produced below the velopharyngeal valve are not impaired by a velopharyngeal leakage, the speaker may not establish a closed velopharyngeal port during these sounds. This opening, which is harmless to the production of glottal stops, may spread to the surrounding oral sounds and thus give a false impression of organically caused velopharyngeal insufficiency. That is, the apparently harmless habit may turn into a very speech-impairing habit. In child I there is no apparent velopharyngeal insufficiency in his speech, so that even if there may be reduced velopharyngeal activity during his glottal stop production, it is of no consequence to his sound production. One may wonder if the apparent velopharyngeal insufficiency in child II, who also employs compensatory glottal stops, may be secondary to a reduced velopharyngeal activity in the glottal stops spreading to the surrounding sounds, and thus complicating the diagnosis of organic versus non-organic causes of his persisting "cleft palate speech".

In normal speech the same connection between glottality and the velopharyngeal mechanism has been observed, namely that the velopharyngeal port may be opened or closed during production of glottal stop and [h] according to the context (Ohala 1972). This optional velopharyngeal leakage has been proposed to explain that nasal vowels - which are normally developed from a context of nasal consonants - may develop from oral vowels preceded by a glottal stop and [h] (Matisoff 1975). From this it may be worthwhile to examine [h] in cleft palate speakers with good or moderate velopharyngeal activity to see if the velopharyngeal behaviour is the same as with the glottal stop.

In conclusion, due to the origin of the passive strategy, speakers suffering from velopharyngeal insufficiency will always show characteristics related to this strategy in their speech. Consequently, the passive strategy can actually not be "chosen". What the speaker can choose is to supplement his/her passive cleft palate speech with some of the characteristics from the active strategies, even though we talk about "choice of strategies" including the passive one. So, it appears that cleft palate speech may contain characteristics from several strategies with some characteristics being the more dominant ones. On the other hand, due to the antagonistic

nature of the two active strategies, they are supposed to be mutually exclusive, even if they, though probably mistakenly, seem to occur simultaneously as seen in child IV. As appears from the preceding discussion this description - including its basis and implications - seems to be profitable to the description of cleft palate speech and its varieties.

B. UNIVERSAL AND LANGUAGE SPECIFIC CHARACTERISTICS IN CLEFT PALATE SPEECH

Characteristics that originate directly from the organic conditions of the speech production apparatus are by definition universal rather than language dependent. As described above, the strategies in cleft palate speech are related to the organic conditions; thus, the strategies per se must be considered universal. Also, within the pressure consonants, differences according to place and manner of articulation have been reported. As to place of articulation, the labial stops are more frequently produced according to or closer to the target norm than stops produced with the tongue. This may be explained by the greater demands in the production of lingual sounds. Within the lingual stops it is further reported that those normally produced with the tip and blade of the tongue are more frequently distorted than are stops produced with more posterior parts of the tongue. These differences may be explained by the reduced tongue tip/blade sensibility observed in some cleft palate speakers. The tongue problems may also explain that sibilant fricatives are observed to be more frequently produced in an abnormal way than stop consonants. These differences as to place and manner of articulation, provided that they originate from the organic conditions, should be observed in cleft palate speakers irrespective of language, and thus be universal in nature. But due to the very restricted amount of data in the literature stemming from other languages than English, our knowledge is only fragmentary in this area. Only the observation that labial stops are more frequently normally produced does occur in the present material (see figure 3), which may be accidental and due to the very restricted number of children. One obvious feature originating from the tongue problem is the interdental articulation occurring in child IV.

Contrarily, since languages differ as to their sounds and sound structure, the consequences for communication induced by the passive strategy will also differ from language to language. Thus, the unavoidable phonetic and phonological difficulties that cleft palate speakers may have are evidently language dependent. However, the question is whether some of these speakers show language dependent reactions to the passively conditioned characteristics. For instance, why do some cleft palate speakers stick primarily to the passive "laisser faire" strategy, whereas others supplement it with one or the other active strategy? Theoretically it could be that the sound structure of the speaker's language is decisive for the choice of strategy(ies), and if so, the choice is by definition lan-

guage dependent. It is a well-known fact that all strategies occur in one and the same language, as also appears from our Danish children. But what we do not know is whether one strategy is more dominant in some languages than in others. For instance, it may be hypothesized that the active compensatory strategy is less frequent in languages which include the compensatory sounds in their sound system. These pharyngeal and glottal sounds have a distinctive function in many languages, but unfortunately, the data available in the literature do not permit any suggestions concerning this question. In the appendix following the present paper a psychological aspect is considered to be a conceivable part of the answer to the question of choice of strategy.

Another question relates to the fact that the phonetic output of cleft palate strategies normally have more or less severe consequences for the target phonological system of the speaker. This aspect of cleft palate speech is obvious in our Danish material. Do cleft palate speakers react to these phonological consequences? Our Danish material illustrates that the phenomenon seems to exist as it appears from child I and his treatment of the stød and glottal stop. After some general comments on glottal stop, this will be described in the following.

Like other sounds occurring in speech, glottal stop can be described with regard to production and to function. Its production is normally described as e.g. by Bergendahl and Fex (1977): *"the vocal folds are pressed forcibly together and separated with a plosive sound by the increased subglottal pressure"* (p. 17). However, this description seems to apply to a forceful realization of glottal stop. For glottal stop - as for all other sounds in speech - the force of articulation depends on factors like overall level of speech intensity, position, and stress. In our material on cleft palate speech it has been observed that perceptually clear glottal stops may in fact be produced without a true closure of the glottis, since weak and irregular voicing appears in the microphone curve. But this needs further examination.

As to the function of glottal stop, it may be phonemic or non-phonemic. When phonemic, glottal stop has a distinctive function, which is found in many languages. When non-phonemic, it occurs before a word initial vowel, and its function may be to signal a boundary or a feature of emphatic speech. Non-linguistic terms for these non-phonemic glottal stops are "hard attack" or "phonatory glottal stop" (Bergendahl and Fex, 1977). The phonemic use of glottal stop does not exist in Danish, whereas the non-phonemic one is optional.

In cleft palate speech the compensatory glottal stop has by definition a phonemic function, but of course the non-phonemic function will also be present if it occurs in the target speech norm. However, compensatory glottal stop used by Danish cleft palate speakers may be perceived as a Danish stød, since this sound is perceptually fairly similar to a

non-forceful glottal stop. And as a stød at the surface level has a phonemic function, there are risks of phonological mistakes. This may be the reason why child I, who uses the phonemic compensatory glottal stop, in most cases either omits the stød or signals it in a deviant manner: because stød in his speech will be understood as a compensatory glottal stop by the listener, it cannot be used also for the stød of the norm language. Furthermore, it is interesting that the non-phonemic function of glottal stop is also avoided in child I's speech, probably in order to avoid confusion with his phonemic use of the glottal stop. However, the [h] that he inserts before a word initial vowel may be interpreted as a different initiation of phonation, namely the 'soft' attack. Furthermore, his [h] instead of [ʁ] may be interpreted as a different initiation of phonation, namely the 'soft' attack with omission of /r/ rather than as the phonemic realization of /r/. The 'soft' attack [h] is probably understood as the phonemic /h/ by the listener. But the number of potential mistakes are in fact much less with [h] than with glottal stop, since [h] is the realization of only one phoneme, whereas the glottal stop is used for many phonemes in the target norm.

It should be added that the literature on language development delay in Danish speakers does not mention stød as a problem, and nor did any of the children with compensatory glottal stop from the 1970-78 status have any problems with stød, which fits very well with the fact that prosodic phenomena are less exposed to be a speech problem than segments. Thus, child I is certainly not representative of the population. However, his treatment of stød illustrates that it seems worthwhile to investigate the question whether cleft palate speakers react to the phonological consequence imposed by the cleft palate strategies. But again, due to want of phonetic and phonological descriptions of cleft palate speech in individuals, one cannot hypothesize which conditions will cause such reactions and how frequent the phenomenon is.

In conclusion, it has been pointed out that some characteristics in cleft palate speech may be observed in all cleft palate speakers, while others depend on the speaker's language. However, it has also become clear that phonetic descriptions of cleft palate speech with speakers of different languages are necessary prerequisites to the solution of the problem of universal and language dependent characteristics in cleft palate speech.

VI. CONCLUDING REMARKS

As quoted from Morris (1979) in the introduction, it may be considered impossible to make meaningful descriptions of cleft palate speech due to the great inter-individual variation in the cleft palate population. However, it seems meaningful to us to describe cleft palate speakers according to the origin and type of their deviating speech production. And such descriptions are useful for the assessment and treatment of these persons.

In practice there are three options for a person with persisting speech problems after primary palatal surgery: extra surgery, treatment with pharyngeal opturator, and speech therapy, sometimes in connection with the two first mentioned options. The advice as to which option to choose rests mainly on impressionistic analyses - as does also much research on cleft palate speech - partly due to the fact that it does not involve much equipment. However, a precondition for impressionistic analysis - in practice as well as in research - is a substantial knowledge about physiological, aerodynamic, and acoustic aspects of cleft palate speech to be related to the impressionistic percept. Many studies on physiological and aerodynamic aspects have been performed, while the number of acoustic studies is rather limited (see, e.g., Philips and Kent 1984). However, many questions as to the nature of the deviant types of speech production described are still unanswered. And other questions concerning cleft palate speech need further research such as the communicative consequences and the question of language dependency. Also, there remain the complications from the psycho-social constitution and background of each individual, and research in this field is evidently needed as well.

VII. NOTES

1. The categories "distortion", "substitution", "omission", and "addition" used to describe language development, are frequently used also in connection with cleft palate speech (excepting "addition"). These categories are ambiguously defined as to phonetic and phonological criteria.

2. A transcription system especially designed for cleft palate speech has been proposed by Vieregge & Jansen (1981).

3. Aspirated stop does not occur after *s*, in spite of the spelling *'ptk'*.

4. The symbol *ɹ̈* is used for the "soft *d*", which is a voiced approximant with a slight raising of the front and back of the tongue, and thus without any true constriction of the vocal tract. There is no symbol for this sound in the IPA-system.

5. [˚] indicates nasal friction, [*] that the sound is weak. These symbols are not included in the IPA-system.

6. For this purpose, syllable initial consonants preceding also a non-full vowel have been examined.

REFERENCES

Bergendal, B.-I. and Fex, S. 1977: "On glottal and pharyngeal articulation and nasality", *Scandinavian Journal of Logopedics and Phoniatrics 2/1*, p. 16-20

Bzoch, K.R., Kemker, F.J. and Wood, V.L. Dixon 1984: "The prevention of communicative disorders in cleft palate infants", in: *Speech and Language, Advances in Basic Research and Practice* (ed.: N.J.Lass), vol. 10, Academic Press, New York & London, p. 59-110

Clifford, E. and Clifford, M. 1979: "Psychological aspects of competence and language development", in: *Communication Disorders Related to Cleft Lip and Palate* (ed.: K.R. Bzoch), Little, Brown & Co., Boston, 2nd ed., p. 148-161

Demark, D.R. van 1974: "Assessment of articulation for children with cleft palate", *Cleft Palate Journal, vol. 11*, p. 200-208

Edwards, M. 1980: "Speech and language disability", in: *Advances in the Management of Cleft Palate* (eds.: Edwards, H. and Watson, A.C.H.), Churchill Livingstone, London, p. 83-97

Erp, A. van 1984: "A phonetic analysis of the speech of ten cleft palate speakers", *Proceedings no. 8, Institute of Phonetics, Catholic University, Nijmegen*, p. 48-59

Fletcher, S.G. 1978: *Diagnosing Speech Disorders from Cleft Palate*. Grune and Stratton, New York

Fox, D., Lynch, J. and Brookshire, B. 1978: "Selected developmental factors of cleft palate children between two and thirtythree months of age", *Cleft Palate Journal, vol. 15*, p. 239-245

Henningsson, G. and Isberg, A. 1985: "A cineradiographic study of velopharyngeal movements for deviant versus non-deviant articulation", *Phoniatric and Logopedic Progress Report No. 4*, Huddinge University Hospital, Karolinska Institute, Stockholm, p. 44-49

Hutters, B. 1985: "Vocal fold adjustments in aspirated and unaspirated stops in Danish", *Phonetica 42/1*, p. 1-24

Lawrence, C.W. and Philips, B.J. 1975: "A telefluoroscopic study of lingual contacts made by persons with palatal defects", *Cleft Palate Journal, vol. 12*, p. 85-94

Lynch, J.I. Fox, D., and Brookshire, B. 1983: "Phonological proficiency of two cleft palate toddlers with school-age follow-up", *J. Speech and Hearing Disorders 48*, p. 274-285

Matisoff, J.A. 1975: "Rhinoglottophilia: The mysterious connection between nasality and glottality", in: *Nasalfest*, (eds.: C.A. Ferguson, L.M. Hyman, J.J. Ohala), Dept. of Linguistics, Stanford University, Stanford, California, p. 265-287

Moll, K.L. 1968: "Speech characteristics of individuals with cleft palate", in: *Cleft Palate and Communication* (eds.: D.C. Spriestersbach and D. Sherman), Academic Press, New York and London

Morley, M. 1970: *Cleft Palate and Speech, 7th ed.*, Churchill Livingstone, Edinburgh

Morris, H.L. 1979: "Evaluation of abnormal articulation patterns", in: *Communication Disorders Related to Cleft Lip and Palate*, (ed. K.R. Bzoch), Little, Brown & Co., Boston

Ohala, J.J. 1972: "Physical models in phonology", in: *Proc. 7th Int. Congr. Phon. Sc.* (eds.: A Rigault and R. Charbonneau), Mouton & Co., The Hague, p. 1166-1171

Ohala, J.J. 1975: "Phonetic explanations for nasal sound patterns", in: *Nasalfest*, (eds.: C.A. Ferguson, L.M. Hyman, J.J. Ohala), Dept. of Linguistics, Stanford University, Stanford, California, p. 289-316

Pannbacker, M. 1975: "Oral language skills of cleft palate speakers", *Cleft Palate Journal, vol. 12*, p. 95-106

Philips, B.J. and Kent, R. 1984: "Acoustic-phonetic descriptions of speech production in speakers with cleft palate and other velopharyngeal disorders", in: *Speech and Language, Advances in Basic Research and Practice* (ed.: N.J. Lass), vol. 11, Academic Press, New York and London, p. 113-168

Shelton, R.L. 1979: "Oral sensory function in speech production and remediation", in: *Communication Disorders Related to Cleft Lip and Palate* (ed.: K.R. Bzoch), Little, Brown & Co., Boston, p. 120-148

Starr, D. 1979: "Dental and occlusal. Hazards to normal speech production", in: *Communication Disorders Related to Cleft Lip and Palate* (ed.: K.R. Bzoch), Little, Brown & Co., Boston, p. 90-100

Trost, J.E. 1981: "Articulatory additions to the classical description of the speech of persons with cleft palate", *Cleft Palate Journal, vol. 18*, p. 193-203

Vieregge, W.H. and Jansen, C. 1981: "Transcription of cleft palate speech": An experimental study", *Proceedings no. 3, Institute of Phonetics, Catholic University, Nijmegen*, p. 1-23

Vieregge, W.H. 1985a: "Transcription of normal speech: Intra- and inter-transcriber agreement", *Proceedings no. 9, Institute of Phonetics, Catholic University, Nijmegen*, p. 5-8

Vieregge, W.H. 1985b: *Transcriptie van spraak: Theoretische en praktische aspecten van de symboolfonetiek*, Foris Publications, Dordrecht.

APPENDIX

CONNIE LOWE NIELSEN*, WENDY COHN*,
AND KIRSTEN BRØNDSTED**

A characteristic feature of cleft palate speech is the tendency to retract the articulation of obstruents normally produced in front, of to a place below the velopharyngeal valve as a compensatory obstruction of the airflow, resulting in glottal stops and pharyngeal fricatives. This deviant pronunciation is often very damaging to intelligibility - because each of the compensatory sounds is often used as realization of several phonemes, whereby the number of contrasting sounds is reduced. Also, this type of deviant pronunciation has proved very difficult to correct, even after surgical normalization of the velopharyngeal mechanism (see, e.g., Bzoch 1979). In the status mentioned above of all individuals with cleft palate from East Denmark, born during the period 1970-78 (N=293), about 10% were found to use glottal stop articulation in their speech at three years of age, and this was only reduced to 7% at six years of age, in spite of logopedic intervention. It would therefore be essential to be able to predict which infants are likely to use compensatory sounds so that prophylactic intervention could be initiated. And since glottal babbling before surgical closure of the palate is sometimes observed (Olson 1965, Henningsson 1981), this intervention could be initiated very early.

From clinical experience we venture the hypothesis that these children have certain features in common: extrovert personality, eagerness to communicate and to be noticed socially, and earlier initiation of language than usual for most cleft palate infants. It was therefore relevant to see if the children in this study who are using glottal stops show features in their communication behaviour which differ from the children who use other "strategies".

From the video-recordings used in the phonetic analysis the communicative interaction between parent and cleft palate child was analyzed. It must be stressed, however, that because of the very limited number of children in the study and also because of the limitations in which the experimental setting was created (see p. 3), the results should only be viewed as a guideline for more comprehensive studies.

The method used for this pilot study is descriptive and derives from an etiological frame of reference (Blurton Jones 1979, Nielsen and Damholt 1982). The first phase consists of describing interaction between the child and parent through a rough analysis of the video recordings. Throughout this phase we did not operate with any specific hypothesis. The

---

*) Psychologist, M.A.
**) Institute for Speech Disorders, Hellerup.

second phase of the analysis required a detailed description of specific elements in the parent-child interaction, and how these elements interact and influence each other. This requires a frame by frame analysis which could not be carried out in this study.

It was not possible, from the rough analysis alone, to determine any definite impressions of the children's behaviour or character, since the experimental setting limited free and natural interaction between the parent and the child. However, it was observed that the children using glottal stops seemed to use their voice more accompanied by gestures, when they lacked words, and they also used more onomatopoeia than the other children.

On the basis of the observations of parent-child interaction a hypothesis could be formulated: The children with glottal stops are exposed to greater expectation regarding their communicative competence than the children who do not use glottal stops.

Communicative competence is defined as the child's knowledge of and ability to use language (Hymes 1971). Thus, it can be considered as a process involving interaction between impressive and expressive language.

An attempt to verify the above hypothesis was made by examining the dialogue between parent and child. A categorization similar to Newport et al.'s (1977) method of scoring Mother's speech into sentence types was used. The sentence types are the following:

Sentence types	
Declarative	You can sing a song.
yes-no question	Can you sing a song?
Imperative	Sing a song.
Wh-question	What can you sing?
Deixis	That's a dog.

Thus, the first fifty utterances of each dialogue in the study were categorized in order to study the differences in the ways in which the parents communicated with their children. In figure 6 is shown examples from the dialogues.

In example 1 the parent's sentence types are the following: yes-no questions - declarative - deixis, in the parent's first utterance. In example 2 the following combination is seen: imperative - declarative - yes-no questions - imperative - imperative, all in the parent's first utterance. In examples 3 and 4 the parents only use one sentence type for each utterance.

The demand on the child to understand the message seems to be greater in ex. 1 and 2 than in 3 and 4. The frequency of the parent's use of successive sentences of different structure for the first 50 parent utterances in each recording was found to be: child I: 36%, II: 36%, III: 6%, IV: 2%, and V: 8%. These results show a significant difference in the ways in which parents communicate with their children. This shows that the parents of child I and child II seem to have greater expectations of their children's communicative competence than the other parents.

Child I and child II were the same children who showed predominant use of phonemic glottal stops in their speech. From clinical and parent information it appears that child I and child II used sentences several months earlier than the other children in the study, and that they used glottal babbling. According to the linguistic analysis child IV used glottal stops as well but in an atypical manner (see p.20).

The observations and the analysis carried out by the psychologists have led to many new questions concerning the communicative behaviour of children with glottal stops, as well as their parents' behaviour. Therefore, before any definite conclusions can be made concerning the findings in this study, further studies must be conducted.

It would be essential to study how parents of children with glottal stop communicate with their children during the years of language acquisition. The parents of children who use glottal stop may also have a significant role in their child's habit, since they may expect the child to understand their long and complicated sentences. They may also expect that their child has the ability to communicate with other, non-linguistic forms of vocal utterance, such as onomatopoeia and sounds accompanying gestures.

Perhaps better insight into this process could show how a child's communicative competence can influence his/her parents' communicative behaviour. Such a study could also include a closer examination of whether a child's eagerness to speak is expressed when he/she first begins using words and sentences, and also whether the language acquisition occurs at a different rate than in other children with cleft palate. Finally, the parents' expectations of their child's communicative competence could be studied through interviews, observations, and analyses of their communication.

To be able to study these suggestions properly, the experimental conditions should be different from those of the present pilot study, which was primarily set up for the linguistic analysis. A study of this kind should include a detailed analysis of the children's spontaneous communication and behaviour in a more natural setting, with an emphasis placed on the behaviour occurring during the communicative dialogue. The number of children should be sufficient, including a control group of normal children and their parents.

The children should be as young as 1-2 years, which is the period in which glottal stops may change from a non-phonemic status in babbling into a phonematic function in emerging language. Also, the rate of language acquisition should be studied. A follow-up study some years later could perhaps be compared to the results of the present observations on 4-5 year old children.

REFERENCES

Blurton Jones, N. 1972: *Ethological Studies of Child Behaviour.* Cambridge

Bzoch, K.R. 1979 (2nd ed.): "Etiological Factors related to Cleft Palate Speech", in: *Communication Disorders Related to Cleft Lip and Palate* (ed.: K.R. Bzoch), Little, Brown & Co., Boston, p. 67-77

Henningsson, G. 1981: "Terapi vid glottalt joller hos läpp-käk-gomspaltbarn", *Scandinavian Journal of Logopedics and Phoniatrics 7/2*, p. 91-94

Hymes, D. 1971: "Competence and Performance in Language Theory in: *Language Acquisition, Models and Methods* (eds.: Huxley, J.R. & Ingram, E.), New York, p. 3-28

Newport, E., Gleitman, H., Gleitman, L.R. 1977: "Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style", in: *Talking to Children* (eds.: Snow, C.E. and Ferguson, C.), Cambridge/London

Nielsen, C.L. and Damholt, H. 1982: *Samtale med et spædbarn (Conversation with an infant)*, Unpublished thesis, University of Copenhagen

Olson, D.A. 1965: *A Descriptive Study of the Speech Development of a Group of Infants with Unoperated Cleft Palate.* Unpublished Ph.D. dissertation, North Western University.

Ex. 1

P:	Make flowers? - But it's not yellow all of it - they are green.		P:	Yes, green to begin with - and then a red flower		P:	hmm, a green one like that, and then a red flower - now you're making it wrong
		C:	Where's green - oh yes, green		C:	Oh yes - green	

Ex. 2

P:	I think you ought to take your boots off - you're not sitting very comfortably, are you? - Yes, boots off - and put them next to you		P:	Oh yes, it's just like the one we just saw on TV about water - what can you use water for - can you remember?
		C:	Look, a little watermill	

Ex. 3

P:	Look here		P:	What's that?		P:	What is it getting to drink?	
		C:	A hat		C:	Some clothes		C: milk

Ex. 4

P:	What are you going to draw first?			P:	First a house, yes, and then you make the wall		P:	Hm-m
		C:	I'm drawing a house			C:	This is inside, and that's outside	

Figure 6

The figure shows examples from the child-parent dialogues (in translation). P = parent, C = child.

ARIPUC 20, 1986, p. 35-44

SENTENCE INTONATION IN TEXTUAL CONTEXT - SUPPLEMENTARY DATA

NINA GRØNNUM THORSEN

The experiments presented here follow up a previous investigation (Nina Grønnum Thorsen "Intonation and text in Standard Danish", J. Acoust. Soc. Am. 77, 1985, 1205-1216) and were designed to show whether a sequential fundamental frequency lowering of individual sentence components is present in a semantically, but not syntactically, coherent sequence (a text), when the number of sentences exceeds three. The results show that such a sequential lowering may appear, though it is not evenly distributed across the text. However, the textual intonation contour is sensitive, not only to the number of sentences that make up the text, but also to the length of individual sentence components.

I. INTRODUCTION

In a previous investigation of intonation in texts in Standard Danish (Thorsen, 1985) it was established that each declarative sentence in a read text is associated with its own declining intonation contour, but together two or three such contours describe an overall slope. Individual sentence contours are steeper in a succession of declarative terminal sentences than in a corresponding string of coordinate main clauses. However, since the texts contained no more than three sentences, the experiment left unresolved the question whether the sequential lowering of individual sentences observed should be ascribed to a particular initial and final effect (text initial and final sentences being higher and lower, respectively, than medial items which under that hypothesis are not differentiated among themselves), or to a truly sequential lowering (evenly) distributed over all the sentences in a text. The experiments reported below were designed to address this issue,

and, furthermore, to see what effect, if any, changing the length of individual sentence components would have.

II. METHODS

10 declarative sentences were made up; 4 sentences with three prosodic stress groups in each (stressed vowels are indicated orthographically with acute accents here):

1. Káren skal séjle ved Sámsø. (Karen will be sailing round Samsø.)
2. Sánne skal på cámping i Skótland. (Sanne is going camping in Scotland.)
3. Kálle skal vándre i Fránkrig. (Kalle is going to trek in France.)
4. Thómas spiller skák i Cánada. (Thomas plays chess in Canada.)

4 sentences with two prosodic stress groups each:

5. Káren skal til Sámsø. (Karen is going to Samsø.)
6. Sánne skal til Skótland. (Sanne is going to Scotland.)
7. Kálle skal til Fránkrig. (Kalle is going to France.)
8. Thómas skal til Cánada. (Thomas is going to Canada.)

2 sentences with four prosodic stress groups:

9. Káren skal séjle med sin fár ved Sámsø. (Karen will be sailing with her father round Samsø.)
10. Sánne skal på cámping med Kálle i Skótland. (Sanne is going camping with Kalle in Scotland.)

The stressed vowels are all low and the initial consonant is either an aspirated stop or an unvoiced fricative, except that in 'Skótland' the stop is unaspirated, and in 'Fránkrig' a devoiced [ʁ] intervenes. This will minimize differences in segmental effects on the stressed vowels (from which the lower lines derive in the figures that follow). Segmental effects on fundamental frequency (F_0) are smaller in unstressed syllables (from which the upper lines derive), though they are not completely negligible, cf. Reinholt Petersen (1979, 1980).

The following combinations of sentences constituted the texts[1] to be recorded:

1	2	4	3
2	1	3	4
3	4	1	2
4	3	2	1
5	6	7	
5	6	7	8
9	10		

Note that each three-stress sentence occurred in all four positions in the four-sentence texts. Sentences no. 1, 3, 5, 7, 9, and 10 occurred in isolation as well. These 13 items were randomized 6 times to make 6 pages of reading material, each page beginning and ending with two filler sentences. Two subjects from the previous investigation (Thorsen, 1985: NRP (male) and NT (the author)) recorded the material in one session lasting about 20 minutes.

The recordings were made with a Revox A-77 tape recorder and a Sennheiser MD 21 microphone in a quiet room. The tapes were processed with hardware intensity and fundamental frequency meters (F-J Electronics) and registered on an Elema 800 Mingograph at a paper speed of 100 mm/s. Measuring accuracy was 1 Hz for the male and 2 Hz for the female speaker, and 0.5 cs in the time dimension.

The first and last local low F_0 point in each sentence (which occurs at or near the offset of the first and last stressed vowel) and the first and last local high F_0 (which occurs in the first post-tonic syllable after the first and last stressed syllable) were measured. The connection of these two highs and lows, respectively, make up the upper and lower lines in the graphs to follow. The distance in time of each of these four measuring points from the onset of the first stressed vowel in each sentence was likewise measured. F_0 and time measurements were averaged over the six repetitions of each item by each speaker. Average F_0 values were converted to semitones (re 100 Hz) and average tracings drawn. Furthermore, an average over the four texts with four three-stress sentences was calculated. For a more detailed account of the procedures, see Thorsen (1985).

The use of only two speakers was justified by the high qualitative agreement across the four speakers in Thorsen (1985). The simplified representation of the data in terms of straight upper and lower lines derived from initial and final points in the tracings is justified by the generally valid, straight line approximation to intonation contours in short Danish sentences like the present ones, cf. e.g. Thorsen (1978, 1985).

The qualitative agreement between the two speakers was high, and a common mean could be calculated. Only those means are presented here. Both speakers generally paused between the sentences in a text. Pause durations range between 20 and 80 cs, averaging around 40 cs, with a tendency to increase through the text. The figures do not depict pause durations.

III. RESULTS

A. ISOLATED SENTENCES

Figure 1 presents upper and lower lines of isolated sentences containing two prosodic stress groups (dotted lines), three prosodic stress groups (dashed lines), and four prosodic stress

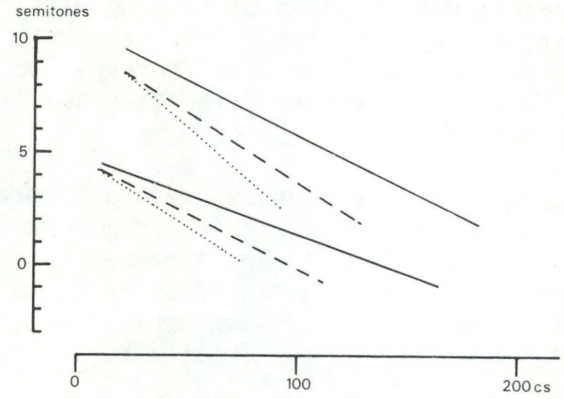


Figure 1

Upper and lower lines in three isolated terminal declarative sentences containing two (dotted lines), three (dashed lines), and four (full lines) prosodic stress groups. Each type represents the mean over four different sentences and two speakers, i.e. N=8. The onset and offset of upper and lower lines are derived from the first and last local high and low point in the F_0 tracings. Zero on the logarithmic frequency scale corresponds to 100 Hz.

groups (full lines). These utterances only confirm what was found in previous investigations (Thorsen 1980, 1981, summarized in Thorsen 1983), that increasing the length of a terminal declarative sentence will affect beginning points (slightly), end points and slope. Beginning points rise slightly, end points lower, and slope gets less steep as the number of stress groups increase. (Increasing the length beyond four stress groups, however, will lead to a resetting of the intonation contour, cf. Thorsen 1983.) For a further discussion of these findings and references to the literature on other languages, see Thorsen (1980).

B. MULTI SENTENCE TEXTS

Figures 2, 3, and 4 present upper and lower lines in the various texts.

1. Sentences with two prosodic stress groups - Figure 2:

In both texts, with three (dashed lines) and four (full lines) sentences, respectively, a clear overall textual slope can be discerned, which is in accord with the previous results (Thorsen 1985). Furthermore, the two medial components in the four-sentence text seem to be differentiated, i.e., the third sentence is lower than the second one, notably in the upper line.

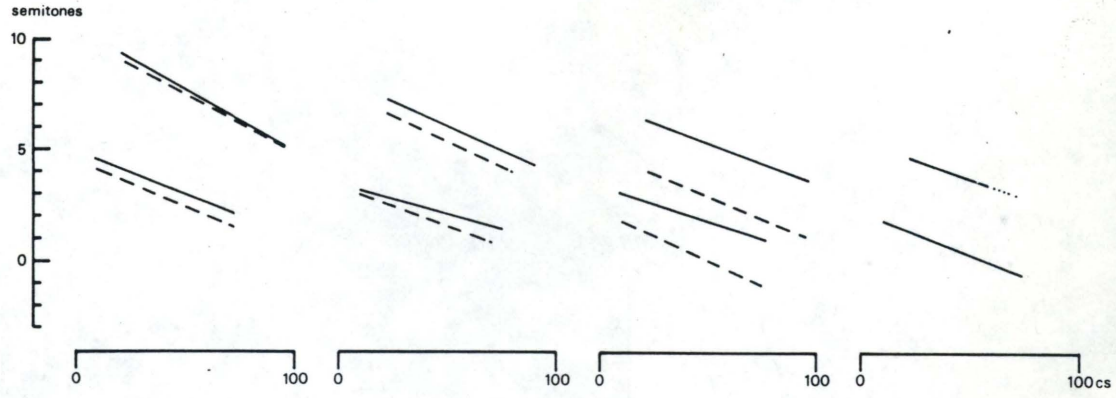


Figure 2

Upper and lower lines in two texts, containing three two-stress sentences (dashed lines) and four two-stress sentences (full lines). Average over two speakers. The upper line offset in the last sentence of four could not be measured with the male speaker. See further the legend to figure 1.

Since the two speakers speak in different ranges, standard deviations of their common means are very great, and statistical significance between such means is not to be expected.[2] However, with the female speaker, the difference between upper line onset and offset values, in the second and third sentences, turn out to be statistically significant (student's t-test) at the 0.05 level of confidence, whereas it does not reach statistical significance with the male speaker. - Simultaneously we should note, though, that the two medial sentences seem to cluster, i.e. the lowering of sentences is not evenly distributed across the whole text.

2. Sentences with 3 prosodic stress groups - Figure 3, full lines:

Here, too, an overall downward slope is discerned, but the slope levels out over the two medial components, which appear to be almost exactly identical as far as upper and lower line onsets and offsets are concerned, in contradistinction to the two-stress sentences (dashed lines).

3. Sentences with 4 prosodic stress groups - Figure 4 - will be dealt with in the discussion section below.

IV. DISCUSSION

Figure 5 summarizes Figures 1-4 and facilitates a comparison of the various texts. The isolated sentences, which are in a sense both text initial and text final, combine the high (upper and lower line) onsets of initial sentences with the low (upper and lower line) offsets of final sentences in multi-sentence

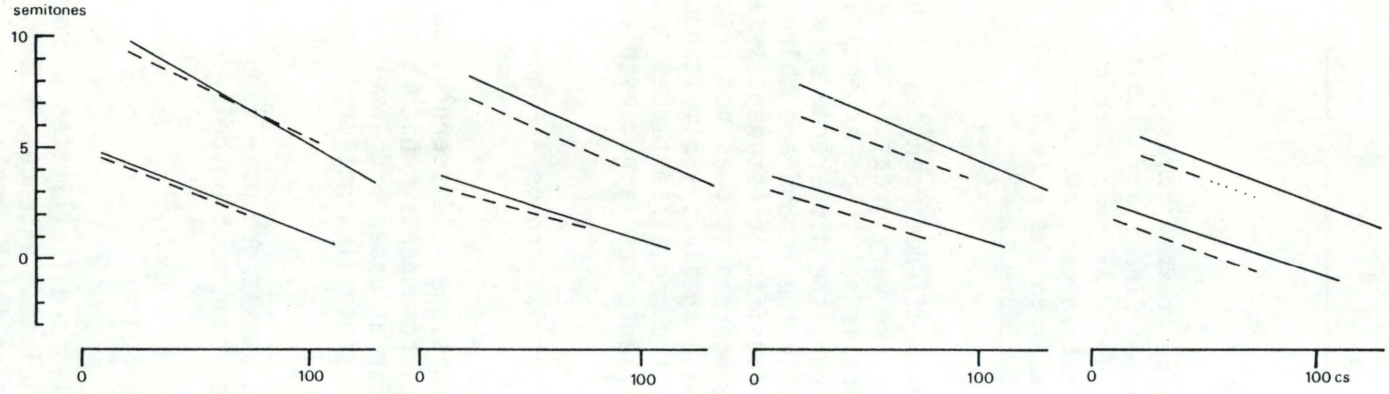


Figure 3

Upper and lower lines in a text with four three-stress sentences (full lines), averaged over four texts and two speakers, i.e. N=8. The dashed lines are a repetition of the full line tracings in figure 2. See further the legend to figure 1.

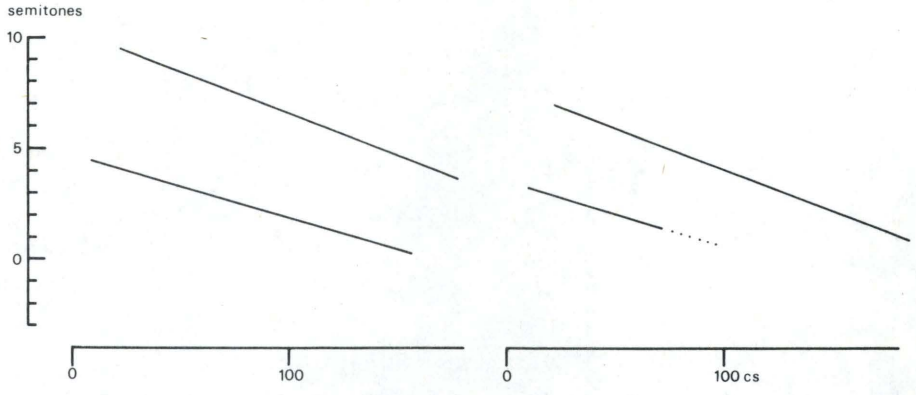


Figure 4

Upper and lower lines in a text containing two four-stress sentences. Average over two speakers. See further the legend to Figure 1.

texts; except that isolated sentences with two stress groups have offset values that do not quite reach the lower limit set by text final sentences. This is in complete agreement with the previous results (Thorsen 1985).

The most salient features of multi-sentence texts are: Onset values (whether of upper or lower lines) are fairly constant across texts of different length, both in terms of number of sentences and number of stress groups in each sentence. These onset values vary within a range of one semitone in the upper lines, slightly less in the lower lines. Considering the range of sentence onset values across a text, which amounts to about 5 semitones in the upper lines and 2.5 semitones in the lower lines, it appears justified to talk about a near-constancy in F_0 onset across texts of different length. Likewise, offset values in text final sentences appear nearly constant. In other words, the total range spanned by any given text is independent of its length, at least within the range of text lengths investigated in this material (and thus the overall slope is inversely proportional to text length). But within this upper and lower limit, set by the first and last prosodic stress group in a text, sentences do distribute themselves differently according to their number and length. A sequence of three sentences (with two stress groups in each sentence) exhibit a smooth and linear fall in onset as well as offset values of both upper and lower lines. This compares well with the previous findings (Thorsen 1985). A text with two sentences with four stress groups in each has its second - and final - component beginning at a considerably higher onset value than any other text final component in this material, presumably in order to preserve a sentence intonation slope across its four stress groups which is steep enough to suit a text final terminal declarative sentence (without falling through the floor of the speaker's F_0 range).

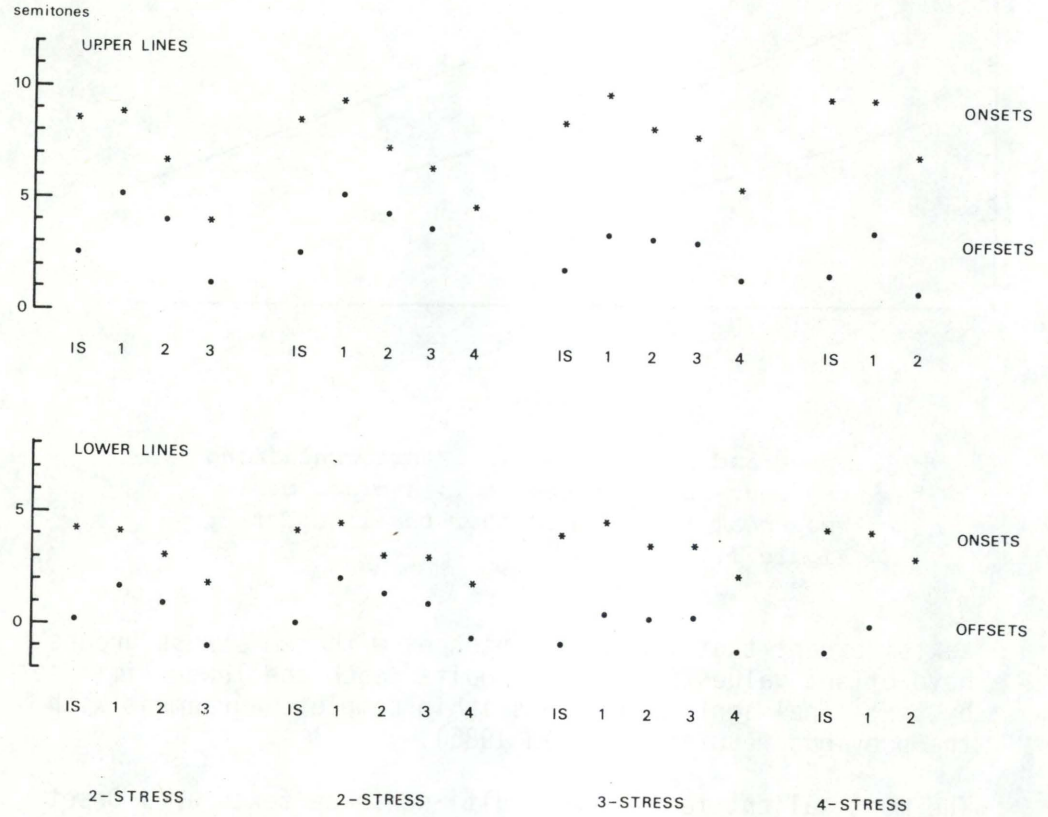


Figure 5

Frequency values of onsets (stars) and offsets (dots) of the upper lines (above) and lower lines (below) in Figures 1-4. The two leftmost sets of data pertain to two-stress sentences in isolation ("IS", derived from Figure 1) and in each of the successive positions in a text with three and four components, respectively (derived from Figure 2), as indicated below each graph. Similarly, the two rightmost sets of data pertain to sentences with three and four stress groups, in isolation (derived from Figure 1) and in a text with four sentences (derived from Figure 3) and with two sentences (derived from Figure 4), respectively.

Another instance of plasticity in the production of sentence and text intonation is offered by the somewhat different treatment of texts with four sentences, where the sentences have two and three prosodic stress groups, respectively. Onset values are higher, and offset values are lower (in upper as well as lower lines) in non-final sentences with three stress groups. This may, again, be seen as a consequence of sentence length, simply: the longer sentences will have higher onsets and lower offsets to preserve a sentence intonation slope suitable for terminal declarative sentences - just as is the case for isolated sentences of varying length, cf. Figure 1. The demand for a certain sentence intonation slope may also explain why a succession of two medial three-stress sentences are produced without any "downstep", in contradistinction to a succession of medial two-stress sentences, cf. Figure 3. If the longer sentences are to preserve a suitable slope, there is no room for a sequential lowering of medial components.

In Figure 2 there is a difference (notably in the upper lines) in the second component in the two texts: it is lower when only one sentence succeeds (dashed line) than when two sentences follow (full line). This difference persists, though diminished, if the two texts are aligned on the frequency scale to have identical upper line onsets. However, it turns out that with one of the speakers there is virtually no distinction between upper (or lower) lines in these sentences, and with the other subject the difference does not quite reach statistical significance. Thus, the evidence is not conclusive but seen together with the remainder of the present data and that published in Thorsen (1983, 1985), it does lend some support to the notion of look-ahead and preplanning in the production of intonation. For a further discussion of such mechanisms and their implications for intonation theory, see Thorsen (1985).

V. NOTES

1. I employ "text" here to mean a sequence of semantically coherent but not necessarily syntactically coordinated sentences.
2. Standard deviations could have been considerably reduced if the data had been normalized, e.g. by a zero-line offset adjustment. However, there are theoretical and practical problems in such a normalization that still await a solution and since, furthermore, it is not really essential to the points I want to make, no normalization has been attempted.

REFERENCES

Reinholt Petersen, N. 1979: "Variation in inherent F_0 level differences between vowels as a function of position in the utterance and in the stress group", *Ann. Rep. Inst. Phon., Univ. Copenhagen 13*, p. 27-57

Reinholt Petersen, N. 1980: "Coarticulation of inherent fundamental frequency levels between syllables", *Ann. Rep. Inst. Phon. Univ. Copenhagen 14*, p. 317-354

Thorsen, N. 1978: "An acoustical analysis of Danish intonation", *J. Phonetics 6*, p. 151-175

Thorsen, N. 1980: "Intonation contours and stress group patterns in 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 in declarative sentences of varying length in ASC Danish - Supplementary data", *Ann. Rep. Inst. Phon. Univ. Copenhagen 15*, p. 13-47

Thorsen, N. 1983: "Standard Danish sentence intonation - Phonetic data and their representation", *Folia Linguistica 17*, p. 187-220

Thorsen, N. Grønnum 1985: "Intonation and text in Standard Danish", *J. Acoust. Soc. Am. 77*, p. 1205-1216.

ARIPUC 20, 1986, p. 45-66

ACOUSTIC-PHONETIC INVESTIGATIONS OF THE INTONATION OF GERMAN AND DANISH: SIMILARITIES AND DIFFERENCES*

ROBERT BANNERT**
AND
NINA GRØNNUM THORSEN

INTRODUCTION

The past ten years have witnessed a growing interest in prosodic studies - phonological as well as phonetic ones. Today the emphasis is on the construction of theories and on empirical investigations, which interact and lead to models of prosody.

We present here the results of instrumental phonetic investigations of German and Danish intonation and account for the prosodic categories required in the description. This is followed by a comparison of the two languages, insofar as the prosodic categories and their manifestation is concerned. The presentation is descriptive rather than arguing. Discussions about intonation theories and models can be found in Bannert (1982) and Thorsen (1983a, 1983b, 1985a, 1985b). Our purpose here is to present to the reader the basic structures of German and Danish intonation and point out the more important similarities and differences. Such descriptions and comparisons may serve as point of departure for further comparative studies and for foreign language teaching. We would like to stress the fact at the outset that the research we each of us have conducted was not intended for comparative (Danish-German) purposes, and this paper thus represents a "rationalization after the fact". That also accounts for the somewhat different disposition of the two sections, which excludes a direct comparison of every aspect treated in Danish and German, respectively.

---

*) *Translation of a paper 'Empirische Studien zur Intonation des Deutschen und Dänischen: Ähnlichkeiten und Unterschiede', to appear in Kopenhagener Beiträge zur germanistischen Linguistik, 1986.*

**) *Department of Linguistics and Phonetics, University of Lund, Sweden.*

TERMINOLOGY AND PROCEDURES

By intonation we mean speech melody, i.e. the fundamental frequency (F_0) variation or pitch course as it appears in F_0 curves. Other phenomena which in linguistic tradition are occasionally encompassed by intonation such as rhythm and voice quality, will not be dealt with here.

Our results derive from systematic, empirical and experimental acoustic investigations. Such investigations have certain intrinsic limitations. They often deal with sentences which have been read by speakers in sound treated rooms. Such is also the case here. The material is often composed of severely limited and manipulated utterance types. Such a procedure may seem inappropriate in view of the fact that the final goal is a description of the intonation of spontaneous speech. However, our methods may be defended on at least two grounds. Firstly, it is easier to investigate the course of Fo in syntactically and pragmatically simple structures which have been produced under controlled circumstances. Secondly, you may expect that natural, spontaneous speech can be described with the same categories and prosodic structures as manipulated, read sentences. In other words, the intonation of non-manipulated, free speech will later be accounted for with the same descriptional devices as the controlled speech situation. The description of intonation in both languages is based on sentences of varying length and encompasses declarative as well as interrogative utterances.

THE BASICS OF DANISH INTONATION

The model in figure 1 derives from analyses of fairly short, simple sentences, read aloud by speakers of Copenhagen Standard Danish. The pragmatically neutral context as well as instructions to the speakers were intended to elicit neutral speech, i.e. speech which is not characterized by any particular emotions or emphases.

The model depicts the structure of intonational phenomena as a hierarchically organized system, where components of smaller temporal scope are superposed on components of larger temporal domain. In other words: global tendencies with more local modifications. For Standard Danish the following components can be discerned:

The text contributes an overall textual contour.
The sentence yields a sentence intonation contour.
The prosodic phrase adds a phrasal contour.
The prosodic stress group[1] contributes a stress group pattern.
The "stød" may involve a tonal modification on the stress group pattern.
Finally, individual segments have intrinsic Fo characteristics (the microprosodic component).

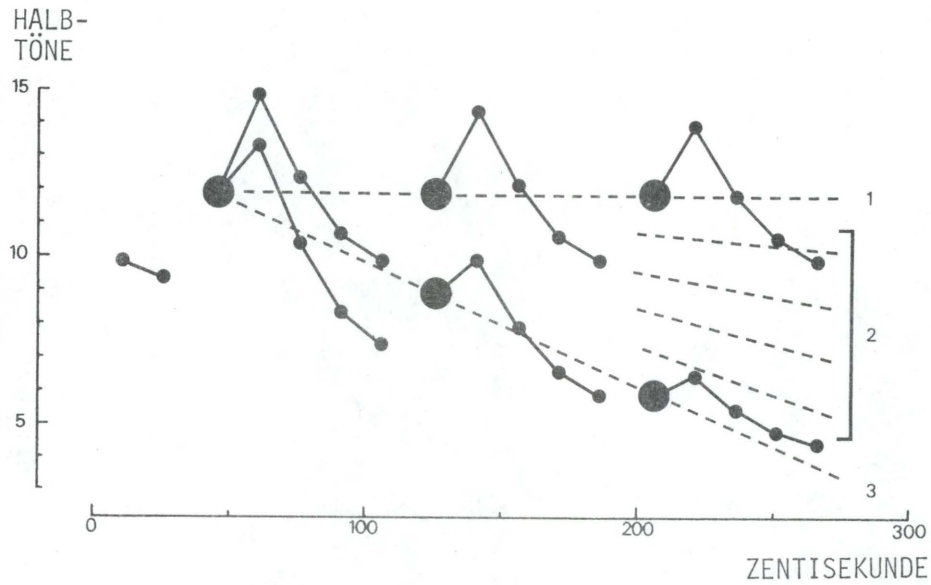


Figure 1

A model for the course of fundamental frequency in short sentences in Standard Danish. (1) Syntactically unmarked questions. (2) Questions with word order inversion and/or interrogative particle; non-final declarative and interrogative clauses. (3) Terminal declarative statements. Large points denote stressed syllables, small points depict unstressed syllables. Full lines represent the Fo pattern of prosodic stress groups; broken lines indicate the sentence intonation contours.

Figure 1 only depicts the sentence and stress group components. If the utterance exceeds a certain length, the intonation contour is decomposed into shorter, falling phrasal contours, which together describe an overall falling contour (figure 2, see also Thorsen 1983a, 1983b). Sentences are coupled together into coherent texts in a similar fashion. Each sentence has its own intonation contour, but together these sentences describe an overall falling course (see figure 3 and also Thorsen 1985a, 1985b). In the following, only the sentences and stress group components are dealt with.

The prosodic stress group

Basbøll (1977) defines a syntactic stress group in Danish as a group of words with one main stress which lies on the last word in the group (with certain exceptions: personal pronouns are unstressed also in final position). The prosodic stress group, on the contrary, consists of a stressed syllable and all succeeding unstressed syllables, i.e. the prosodic stress group boundary, or foot boundary, lies immediately before the stressed

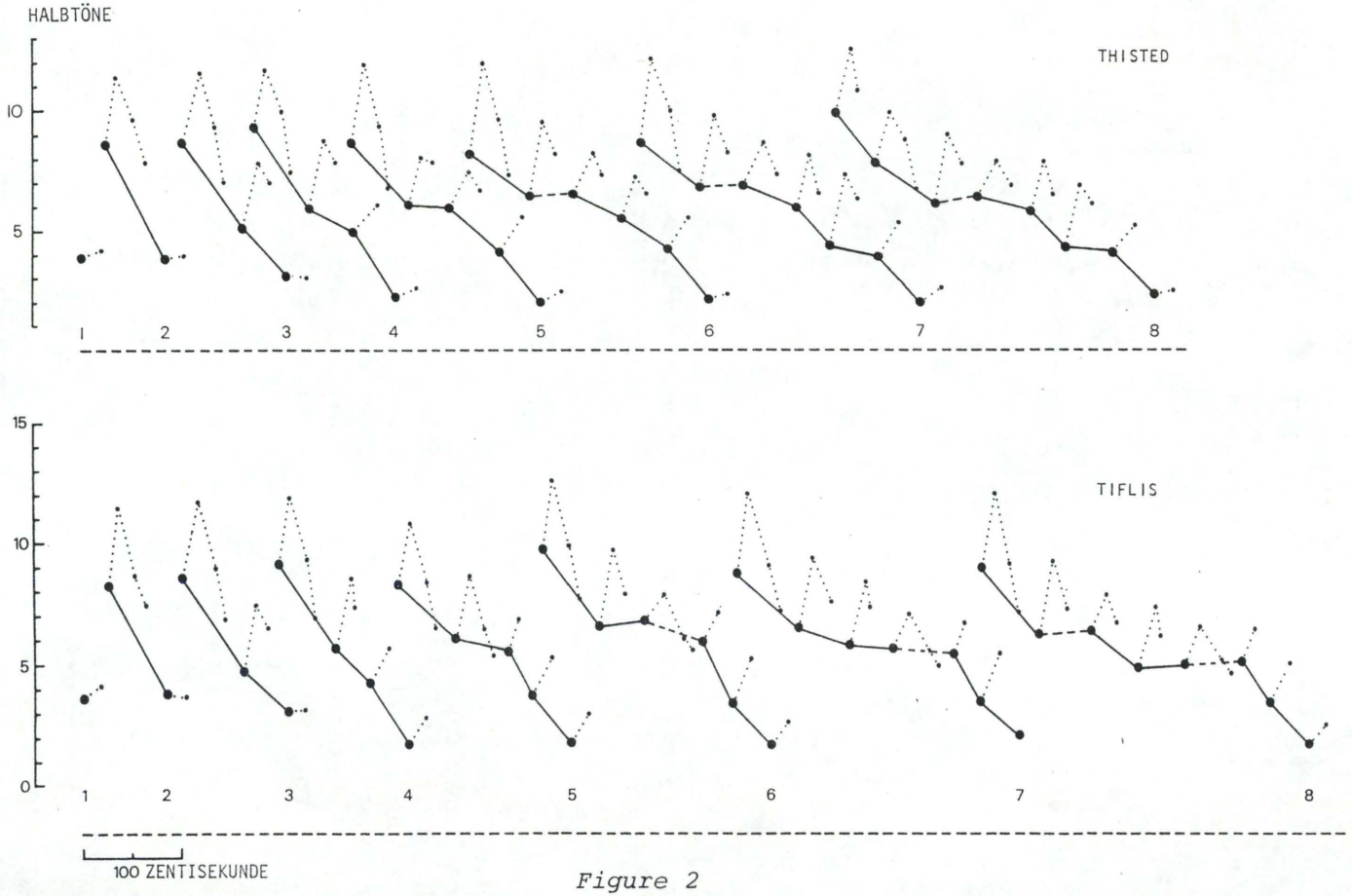


Figure 2

Intonation contours (full/broken lines) and stress group patterns (dotted lines) in two groups of terminal declarative utterances, consisting of one to eight prosodic stress groups. Average over four speakers. Large points denote stressed syllables, small points unstressed syllables. Broken lines denote the boundaries between prosodic phrases.

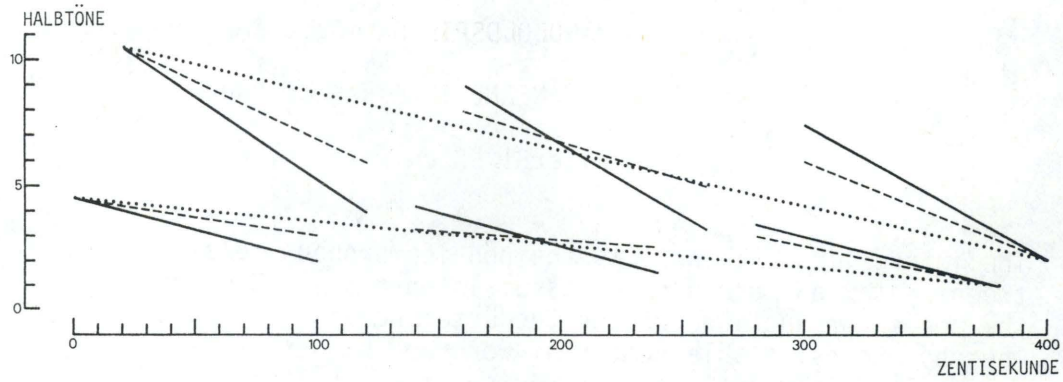


Figure 3

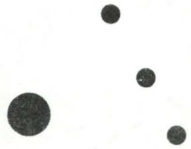
Stylized model of the course of Fo in texts with three terminal declarative sentences (full lines) and three coordinate main clauses (broken lines). The lower lines connect the stressed syllables, i.e. they represent the intonation contours (cf. figure 1), the upper lines connect the first unstressed syllable in each stress group. The dotted lines represent the overall slope through the text. A text with two components is arrived at by leaving out the middle component. An isolated sentence arises when the dotted lines are suitably compressed in time.

syllable, independently of the number and type of syntactic boundaries in the sentence. The parsing of an utterance into prosodic stress groups can be illustrated with the following example (/ depicts the boundaries between two main clauses and between verb phrase and complement within each clause; + denotes prosodic stress group boundaries, and ´ marks the stressed syllables):

Han lagde sig / på chaiselongen / og tændte / en Caminante.

Han + lágde sig på chaise + lóngen og + téndte en Cami + nánte.

(He lay down on the sofa and lit a Caminante (a cheroot).)
The definition of the prosodic stress group as a stressed syllable plus all succeeding unstressed syllables (if any) within the same intonation contour derives from the analysis of tonal patterns. The prosodic stress group is the carrier of a recurring and fairly constant Fo contour consisting of a (relatively) low stressed syllable followed by a high-falling tail of unstressed syllables. Thus, the three capitalized sequences below are all realised with the same Fo pattern:



HÅNDBOLDSPIL ER méget ánstrengende

baNÁNERNE I kássen er rådne

allerGÍ ER EN INfám sýgdom

The interesting fact here is the non-isomorphous relation between syntax and prosody (in casu: intonation). Somewhere in the speech production process the utterance is re-structured, whereby a considerable number of word and higher syntactic boundaries are deleted in the course of Fo. (This is true at least of the type of monitored and fluently read speech which is the basis for the analysis.)

If tonal relations are accepted as criteria for locating boundaries, it is clear that the prosodic stress group cannot, e.g., end with the stressed syllable: the tonal relation between a stressed syllable and the preceding unstressed one is highly variable, as is apparent from figure 1. The preceding syllable may be higher than, on a level with, or lower than the stressed syllable, depending on how many unstressed syllables precede.

Stress group pattern variation

The prosodic stress group pattern is subject to a certain quantitative variation, depending on

1) its position in the utterance (the magnitude of the rise from stressed to post-tonic syllable decreases from beginning to end)
2) the intonation upon which the pattern rides (the rise is higher on less falling contours, cf. figure 1).

Furthermore, the number of unstressed syllables in a prosodic stress group may vary between zero and rather large numbers, which naturally influences both the Fo pattern and the time interval between the stressed syllables. If there are no unstressed syllables in the stress group, there is no material on which the pattern can rise (and fall), so the Fo pattern is truncated (rather than being compressed in time to be contained within the single, stressed syllable). Finally, there is an interspeaker variation in the magnitude of the low-to-high interval and in the steepness of the slope of the falling unstressed syllables. See further Thorsen (1984).

The slope of the intonation contour

When the stress group pattern is a recurrent, predictable, qualitatively constant LOW + HIGH-FALLING pattern, the intonation contour can be defined in terms of the stressed syllables alone. That is not to say that the course of the unstressed

syllables is irrelevant for the identification of intonation contours, only that they are redundant in the strict sense of the word (see further Thorsen 1980a). In short utterances the intonation contours approach straight lines whose slopes vary in close correlation with the type and function of the utterance, as shown in figure 1: terminal declarative utterances have the steepest slopes, syntactically and lexically unmarked questions have horizontal contours. In between we find other types of questions and non-terminal sentences, with a tendency towards a trade-off between syntax/lexicon and intonation: the more syntactic or lexical information about the interrogative or non-terminal function of the sentence, the more falling, i.e. the more declarative-like, is the intonation contour, and vice versa. Furthermore, it appears that, just like in German, the tonal course is the same in utterances with identical prosodic structure, independently of syntactic constituents. See also above under "The prosodic stress group".

Longer utterances

If a declarative utterance contains more than three or four stress groups, most speakers will decompose the contour into several, shorter phrase contours, each with its own declination - which together describe an overall falling slope (figure 2). The boundaries between prosodic phrases in an utterance are related, though not directly subordinated, to syntactic boundaries. The relationship between utterance length, prosody, syntax and semantics is rather complex. See further Thorsen (1983a).

Global signalling of sentence intonation

The description of intonation above implies that sentence intonation is signalled globally rather than locally, i.e. the difference between, e.g., a declarative and interrogative sentence does not reside in a special movement at the end of the sentence but is distributed over the whole utterance. In this matter, Danish is different from most of the related Germanic languages.

Sentence accent

Danish displays yet another peculiarity: it lacks an obligatory sentence accent, or focus, or nucleus. In pragmatically neutral speech all stressed syllables have the same weight or prominence. An extra prominence somewhere in the utterance is not present acoustically, nor perceptually. Pragmatically and prosodically neutral utterances are neither incomplete nor unnatural. Even if they do not occur very often in spontaneous speech, they are in no way unnatural and they are very easy to elicit in recordings. See further Thorsen (1983b).

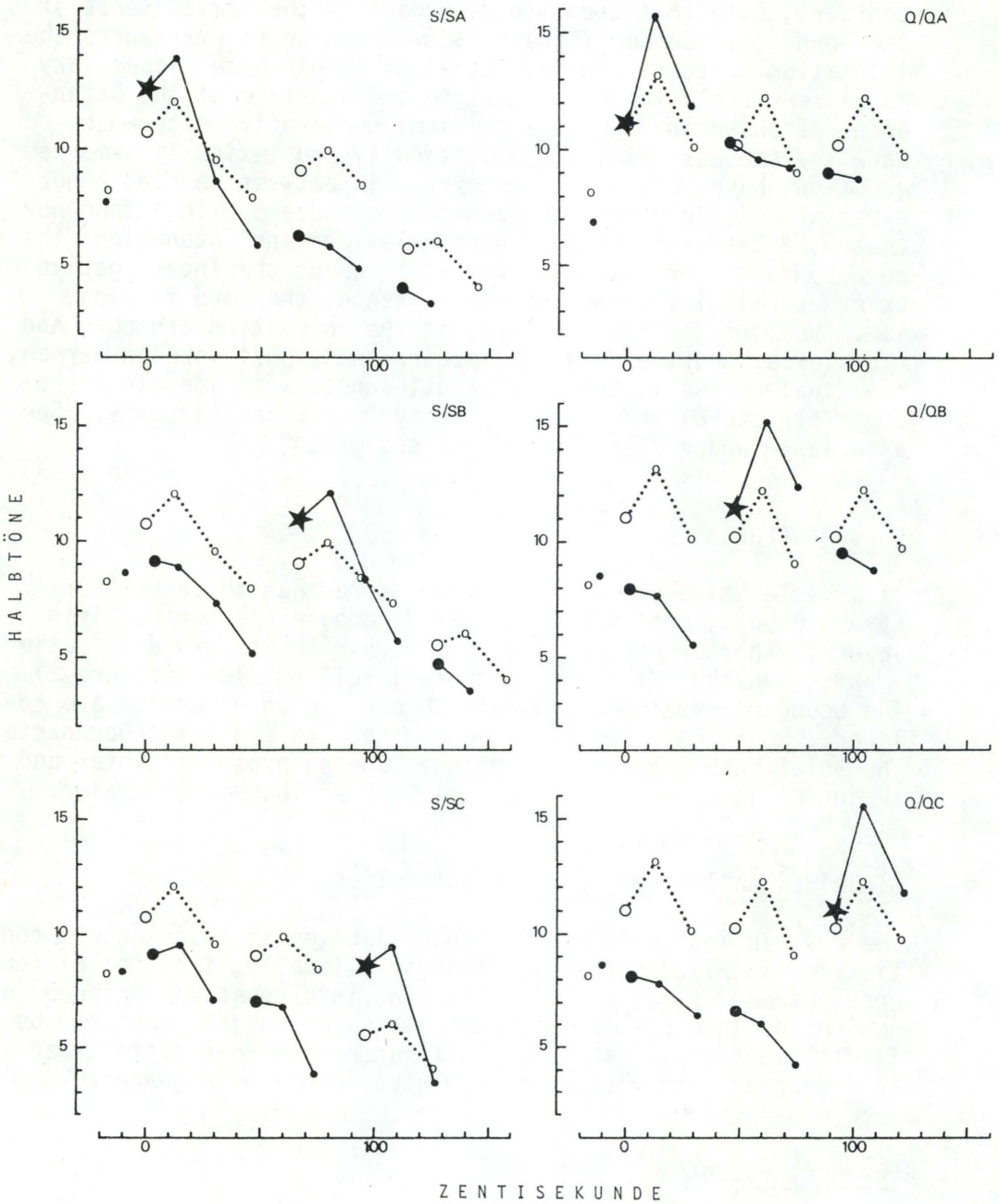


Figure 4

Stylized tracings of the course of Fo in statements (S - left) and questions (Q - right), which are prosodically neutral (open circles and dotted lines) or have emphasis for contrast (stars, points and full lines) in initial position (top), medial position (mid), and final position (bottom). Stars denote the emphatically stressed syllables, large points denote other stressed syllables, and small points depict unstressed syllables.

Contrast

Emphasis for contrast is manifested in Standard Danish by a raising of Fo of the stressed syllable of the emphasized word, together with a shrinking of the Fo patterns in the neighbouring stress groups (figure 4 and further Thorsen 1980b). Thus, emphasis for contrast in Danish shares certain features with the sentence accent of other languages, without being identical with it. One difference remains, however: in pragmatically neutral utterances of Standard Danish there is no particular prominence attached to the last (or any other) stressed syllable.

In the description of Danish, as opposed to German, the following points should be noted: In German there is a distinction between emphasis as an expression of the speaker's involvement - which is signalled globally, i.e. with larger tonal movements throughout the utterance - and contrast, which emphasizes a particular word and which is expressed in and around the stressed syllable. In Danish the parameter "speaker involvement" has not yet been subjected to investigation. The German parameter 'contrast' corresponds to the Danish 'emphasis for contrast'.

Furthermore, in the description of German intonation a distinction is made between accented and stressed syllables. Accented syllables are those lexically stressed syllables which are associated with a tonal change, i.e. a skip (up or down) in frequency. Any German word may contain a stressed syllable; this is the word accent. In the sentence, however, words will often be de-accentuated due to the context, i.e. they will appear without any tonal movement. Such words, with their meanings, are not made prominent. This distinction in German corresponds, in the description of Standard Danish, rather to the difference between stressed syllables, which are signalled through a rise in Fo, and syllables with secondary stress which have no such rise and which we find, e.g., in prosodic stress groups which surround a word with emphasis for contrast. This is a point where the description of Danish may have to undergo a revision; but until the concept of accent is better understood in Danish, the present account will be maintained.

Finally, we wish to point out that certain phenomena, which are termed emphasis for contrast in Danish might correspond to the German sentence accent, insofar as the difference between sentence accent and contrast in German is not a categorical one, neither semantically nor phonologically/phonetically.

GERMAN INTONATION

The following description of German intonation is based on phonetic investigations of the speech of ten North German speakers, where the Fo course is attained through acoustic analysis and measurements.[2] A more detailed description of the phonetic background can be found in the references.

Units of description

The basic units for the description of German intonation, apart from the whole text, are (1) the complete utterance, whose end is signalled by a very low Fo in answers and by a very high Fo in questions, tones which simultaneously reflect the terminal juncture, i.e. the boundary at the end of the utterance; (2) the phrase, the part of an utterance which is subordinated to it through tonal (and temporal) means (parentheses, subordinate clauses of various types); (3) the foot, i.e. the part of a phrase which begins with an accentuated syllable and encompasses all succeeding unstressed syllables. In certain cases units may merge. For instance, utterances like Es hat geklópft, Ích doch nicht!, Já, Wiesó? consist of phrases which are simultaneously feet.

The intonation, i.e. the Fo curves of utterances of varying length (1, 2, 3 and 8 accents), and the three utterance types: neutral answer, neutral information question, and surprised echo question by a North German speaker are shown schematically in figure 5. Every content word has the same weight. The time axis (horizontal) is normalized. The Fo rises in all accentuated syllables until the last one of the answer (declarative), where it falls. The two question types differ from the answer mainly at the end, through the rising versus falling tonal movement. The utterance with three accentuated syllables, in the three intonational versions is as follows:

ANSWER	Der Müller will die Männer immer Lümmel nennen.
	(The miller always calls the men louts.)
ECHO QUESTION	Der Müller will die Männer immer Lümmel nennen??!!
INFORMATION QUESTION	Will der Müller die Männer immer Lümmel nennen?

The different utterance lengths are achieved by reduction and expansion, respectively, of these sentences.

For the production of intonation in an utterance in German, the following prosodic information is necessary and sufficient:

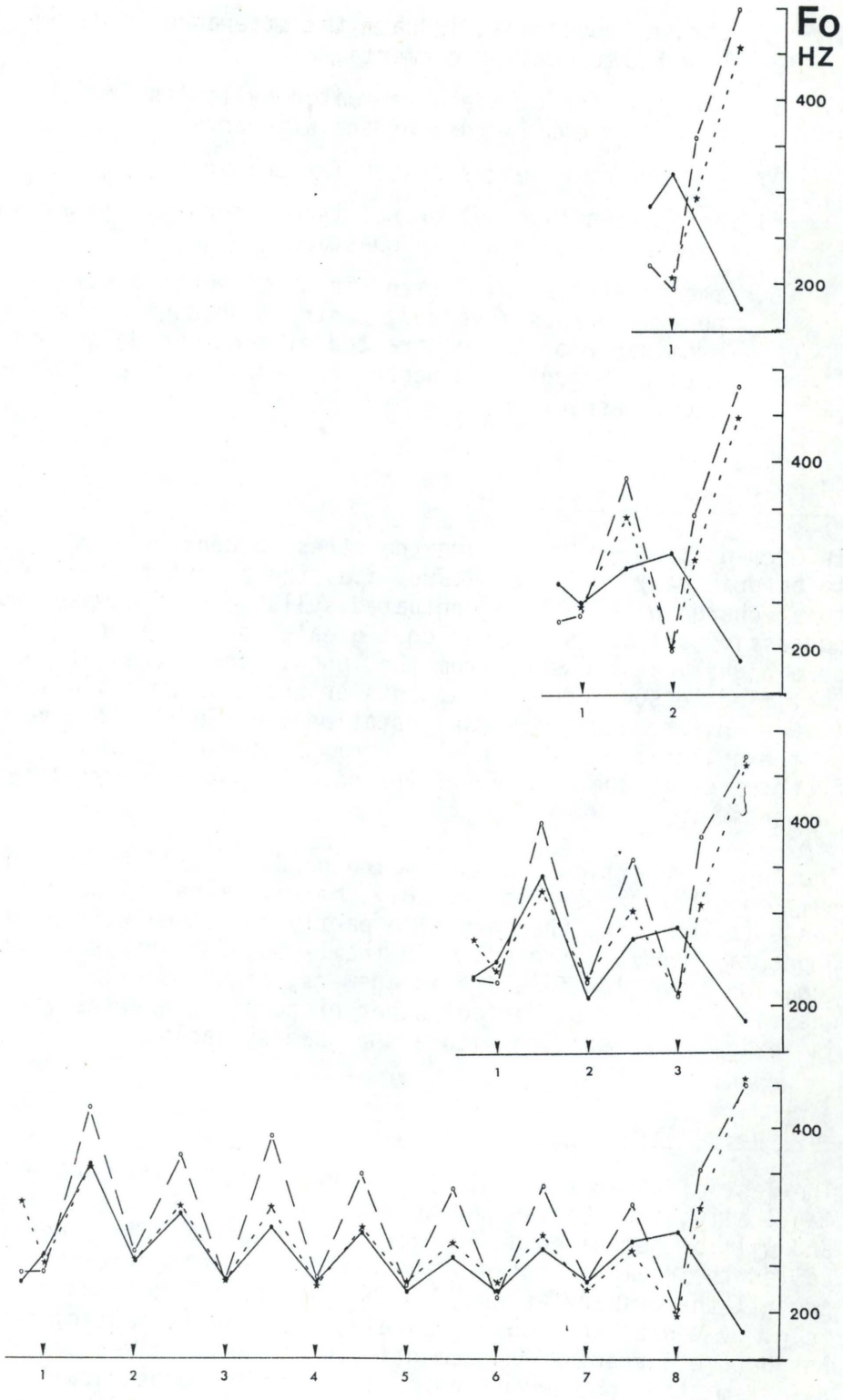


Figure 5

Normalized and superposed Fo tracings of answers, information questions, and echo questions of varying length. One to eight accents. The last accent is the line-up point.

1. Beginning and end of the utterance (initial and terminal junctures),
2. phrase boundaries, in case the utterance is decomposed into smaller elements,
3. the position of the accentuated syllables in the lexical items (words) of the utterance,
4. the semantic weight of the lexical units,
5. intonation type (utterance type): terminal in answers versus non-terminal in questions,
6. the speaker's involvement in what is being said: neutral versus involved, contrast through particular prominence of the contrasted element (word) and emphasis through prominence of a whole phrase or utterance, respectively.

Prosodic characteristics

In German, too, we may assume the stress accent or word accent to be primarily tonal in nature, i.e. the accent relies on a tonal change within the accentuated syllable or between two successive syllables and not on a greater articulatory force or a higher air pressure from the lungs. Therefore, it seems justified to speak only of accents or accentuated syllables in German, whose primary manifestation lies in the Fo movement of the syllable or in a skip up from the low Fo of the pre-accentuated to the high Fo of the post-accentuated syllable, respectively.

The tonal properties are partly independent, i.e. they do not rely on other properties, notably the rhythmical structure, of the utterance; but they are also partly connected with other properties and/or interact with them. So, for instance, intonation type (low final Fo in answers, high final Fo in questions) is independent of other properties, whereas accent is lexically associated with a stressed syllable.

Nucleus or main accent

The last of several accents, intended to be equally heavy semantically, in a German phrase will stand out from the others, and will be perceived as heavier or stronger. This accent is called the rhematic accent, the sentence accent, the main accent, the primary accent, or the nucleus ("Schwerpunkt") (e.g. by von Essen, 1956). Lately, the term (semantic) focus has become current. The semantic focus, which makes one lexical item stand out from others, is signalled prosodically through the main accent. This nucleus may occur with both intonation types: with terminal utterances it is manifested in the tonal fall (in the low end point), with non-terminal utterances it is signalled in the tonal rise (in the high end point). The stronger auditory impression of the main accent derives from two properties: firstly, from the tonal

change in the final accent itself, secondly, in the tonal course from the accent to the end of the utterance.

After the nucleus, no more accentuated syllables will appear (though there may be stressed syllables), i.e. after the tonal change associated with the primary accent the Fo curve runs smoothly towards the low or high end point. Thus, a phrase contains - besides possible accentuated syllables - only one nucleus or main accent. The nucleus is no independent unit: it can be predicted with semantic-pragmatic rules.

An example of nucleus displacement to the left, departing from an utterance with four accents which was pronounced as one phrase, is shown in figure 6, (a) as a terminal utterance type (an answer to a corresponding question or as a general statement) and as a non-terminal utterance type in two editions: (b) as information question with word order inversion and (c) as echo question, which expresses, inter alia, surprise and wonder. It is clear to see that the Fo curve, after the last, major change in the 4th, 3rd, 2nd, and 1st accentual position, respectively, runs smoothly low and high, respectively, to the end.

Special prominence

Intonation may also express the speaker's attitude towards the contents of his utterance, i.e. his involvement in what he is saying. A strongly participating, involved, emphatic attitude, e.g. surprise, wonder, indignation, is expressed tonally in a larger Fo range. The increase of the range, within which the pitch curve of the utterance varies, is primarily brought about by a raising of Fo maxima, but also by a lowering of the minima. This increase of the tonal range covers the whole phrase with all its accents. Figure 6 shows the echo question (c) in contrast to the information question (b) and to the answer (a).

Contrast, on the other hand, does have a semantic-lexical function, which is signalled tonally to the listener by the same means, namely an increased Fo range. However, its domain is limited to the lexical unit, the word. Contrast often expresses an opposition or the correction of an assumption. In phrases, which consist of only one foot, contrast and emphasis merge as far as the tonal manifestation goes. The difference lies, apart from the domain, in the function of the tonal expression. Contrast and nucleus are also different phenomena. This appears from figure 7. The curves derive from utterances which are of identical lexical, syntactic and phonological structure, but which occur in different prosodic contexts. The contrast appears tonally as prominent and marked peaks in each accent position.

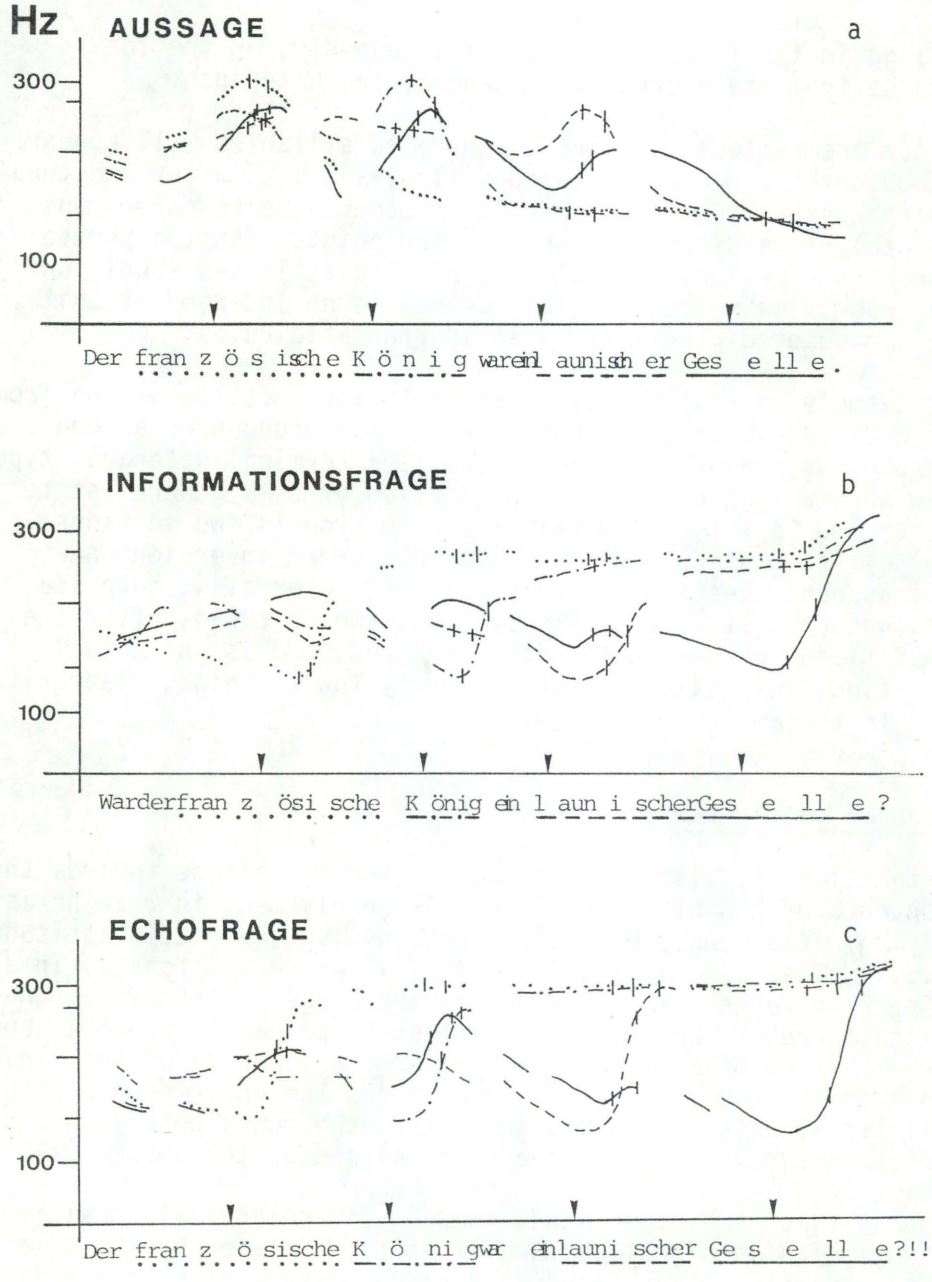


Figure 6

Sentence accent in answers, echo- and information questions with four possible accents (arrows). Superposed tracings from four versions with the sentence accent in each of the four positions. The arrows along the horizontal time axis denote the boundaries between the initial consonant and the accentuated vowel. The respective accentuated words are underlined in the same type of line as the Fo tracing. The sentence accent contains the last tonal change of the utterance.

Phrase intonation

Two junctures (boundary signals), the terminal (utterance final) and the non-terminal (phrase final), have their characteristic tonal expressions. While the former expresses itself through an extremely low or high Fo, respectively, the latter is manifested by an Fo in the middle to upper range of the utterance, the so-called phrase tone. The relation to the Fo course preceding and succeeding the juncture is also part of the total picture of these tonal-junctural phenomena.

Figure 8 shows the syntactically complex sentence "Mánchmal aber, und dann óhne grósse Ánkündigung, blíckte er sie fínster unter einer gerúnzelten Stírn an, dass sie béinahe das Fúrchten bekamen" (Often, however, and without particular warning, he looked severely at them, under dark and knitted brows, so that they were almost afraid.) as a part of a text. The utterance consists of four phrases. Each of the three non-final phrases ends in a high Fo which is reached either in the last syllable of the phrase ("Ankündigung, an") or already in the last accent ("aber"). The end of the last phrase is identical to the end of the utterance. The low Fo in the last syllable expresses simultaneously the terminal juncture and the intonation type 'answer' (statement). This figure also illustrates a natural Fo curve, as delivered by the acoustic analysis. The empty

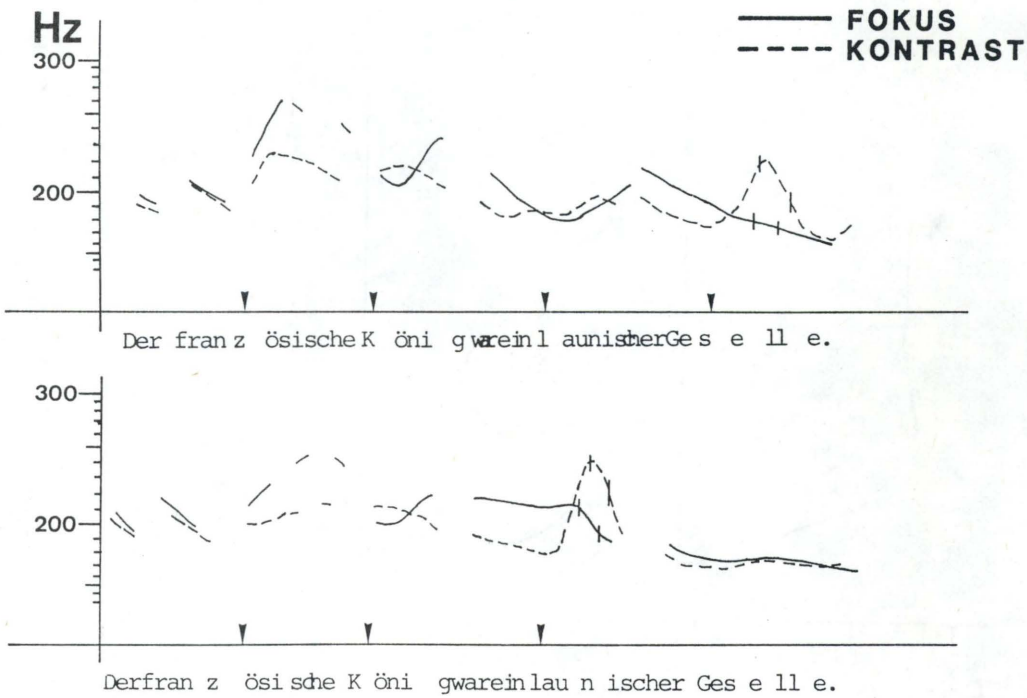


Figure 7

Contrast and sentence accent in a declarative. Fourth and third accentual position.

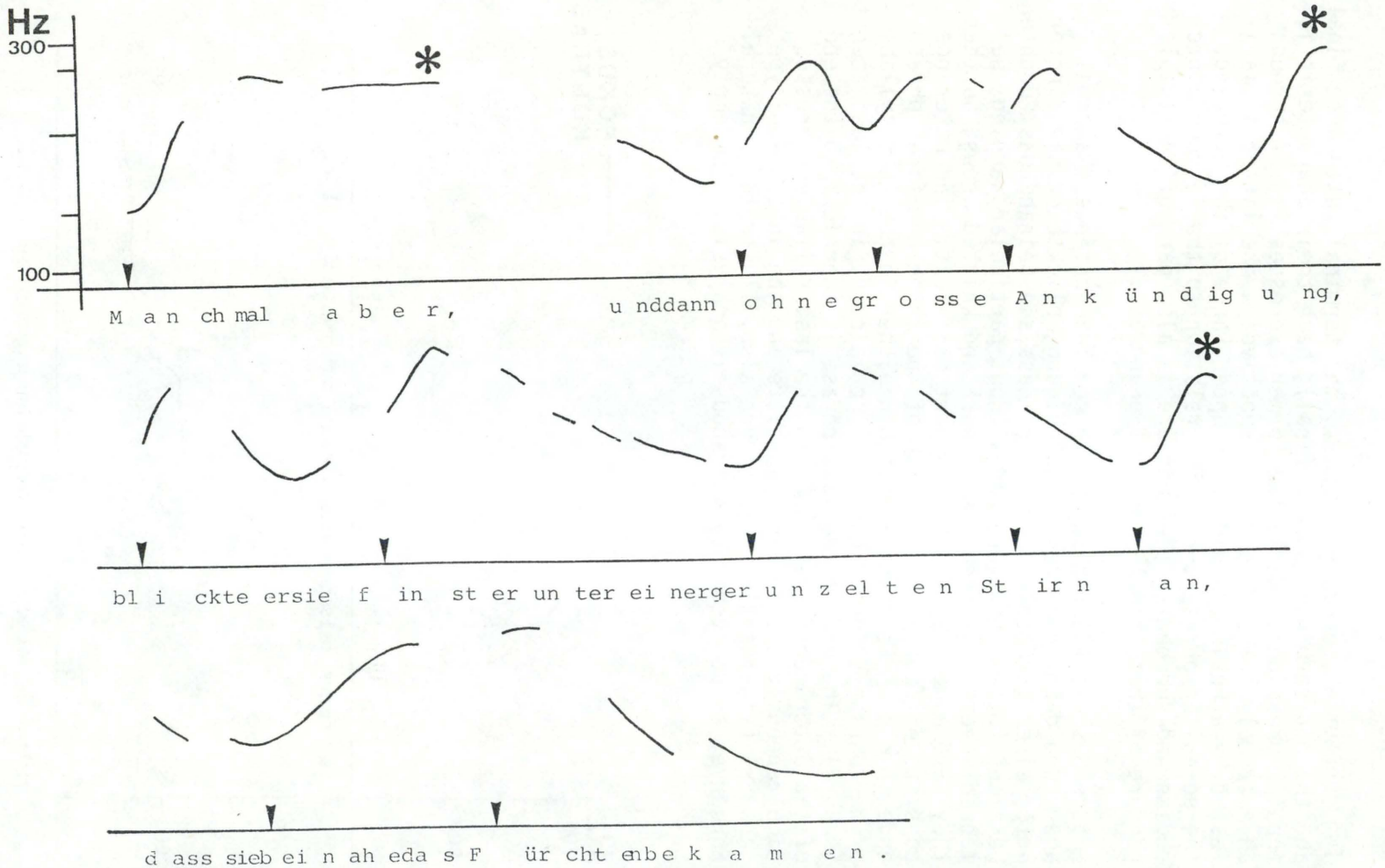


Figure 8

Phrasal intonation and phrase tone. The syntactically complex utterance (declarative) consists of four prosodic phrases. The three non-terminal phrases end in a high, continuation phrase tone, which is denoted with a star.

spaces in the curve correspond to unvoiced segments and phonatory pauses. Both junctures may be, but are not necessarily, accompanied by a physical pause. However that may be, we may assume that both junctures, beside their tonal manifestation, also involve a temporal lengthening of the last syllable, the so-called final lengthening, which has been demonstrated in other Germanic languages like Dutch, English and Swedish as well, but which is not a feature of Danish.

The intonational system

The course of Fo in phrasally identical utterances with equally prominent accents appears in German to be fairly simple and systematic. Apart from the last accent, the nucleus - which simultaneously expresses the intonation type - the tonal course of utterances with several accents appears, grossly simplified, as a repetition of the sequence Low-High(-Low) for each non-utterance final accent. This system for all three sentence types and for sentences with one to eight accents appears clearly from figure 5. The low points (Fo minima) of the accents in the three utterance types form a rather firm floor while the peaks in the utterances may be raised (so for instance in the echo question).

With unequally prominent accents, i.e. when an accent in the phrase, due to less semantic weight, appears weakened relative to a neighbouring one, the high and low points of this accent are displaced towards the middle of the range. The tonal movements of this accent diminish. A complete loss of semantic weight of a lexical element leads to de-accentuation and thereby also to the absence of Fo movements in the potentially accentuated syllable.

The tonal course through an utterance, i.e. the Fo minima and maxima with rising and falling movements between them, is characterized by an overall falling slope. As seen in figure 5, the Fo curve begins with a higher value of the minimum of the first accent than the one on which it ends. Intervening minima are distributed between these two values. This slope on the course of Fo is more or less identical for both utterance types and for the dimension 'involvement', although the echo question often begins with a lower Fo minimum in the first accent than the corresponding statement. Even when this difference is present, it should not be associated with utterance type differences, which are primarily signalled locally at the end of the utterance, but rather as an expression of involvement (in casu: surprise).

In the longest utterances, from five accents and upwards, Fo minima are not evenly distributed between the first and last minimum, as they are in shorter utterances. In the longest utterances a certain re-structuring appears to take place. The Fo minima of the penultimate and antepenultimate accent lie somewhat higher than the preceding ones. It is possible

that this upward displacement of the tonal movements - the peaks, at least, are raised - takes place in anticipation of the end of the utterance, which is so rich in tonal information. Thereby it is assured that the tonal change of the last accent may be really extensive and clear. A comparison with figure 2 shows that the floor, i.e. the line which connects the Fo minima, is clearly different from the corresponding sentences in Danish. The interval between the first and last Fo minimum in an answer is considerably smaller than in Danish. This behaviour is a natural consequence of the different tonal signalling of the intonation types answer and question, respectively, in the two languages, as will appear from the next section.

Intonation types

In German, answer and question differ very clearly and unambiguously with respect to their tonal manifestation. The difference between answer (terminal) and question (non-terminal) appears locally at the end of the utterance. The terminal intonation type ends with a low Fo. As a matter of fact, this point represents the speaker's absolute low, which he reproduces fairly accurately and constantly. The non-terminal utterance type ends high and may constitute the auditory pitch peak of the utterance. As appears clearly from figure 5, the Fo minima of the various sentence types practically coincide. From this we may conclude that the overall slope of the pitch curve, respectively the tonal floor (of the Fo minima), does not, in German, constitute the signal for intonation type. In Danish, on the contrary, where intonation type is not signalled locally at the end, the global tonal difference of the total utterance must serve to distinguish intonation types.

Tonal range

The course of the Fo curve, i.e. the timing of tonal minima and maxima, signals linguistic categories, but apart from that the speaker also exploits the possibility to change the range of his tonal movements for communicative purposes. With the speakers investigated so far the tonal movements in answers and information questions had the same range. Both utterance types were spoken with normal, neutral involvement. The information question is a true question, with which the speaker seeks real, to him unknown information. By strong involvement, as e.g. in the expression of surprise and wonder or incredulity in the echo question, most speakers increased their tonal range, though maintaining the basic structure of the tonal course. Furthermore, the syllables of the total utterance will be considerably lengthened in some cases. The tonal range is increased primarily by the raising of the Fo maxima. Thereby the Fo rises and falls in any given accent are increased. This observation leads to the assumption that the tonal course rests on a constant Fo bottom or floor, while the height of the ceiling

can be displaced upwards. These tonal relations express physiological-articulatory mechanisms and constraints.

Intonation and syntax

Prosody in general and intonation in particular have a double function in the grammar. They are associated with other grammatical components, on the one hand, but they lead separate lives, on the other. Accents occur only in lexically determined syllables; prosodic phrases often coincide with syntactic phrases, but not necessarily. The tonal course of phrases with the same prosodic structure is identical, whether the phrase be a nominal, a prepositional, or an adverbial one. Word boundaries have no influence on the Fo course, as is also the case in Danish.

SURVEY FOR COMPARISON

From the descriptions of German and Danish intonation above it appears that similarities and differences exist between the two languages. By way of conclusion they are summarized here:

Alike or similar are the following properties of German and Danish intonation: The basic units of description, namely text, utterance (sentence), phrase, foot, and syllable, which are hierarchically organized from the largest to the smallest unit. The (word-) accent is mainly realized through an Fo change in the accentuated syllable, or an Fo jump. Contrast and emphasis, i.e. a particular prominence of a word or an utterance, are expressed through greater Fo range, which is primarily achieved via a raising of the Fo peaks. Between prosody, in particular intonation, and syntax relations exist which, however, have not yet been investigated directly or to any great extent (but see Tropf 1984).

German and Danish intonation differ on the following points: In German, the last of several semantically equally prominent words in an utterance receives the nuclear or sentence or main accent; this is not so in Danish. In a neutral statement in Danish all accents sound equally prominent. The intonation types answer and question, respectively, demonstrate major differences. In German their tonal manifestation is located at the end of the utterance (from the last accent to the end) - the answer ends in a low Fo, the question in a high Fo. The preceding tonal course is identical. In Danish, on the contrary, answer and question differ globally, in the total course of Fo which is falling in terminal declaratives and horizontal in syntactically unmarked questions. At the end, both utterance types are alike tonally.

These similarities and differences are summarized in Table I:[3]

Table I

Similarities and differences in the intonation of German and Danish

Property	German	Danish	similar?
Units of description	Text (period) Utterance (sentence) Phrase Foot Syllable		yes
(word-)accent	tonal manifestation as fundamen-frequency change (jump)		yes
Nucleus (primary accent)	compulsory	non-compulsory	no
Contrast	increased tonal range in the accentuated syllable and simultaneous tonal reduction in the neighbouring syllables		yes
Intonation types	local signalling: lower final Fo in answers, higher final Fo in questions in otherwise similarly sloping pitch curves	global signalling: falling in terminal declaratives, less falling in other sentence functions	no
Intonation and syntax	no direct, but a rather complex relation		yes

ACKNOWLEDGEMENT

We are very grateful to Hartmut Haberland for valuable comments and suggestions for improvement.

NOTES

1. The stress group or foot consists of the accentuated syllable and all succeeding unstressed syllables (if any).

2. The present description deals exclusively with the basic linguistic units of intonation, i.e. with the basic intonational structure of utterances with a purely linguistic function. Paralinguistic aspects, like emotions or irony as well as voice quality, are excluded. Note also, that as far as we know, the acoustic data on German intonation with their integration in an intonation or prosody model, are the first of their kind. Consequently, the statements made here about German intonation cannot be exhaustive. On the other hand, we assume that the description contains fundamentally valid observations and characteristic features of Standard German intonation. Further data and knowledge in this area are anticipated: due to a recent financial support for intonation research from the German Research Council, there is a lively activity in intonation research going on in the Federal Republic.

3. This comparison is, for various reasons, not complete.

REFERENCES

Bannert, R. 1982: "Temporal and tonal control in German", *Working Papers 22*, Lund University, Institute of Linguistics and Phonetics, 1-26

Bannert, R. 1983a: "Modellskizze für die deutsche Intonation", *Zeitschrift für Literaturwissenschaft und Linguistik 49*, 9-34

Bannert, R. 1983b: "Some phonetic characteristics of a model for German prosody", *Working Papers 25*, Lund University, Institute of Linguistics and Phonetics, 1-34

Bannert, R. 1984: "Towards a model for German prosody", *Working Papers 27*, Lund University, Institute of Linguistics and Phonetics, 1-36

Bannert, R. 1985: "Fokus, Kontrast und Phrasenintonation im Deutschen", *Zeitschrift für Dialektologie und Linguistik 52*, 289-305

Basbøll, H. 1977: *Dansk Fonetik og Fonologi - Skitse til en Systematisk Indføring*, Nordisk Institut, Odense University

Bierwisch, M. 1966: "Regeln für die Intonation deutscher Sätze", *Studia Grammatica VII*, 99-201, Berlin

Delattre, P., Poenack, E. and Olsen, C. 1965: "Some characteristics of German intonation for the expression of continuation and finality", *Phonetica 13*, 134-161

Ehlich, K. 1981: "Intonation des gesprochenen Deutsch: Aufzeichnung, Analyse, Lehre", *Kopenhagener Beiträge zur Germanistischen Linguistik 18*, 46-93

Essen, O. von 1956: *Grundzüge der hochdeutschen Satzintonation*, Ratingen

Isačenko, A.V. and Schädlich, H.-J. 1966: "Untersuchungen über die deutsche Satzintonation", *Studia Grammatica VII*, 7-67

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

Thorsen, N. 1980b: "Neutral stress, emphatic stress, and sentence intonation in Advanced Standard Copenhagen Danish", *Ann. Rep. Inst. Phon., Univ. Copenhagen 14*, 121-205

Thorsen, N. 1983a: "Standard Danish sentence intonation - phonetic data and their representation", *Folia Linguistica 17*, 187-220

Thorsen, N. 1983b: "Two issues in the prosody of Standard Danish", in: *Prosody: Models and Measurements* (eds.: A. Cutler and D.R. Ladd), Springer, Heidelberg, 27-38

Thorsen, N. 1984: "Variability and invariance in Danish stress group patterns", *Phonetica 41*, 88-102

Thorsen, N. 1985a: "Intonation and text in Standard Danish", *J. Acoust. Soc. Am. 77*, 1205-1216

Thorsen, N. 1985b (forthcoming): "Intonation and text in Standard Danish - with special reference to the abstract representation of intonation", *Phonologica 1984*

Tropf, H.S. 1984: "Zur Intonation fokusierter Nominalphrasen in Antwortsätzen: Eine Fallstudie", Mimeo. Nimwegen.

ARIPUC 20, 1986, 67-78

DEVOICING OR STRENGTHENING OF LONG OBSTRUENTS IN GREENLANDIC*

JØRGEN RISCHEL

It is a characteristic feature of Modern West Greenlandic that long nonnasal consonants are voiceless, whereas these have short voiced counterparts in the case of continuants. Seen in isolation this distribution is suggestive of spontaneous devoicing of long obstruents, but on the basis of evidence from dialects and from old spellings it is suggested that there may be an old "strengthening" process (segmentalization) underlying this modern feature of voicelessness.

I. INTRODUCTION: THE OBSTRUENT PATTERN

It is a well-known characteristic of West Greenlandic Eskimo that this type of Eskimo (unlike North Alaskan Inupiaq, for example) has complementarity of voiced and voiceless fricatives, short fricatives being voiced and long ones voiceless (in the new Greenlandic orthography, which is used in this paper, g and r symbolize voiced velar and uvular continuants, and gg and rr their long voiceless counterparts, whereas the corresponding labial set is distinguished orthographically as v versus ff). In accordance with the principle of feature distribution just stated we find that voiced fricatives are automatically replaced by voiceless segments in environments where they are long by

---

*) *This paper has appeared in B. Brendemoen, E. Hovdhaugen and O.H. Magga, eds.: Riepmočála, Essays in Honour of Knut Bergsland (Oslo: Novus Forlag 1984), p. 122-135.*

true gemination (or by assimilation), e.g. iga- 'to cook' - iggavik or igaffik 'kitchen', neripput 'they ate' - nerripput 'they shared a big meal'.

Morphophonemically, this apparently straightforward pattern is complicated in that long and short fricatives regularly participate in deviating alternation sets. Although the short (voiced) fricatives v, g, r and the long (voiceless) fricatives ff, gg, rr all occur quite frequently in West Greenlandic, the source of a long geminate fricative does not necessarily appear synchronously as a short fricative, nor does a consonant which appears synchronously as a short fricative necessarily remain fricative under gemination. We rather find a preponderance of alternations between a "weak grade" short alternant and a "strong grade" long alternant. Diachronically, this has to do with weakening process, as appears a.o. from the pioneering comparative studies of Hammerich and Bergsland. As for old short fricatives, the normal situation for g and r was to be lost intervocalically, the result being that geminate gg and rr frequently alternate with zero (cf. naggat 'end' derived from naa- 'to end'; unarrat (old) pl. of unaaq 'harpoon shaft') rather than with g and r. On the other hand, a large number of new occurrences of voiced fricatives arose by lenition of old short stop consonants in intervocalic position, whereas stops were preserved when geminated or occurring in clusters. In this way we get a number of alternations between v and (p)p, cf. kujavarpoq 'has moved further to the south' vs. avannarparpoq 'has moved further to the north' (-var- vs. -par-), and particularly between g and (k)k, r and qq (cf. aalisagaq 'fish' - pl. aalisakkat; ujarak 'stone' - pl. ujaqqat). In a number of cases, finally, we encounter old short stop consonants which have been preserved, i.e. protected from lenition, in the position after the first (single) vowel of roots (cf. Bergsland 1955, p. 12), this position after the first mora being in general a "strong" position. Such short stops, then, engage in alternation with long stops if they stand (or used to stand) in the just mentioned strong position (cf. napavoq 'is in upright position' - napparpaa 'puts it in upright position'; nukik 'strength' - nukkiorpoq 'forces himself, uses his strength'; niaqoq 'head' (as suggested by Bergsland probably from *nayquq) - pl. niaqqut).

In the dental or (with a more adequate labeling) coronal series of obstruents there is the complication that West Greenlandic exhibits two sibilants, which are voiceless even when short. One is a plain s, the other has a more retracted point of articulation. In the (new) orthography they are not distinguished, and they are actually kept distinct only in a geographically limited area (within which they are expected to merge eventually). Both s-phonemes occur phonemically long and short, but in cases of morphological alternation involving gemination the regular and frequently occurring patterns are: (a) short plain s alternating (like j) with the long affricate ts (cf. nasaq 'cap' - pl. natsat), (b) zero alternating with retracted ss (cf. iluliaq 'iceberg' - pl. ilulissat).

The development of the sibilant subsystem is a very complex issue, which is much debated in current work. In the present context it may suffice to mention that in forms such as nasaq the plain s derives from a former affricate, in some other forms from a stop consonant *t, and that the retracted s derives from a voiced (post)alveolar or palatal fricative. All of this is supported by the retention of the older state of affairs in other Eskimo languages or dialects. What is suggested by authoritative linguistic reconstructions (based above all on Bergsland's work, cf. Bergsland 1959, 1966) is that the proliferation of voiceless sibilants in West Greenlandic is a secondary phenomenon. It is noteworthy that we have within this subsystem a genuine instance of seemingly unconditioned, spontaneous devoicing of a fricative taking place also if the fricative is phonetically (and phonemically) short, viz. in numerous forms such as isi 'eye', aasaq 'summer', puisi 'seal' (all with the retracted s).

Finally, West Greenlandic has also a lateral l, which switches manner of articulation and crucially enters the obstruent pattern under gemination (see later), although the nasal sonorants geminate without any change in manner of articulation, cf. ameq 'skin' - pl. ammit with a voiced nasal in both cases.

II. GEMINATION AND DEVOICING

Consonant gemination as a kind of syllable strengthening process is a very old feature of Eskimo. It is shared by the other dialects of the Eastern (Inupiaq) branch of Eskimo, and as pointed out by Bergsland its previous existence also in the Western (Yupik) branch may be inferred even though the quantity pattern which still survives in Eastern Eskimo was eventually replaced by a very different quantity pattern in Yupik (with drastic changes such as shortening of old geminates and secondary processes of new gemination or lengthening affecting both consonants and vowels). On this point, then, the dialects of the Eastern branch are more archaic. On the other hand, the old pattern must obviously have involved straightforward alternations of single and geminate consonants with the same manner of articulation, unlike the situation in Modern West Greenlandic, where - as already stated - we have a complex situation with four types of manner alternation accompanying gemination, viz.

(i) zero alternating with long voiceless fricative
(ii) short voiced fricative alternating with long voiceless fricative
(iii) short voiced fricative alternating with long voiceless stop
(iv) short nasal or stop alternating with long nasal or stop (without change in voicing).

Whereas the fricatives in set (iii) must be assumed to have acquired voicing as a corollary of lenition of old stops, com-

parative evidence is unanimously in favour of voicelessness being a secondary feature in sets (i) and (ii). Voicelessness in these sets is, in a dialect geographic perspective, a feature which is highly characteristic of Greenlandic.

In North Alaskan Inupiaq, for instance, there is alternation between short and long voiced continuants in cases corresponding to those above. Since the voiceless member of each pair is obviously secondary compared to the voiced one, there is overwhelming evidence both from internal reconstruction and from dialect comparison of a Greenlandic sound change replacing voiced by voiceless articulation under certain conditions involving length. It is very natural that Greenlandic Eskimo is often mentioned as evidence in support of the claim that there is a universal tendency for long fricatives to become voiceless by a spontaneous sound change.

There are some inherent difficulties with the assumption of spontaneous devoicing, however. It is true that one may well expect voicing to be hampered if there is a long interval of strong constriction, as in true fricatives (because most of the pressure drop then takes place across the oral constriction instead of taking place across the glottis, as is necessary for sustained vocal fold vibrations), but the voiced continuants in contemporary Greenlandic are not all that fricative. This, of course, does not prevent us from assuming that these sounds used to be strongly fricative obstruents, at least in their geminate versions (this is consistent with their articulation in North Alaskan Inupiaq). Still, one may ask why across-the-board devoicing - if it is such a natural process - should happen only in Greenlandic, considering that there is no direct evidence that the fricative articulation in old Greenlandic was particularly favourable to devoicing.

It is a much more serious challenge to the devoicing hypothesis that the voiced-voiceless alternation occurs also with the lateral l, as in ukaleq 'hare' - pl. ukallit (where l is a short voiced lateral or slightly flapped sound, and ll a long voiceless lateral fricative). All evidence suggests that this is "from the start" a true lateral consonant, and it differs morphophonemically from the ordinary fricatives in that there is just one alternation set, viz. (voiced) l - (voiceless) ll, i.e. no alternation involving either zero or a geminate stop in (Central) West Greenlandic (in cases such as atserpaa 'signs it' from ateq 'name' and -(l)erpaa 'provides it with sth.' the l enters the process by which ts arises, but compare analogous examples such as imaq 'content' - immerpaa 'fills it out').

If we wish to understand the diachronic developments underlying such voiced - voiceless alternations as v - ff, g - gg, r - rr, and indeed l - ll, in Greenlandic, there are three different sources of insight. The first is comparative evidence from other types of Eskimo outside Greenland, and above all around the Bering Strait, where more archaic phonologies are found. We have seen already how this is suggestive of

spontaneous devoicing of long obstruents and even of the long lateral in Greenlandic. Let us consider now what the other types of evidence have to tell.

The dialects of Greenlandic Eskimo proper have until quite recently received little (if any) attention in comparative Eskimo studies. By and large the central dialect of West Greenlandic has been taken to be representative of Greenlandic Eskimo as a whole. Textbooks on Greenlandic have contributed to this notion of Greenlandic as being more or less one dialect by hardly mentioning the issue, but part of the reason is that there has been a lack of adequate and manageable descriptions of other dialects and a lack of concise dialect geographic surveys. The pioneering work with regard to phonological dialect surveys was done by Petersen (1969/70, 1975) less than two decades ago (also cf. the expedient survey in Fortescue 1983).

If we look at Petersen's chart of phoneme correspondences it is noteworthy that the voiceless fricatives ff, gg, rr of West Greenlandic are matched by stops in the northermost (Upernavik) and southernmost (Cape Farewell) fringe areas of West Greenland and in the East Greenlandic dialect. Voiceless ll is matched by a stop in southernmost West Greenlandic and in East Greenlandic. - To be strictly correct, exception must be made for some instances in which the easternmost Inuit dialects had a voiceless lateral already before the specifically West Greenlandic devoicing; in these instances East Greenlandic has developed (t)s directly from the old voiceless lateral, example: East Greenlandic atsinaaq, West Greenlandic allunaaq 'rope, string' from something like *akLunaaq with "L" symbolizing the voiceless lateral.

The stop consonant reflexes in these Greenlandic dialects far off Central West Greenlandic are straightforward for gg and rr, which are matched by kk and qq, respectively. As for ff, Upernavik and East Greenlandic have pp, but Cape Farewell has dorsal kk (rf is matched by qq). As for ll, East Greenlandic has tt (now often written "dd"), but Cape Farewell has a supradental affricated stop which is phonemically distinct from tt. These discrepancies (which can be explained) are not at issue here; the general observation to be made is that there are two main categories of reflexes of old long non-nasal sonorants in Greenlandic: (a) as long voiceless continuants and (b) as long stops with or without affrication. As a sweeping statement we may say that West Greenlandic exhibits the former, East Greenlandic the latter, although the fringes of West Greenlandic largely side with East Greenlandic.

How does this tie in with the explanation of voiceless ff, ll, etc. in West Greenlandic as being due to spontaneous devoicing? If one insists that we have such devoicing in the West Greenlandic examples it is appropriate to search for an unrelated explanation for the development in East Greenlandic, for it does not seem particularly convincing that the newly developed long voiceless continuants spontaneously went on and became stops (although such a further shift cannot be excluded).

There is certainly no evidence that the development from old voiced geminate continuants to stops ever went through anything like the stage encountered in modern Central West Greenlandic. On the other hand, it would seem utterly strange for Greenland to have two basically different processes affecting the long voiced continuants considering how unaffected these consonants are in most of the remaining Eskimo area. Neither of the two alternative approaches being overwhelmingly attractive, the situation invites a diachronic hypothesis according to which the Greenlandic dialects as a whole share a strengthening of some kind of the old voiced geminates into a series of reflexes which provide a natural starting point for both of the ultimate developments: into voiceless fricatives and into stops.

Here the third source of evidence comes in, viz. spellings in old Greenlandic sources.

III. OLD GREENLANDIC SPELLINGS

The first truly systematic and phonetically interpretable transcription of Greenlandic Eskimo (and of Eskimo in general) was the one developed by Samuel Kleinschmidt around the middle of the 19th century and used even today, though the new orthography of 1973 is gradually taking over. There is a considerable bulk of sources before Kleinschmidt, but the (17th and) 18th century sources have not been much utilized in diachronic phonology because of glaring inadequacies with regard to such features as length, the distinction between velar and uvular stops, and the rendering of vowel qualities. Nevertheless, there are very important insights to be gained from this early material, and this is very much true with respect to the question of what happened to the old voiced geminates in Greenlandic.

With regard to voicedness or voicelessness of fricatives the earliest useful sources from the second half of the seventeenth and the first half of the eighteenth century (i.e., sources which in part antedate, in part postdate Hans Egede's arrival in Greenland in 1721) are clearly suggestive of voicedness of both short and long fricatives. For example, in a 1654 word list preserved in manuscript (of P.H. Resen's Danske Atlas) from the 1680ies we find short intervocalic fricatives rendered as v, g, r in accordance with modern orthographical usage: Ivirning (now: iviannggit with plural -t) 'breasts', Acago (i.e. aqagu) 'tomorrow', Kameresin (i.e. qimerissat) 'eyelashes', though deviant spellings also occur. As for the long intervocalic fricatives, we find such spellings as b, g(j), gg corresponding to ff, gg of the modern orthography, e.g. Sibian (siffiaq, or rather siffiat 'your-') 'hip'. Sigju (i.e. sigguk) 'trunk', Naggesung (i.e. nerrersooq) 'inclined to eat a lot' (in the old source translated as Essen, i.e. 'to eat').

Although the earliest spellings are understandably inconsistent they definitely suggest that the long (geminate) consonants in such cases were not voiceless fricatives (the old evidence is very conflicting with regard to rr, however, which was sometimes spelled ch etc., probably to indicate the rasping quality of the uvular fricative which was more audible when it was a long segment). It is, on the other hand, impossible to see from spellings such as those illustrated above whether the long segments were fricatives or lax (voiced) stops. Offhand the (for a long time dominant) spelling b or bb for modern ff seems strongly suggestive of realization as a bilabial stop which, however, must have been different from the "true" stop consonant /p/ (the latter is mostly rendered as p or pp in old sources). If that is true, the same might be expected for the long velar segment and possibly for the long uvular segment. This would be in agreement with the later reflexes in East Greenlandic (though not with the reflex /kk/ for */vv/ in the Cape Farewell dialect), but it would be hard to reconcile such an interpretation with the fact that these long segments started out as fricatives and have fricative reflexes in modern West Greenlandic. A development voiced fricative > stop > voiceless fricative seems too far-fetched to deserve consideration. It seems much more likely that the long labial fricative sounded too different from the labial fricatives occurring in Danish or German to warrant a symbolization such as v or w; in all probability this long fricative was articulated as a narrow unrounded (slit-articulated) bilabial fricative which to a Danish or German ear sounded more like a voiced b than a v or w.

The orthography of the manuscripts and published works following after the colonization in 1721 is of course characterized by considerable improvements, but there are no dramatic changes in the rendering of the fricatives. It is noteworthy that as late as the beginning of the nineteenth century Otho Fabricius in his dictionary (Fabricius 1804) gives spellings such as "Sabbiorbik v. Sagviortarbik" for what is now spelled saffior-(tar)fik 'forge' (Fabricius also: 'anvil'). His vacillation here and elsewhere between b(b) and gv reflects a dialect difference in the pronunciation of the Greenlandic labial fricative as pointed out by Petersen, the spelling gv representing a South Greenlandic labialized velar fricative (cf. the development of */vv/ into /kk/ in the Cape Farewell dialect mentioned above). What is interesting in the present context is that neither of the alternative spellings is suggestive of voicelessness (since the letter f would be the obvious way to indicate voicelessness according to its use in Danish). It is hard to know to what extent this seemingly very conservative transcription reflects stability in the phonology and to what extent it just reflects loyalty towards an orthographical tradition for which the early word lists of 1654 may have played a considerable role.

But in any case, the spelling gv was not customary before Fabricius entered the scene with his special knowledge of the

southern Frederikshaab (Paamiut) dialect, so the odds seem in favour of the assumption that voicelessness of such long segments or clusters postdates the beginning of the nineteenth century.

For the long velar fricative the spelling is regularly gg (which also occurs along with g to represent the short voiced velar fricative); it is only with the uvular point of articulation that the spelling is suggestive of a specific (strident or voiceless?) manner of production of the fricative when long (cf. the remark on earlier spellings above). For the long uvular Fabricius has such spellings as Noṙak 'calf (specifically of caribou)' (now spelled norraq) with a diacritic mark on r to distinguish it from the short uvular of such forms as torā̊juvok (now: toraajuvoq) 'is certain of aim when shooting'. Similar spellings occur until Kleinschmidt's orthography settles on rr vs. r by the middle of the nineteenth century.

If we turn now to the lateral continuant the spellings of the 18th and 19th centuries are much more suggestive of a phonological development. The preferred spelling in early sources is l or ll (consonant doubling did not serve to indicate length in pre-Kleinschmidt orthography), no matter whether the lateral is long or short, but only in instances where the consonant or cluster was voiced. Clusters appearing at that time with a voiceless lateral were spelled differently, as seen from numerous, more or less consistently used spellings with tl or kl in Paul Egede's dictionary of 1750.

> *As for clusters with a preceding velar, those with a voiceless lateral were spelled with kl (sometimes ktl), e.g. Aklunák 'rope, string' (now allunaaq; with early voicelessness as evidenced both by dialects west of Greenland and by the East Greenlandic reflex atsinaaq mentioned earlier) as against Segluvok (now salluvoq) 'he lied'. With a preceding uvular the apparent voiceless-voiced distinction is mostly expressed by rtl or rkl vs. rl, as in Tórtlorpok (now torlorpoq) 'called out', torkluluanga (i.e. torlulavaanga in modern spelling) 'he was calling for me' versus Korlórpok (i.e. qorlorpoq) 'poured down (e.g. through a funnel)'. In cases where there is etymologically no labial, velar, or uvular preceding the lateral, the long lateral (or dental stop plus lateral?) is expressed in either of two ways, viz. tl contrasting with l (or ll), as in Itlerbik (now illerfik) 'chest', versus Sillit 'whetstone' (also occurring with the spelling Silii in 17th century lists). The former spelling is particularly well attested with affixes such as -tlar... 'forcefully' or contemporative -tlu- (only sporadically -(l)lu-; -tlu being the regular post-consonantal alternant) both of which occur in numerous illustrative phrases in Paul Egede's dictionary (1750), as against other words with -l- or -ll-.*

> *Examples are, on the one hand: Usitlarau (now usillaaraaq) 'is generally heavily loaded', torartlugo (now toraarlugu) 'heading towards it', on the other hand: -ngilet (with indisputable gemination due to plural inflection) as attested in several forms such as ajungilet (i.e. ajunngillat) 'they are good (literally: not bad)', and perfective stems with gemination such as Aúlarpok or aularpok (now aallarpoq) 'he left' (derived from the stem *aula- 'to move'). There can be no doubt whatsoever that tl and l(l) reflect two formerly contrasting long items.*

Although the spellings involving the expected voiceless lateral are not phonetically transparent they can be construed to indicate that there was either a fully developed stop segment in the beginning or a tendency toward such "segmentalization". This would off-hand explain the East Greenlandic reflex /ts/ as a modification of a cluster */tL/. As for the forms with an expected voiced lateral the evidence is overwhelmingly in favour of a plain, long voiced segment in 18th century Greenlandic.

Eventually, however, spellings with -dl- take over though exhibiting some confusion with -gl- (eventually written -gdl-). Fabricius (1804) writes sidlit for 'whetstone', Aûdlarpok for 'he left', etc., a type of spelling which is virtually (totally?) nonexistent in Paul Egede's dictionary of half a century earlier. There is reason to believe that this signals a development of a stop segment in the beginning of the long lateral, i.e. "segmentalization" in a straightforward sense.

It is hard to determine when exactly the voiceless clusters merged with the voiced ones since Greenlandic orthography through the ages is characterized by considerable loyalty towards earlier spellings of individual words. Variant spellings in Fabricius (1804) such as tôrdlorpok = tôrklorpok for Egede's Tôrtlorpok are suggestive of ongoing merger around 1800. Taken together with forms such as iklerbîk 'chest' for Egede's Itlerbik the above example suggests that "kl" tended to be generalized (due to spurious etymologizations, or because of sporadic development of a velar before the long lateral?) as the way to render the voiceless lateral. It is less likely that the variant spellings with "rkl", "rdl" indicate a development from voiceless long lateral to voiced lateral. Rather, they may be indicative of the opposite, viz. that clusters stemming from a long voiced lateral, or from a stop plus a voiced lateral, were being devoiced thus obviating the need for distinct spellings of those clusters that already had a voiceless lateral.

By this merger "dl" came to symbolize a voiceless lateral, mostly with a preceding stop consonant symbol. This usage was generalized by Kleinschmidt, and in some cases he combined the

tendency to use a velar symbol in front of the lateral with the use of "dl" to symbolize voicelessness, spellings such as igdlerfik resulting from this fusion of earlier usages. Kleinschmidt, however, distinguished between vdl, tdl, and gdl to indicate what is now a plain long voiceless lateral, although it is not clear to what extent this was based on phonetic reality. (The "g" in igdlerfik cannot be a priori excluded as a spontaneous development, a differentiation of some kind, but the "v" in Kleinschmidt's avdla 'other' must be totally erroneous. A check of labialization of consonant clusters in an archaic type of Upernavik dialect which otherwise has preserved quite many instances of labialization, gave a negative result for this word as expected on comparative grounds, and Kleinschmidt's spelling is glaringly at variance with Fabricius' Adla half a century earlier.) One must fully subscribe to Bergsland's characterization of Kleinschmidt's spellings of such complexes as being "often without etymological foundation" (1955, p. 2).

IV. CONCLUSION

The orthographical evidence strongly suggests that the voicelessness of the long lateral developed not as spontaneous devoicing but as "segmentalization" (differentiation of the continuant into stop plus continuant). The next step was devoicing of all such clusters with an initial stop, and, finally, as a quite late process which has not yet been completed in all dialects, the clusters underwent regressive assimilation (irrespective of their origin as a cluster or as a geminate). The last step in this development is well attested on independent grounds since simply all consonant clusters have undergone it (with the exception that ts has remained an affricate, and that the cluster initial uvular of clusters such as rl, rf, etc. has left a trace affecting the preceding vowel). Thus, the only weak link in this explanation is the contention that clusters of stop plus voiced lateral became fully voiceless. Is such devoicing inherently more likely than spontaneous devoicing of a long intervocalic lateral? I think it is. The conditions for devoicing of the lateral would presumably be favourable after a (possibly voiceless) stop, particularly since one may expect a more fricative articulation immediately after a (homorganic) stop than elsewhere.

Taking a well-known word such as illu 'house' (the source of the term igloo), we may now follow the development of an old cluster. The contention is that the spelling forms occurring over time should be taken essentially on face value: we start with a cluster "gl" in the oldest sources, as etymologically expected (P. Egede: Iglo), then in the 19th century comes the spelling with "gdl" indicating possibly both segmentalization and devoicing (this is the spelling found in Kleinschmidt, but actually not invented by him, cf. that Steenberg 1849 has igdlo though his spelling is otherwise completely pre-Kleinschmidtian), and today we have the new spelling "ll", reflecting complete assimilation and merger of voiced and voiceless

clusters (so that the voiceless lateral can be taken as an allophone of /l/ in modern West Greenlandic). - Although the old spellings have been almost ridiculed, they may not have been far off the mark, and if the new orthography is scorned by some Greenlanders because of spellings such as "ll", which allegedly fail to indicate the special character of the long lateral, this controversy reflects the strange fact that the digraph "dl" had come to mean "voiceless lateral" although it probably started out as a straightforward and indeed adequate marking of segmentalization (cf. the pronunciation of words such as kalla ("kadla") in various West Scandinavian languages and dialects).

With regard to such segmentalization and subsequent devoicing and assimilation interesting support may be found in the observations of Holtved (1952) on the Polar Eskimo dialect, where he reports about variant forms with regard to stop plus lateral (as in illu 'house') - exhibiting more or less exactly such a chain of development under way (and apparently compressed into a short time span, as would also have been the case with West Greenlandic some 100-200 years ago).

The old spelling evidence for West Greenlandic is sadly non-informative with respect to the various fricatives dealt with earlier, but it is tempting to generalize from the above explanation of how the lateral devoiced and suggest that such segmentalization into voiced and eventually devoiced affricate-like clusters occurred also in other instances (spellings such as bb, gg do not at all rule out the possibility that there was a change from long narrow fricative to stop plus homorganic fricative without change of orthography). The only cluster for which the spelling evidence is clearly in disfavour of this interpretation, is /rr/. Spellings such as kr, which might be expected in case there were segmentalization, are conspicuously absent, so maybe we must reckon with spontaneous devoicing here (which is not very surprising, considering the natural tendency for uvular frication to involve much stridency).

It must probably remain an open question whether the long labial and velar fricatives were segmentalized into stop plus fricative or just developed very narrow allophones favouring spontaneous devoicing. The virtue of the former assumption is that it permits a generalized account for old clusters as well as old long continuants (involving fricatives and the lateral, though apparently with the uvular taking a separate course, as it does rather generally in Greenlandic phonology), and that it provides a basis for deriving the West and East Greenlandic forms rather directly from a common denominator (the old long */ll/ of aallarpoq going via the common base /dl/ or /tl/ to East Greenlandic /tt/ by progressive assimilation only, and to West Greenlandic voiceless /ll/ by progressive devoicing assimilation followed by regressive articulation with regard to all other features).

Such "generalizability" is certainly no proof that the truth about these developments has been disclosed; for one thing, the alleged common denominator accounting for the strange developments in East and West Greenlandic may be at best a matter of drift, meaning that the two dialects before moving apart have passed independently through related or even identical steps of change (which may not at all have been simultaneous in the two areas, cf. the related but much belated developments in 20th century Polar Eskimo, or the now ongoing regressive cluster assimilation in Canadian Eskimo which, as it were, repeats the development of West Greenlandic a century ago).

Much is totally unknown about the earlier stages of the Greenlandic language, and especially with regard to phonology this is true even of the historical period. If progress is to be made it requires the combination of comparative expertise and painstaking philological analysis of the old sources so outstandingly reflected in Bergsland's work, going hand in hand with utilization of recent explorations in Greenlandic dialectology. The present paper just outlines one of the controversial topics on which Greenlandic diachronic phonology is so rich.

REFERENCES

Bergsland, Knut 1955: *A grammatical outline of the Eskimo language of West Greenland* (Oslo)

Bergsland, Knut 1959: "The Eskimo-Uralic hypothesis", *Journal de la Société Finno-Ougrienne 61*, p. 1-29

Bergsland, Knut 1966: "The Eskimo Shibboleth inuk/yuk", *To Honor Roman Jakobson (Festschrift)*, p. 203-221 (The Hague)

Egede, Paul 1750: *Dictionarium Grönlandico-Danico-Latinum* (Copenhagen)

Fabricius, Otho 1804: *Den grønlandske Ordbog* (Copenhagen)

Fortescue, Michael 1983: *A comparative manual of affixes for the Inuit dialects of Greenland, Canada, and Alaska*, in: *Medd. om Grønland, Man & Society 4* (Copenhagen)

Holtved, Erik 1952: "Remarks on the Polar Eskimo dialect", *International Journal of American Linguistics XVIII*, p. 20-24

Petersen, Robert 1969/70: "On phonological length in the Eastern Eskimo dialects", *Folk 11-12*, p. 329-344

Petersen, Robert 1975: "Sprog og dialekter", in: *Grønland* (ed. P. Koch) p. 194-204 (Copenhagen)

Steenberg, K.J.O. 1849: *Grønlandsk Grammatik* (Copenhagen).

ARIPUC 20, 1986, p. 79-98

CAN "THE GREAT TONE SPLIT" IN THAI BE PHONETICALLY EXPLAINED?*)

JØRGEN RISCHEL

This paper deals with the so-called "Great Tone Split" in Thai, which in most dialects has taken a course contrary to the general contention that initial voiced and voiceless consonants, if they have any conditioning effect on tonogenesis, will trigger a difference between a lower tone in the former case, and a higher tone in the latter case. On this background it is discussed to what extent tonogenesis of this type is at all phonetically explicable at present.

I. INTRODUCTION

Explanation of sound change is one of the major tasks of current theoretical work, although researchers do not agree on how to deal with this issue. Among the types of change which seemingly have been explained successfully in recent years, a certain type of "tonogenesis" occupies a prominent place. Does it so rightfully?

The tone systems of numerous Southeast Asian languages are known to have developed out of an earlier pattern of syllable-final and syllable-initial features in the language in question.

---

*) *A substantial part of this paper was presented under the title "Tonogenesis in Thai: A Phonetic Paradox", at the Sino-Tibetan Conference, which took place in Bangkok in August 1985. For a more general presentation of some of the controversial features of tonogenesis in Thai, see Rischel (1985), especially p. 64-68. The reader is referred to that source for further references to the relevant literature on Thai (Bibliography loc.cit. p. 81-93).*

This obtains not only for Sino-Tibetan but also for Austro-asiatic and Tai languages (as shown by the work of Haudricourt and others), i.e., it is an areal feature. What has most generally attracted the attention of phonologists interested in "tonogenesis" (to use the excellent term coined by Matisoff) is the role played by initial consonants via the different (more or less automatic) pitch perturbations they induce on vowels, depending on their inherent manner features such as voicing, aspiration, and glottalization.

In discussing how such pitch perturbation by initials can give rise to tones one must constantly pay attention to the remaining, relevant properties of the syllables. Phonologization of initial pitch perturbation may well make a non-tonal language tonal (Modern Tibetan is interpretable as a case in point), but often the result is a tone split within an already existing tonal system, as is indeed the case in many East Asian languages and dialects. Obviously, sound change has both an articulatory, a perceptual, and a high-level organizational aspect. Therefore, one cannot hope to arrive at a unified explanation of any mechanism of tonal change without taking into consideration the psychological role played by already existing or incipient tonal distinctions (e.g. such that have to do with syllable quantity or types of syllable termination), and the degrees of freedom inherent in the tonal development. Thus, the conditions for tonogenesis must be considered anew in all detail for each particular language, and with reference to all chronological stages that seem relevant to the explanation.

Research on Tai languages and on the dialects of Thai proper within the last quarter of a century has shown that the interplay among such factors conditioning tonal development may lead to widely different tone patterns for closely related languages or dialects. The possibility of nonetheless finding general principles of development for these languages and dialects has been the object of important recent work (Egerod 1971, Brown 1975, Strecker 1979).

As for the role of initial consonant type, the by now trivial phonetic observation is that after a voiced initial (at least after a voiced stop) the pitch on the following vowel tends to start relatively lower than after a voiceless initial. It is, however, important to bear in mind that the degree of (negative or positive) pitch perturbation may depend on very specific articulatory properties of the consonants (including such that are not always indicated in phonemic descriptions of languages), so that one cannot expect to be able to predict the degree of pitch perturbation simply from a knowledge of the voicing state of the consonant. Thus, it is a controversial issue how the series of Tai consonants reconstructed by Li (1943) as glottalized voiced stops might be expected to influence the pitch (this particular issue has been discussed at length by Donna Erickson (1975) with reference to tonogenesis in Thai). More generally, it must be stated that there is not yet much of a general phonetic theory explaining why voiced and voiceless initials condition a pitch difference in the following segment.

Hombert, Ohala and Ewan (1979, p. 42-45) evaluate the relative merits of three types of phonetic explanation, viz. the Aerodynamic Hypothesis, the Halle-Stevens type of Vocal-Cord Tension Hypothesis, and the Vertical Tension Hypothesis: (1) The aerodynamic hypothesis is based on the observation that the initial airflow in voiced stops just after release is low compared to that of voiceless (particularly aspirated) stops, which would make the vocal cords vibrate more slowly just after release in the former than in the latter case. As pointed out by Hombert et al., this fails to explain why the pitch perturbation is normally seen to extend far into the vowel even though the aerodynamic effect should disappear shortly after the release. - (2) The first tension hypothesis takes its point of departure in Halle and Stevens' suggestion that there is a difference in horizontal vocal cord tension, voiced stops being articulated with slack but voiceless stops with stiff vocal cords, and that these states affect the Fo of adjacent vowels. This should be true preconsonantally as well as postconsonantally, which is definitely not corroborated by empirical evidence (Kohler 1985, p. 22 suggests that the similar effect of French unaspirated and English aspirated voiceless stops may be due to a shared feature of tension in initial position, however). - (3) The second tension hypothesis suggests that Fo perturbation has to do with a difference in vertical tension of the vocal cords. This is apparently corroborated by the observation that the larynx position is lowered for voiced as opposed to voiceless stops and that this difference in larynx position persists well into the following vowel. Unfortunately, Hombert et al. state: "the voiced stops' effect on Fo is like that of the sonorants, and it is the Fo after the voiceless stops which is perturbed, i.e. raised above them" (p. 45). Nevertheless, the authors lean to this hypothesis with a view to the difficulties facing the aerodynamic hypothesis.

Recent research on larynx height (Riordan 1980, Reinholt Petersen 1983) does not, however, support the assumption that there is a direct and simple relationship between larynx height and Fo perturbation in vowels after obstruents.

Kohler (1985, with references also to recent work by Haggard and others), on the contrary, finds more support for the aerodynamic hypothesis, possibly coupled with horizontal tension differences. He points out that most earlier data muddle the issue by mixing up sentence intonation contour with local pitch perturbation. - At present it seems safest to conclude that although there is obviously a principled connection between voicing state and Fo, the mechanism involved is not (at least not fully) understood. In fact, it is still an open question to what extent such pitch perturbation is at all an automatic phenomenon (whatever its mechanical cause), and to what extent it is actively programmed as a more or less independent variable in speech production.

It may be that even with absolute sameness of consonant articulation and voice quality the absolute degree and duration of pitch perturbation is variable, and hence language specific (though, on the other hand, such language specific differences may often be due to the effect of additional articulations, e.g. glottalization). To the extent that pitch perturbation is an independent variable it must compete with the exploitation of Fo for other purposes than cueing the category membership of initial stops, e.g., one may expect the duration of such perturbations in postconsonantal vowels to be shorter or less pronounced in languages with lexical tone. There is very little data, and it is not clear to what extent there is empirical support for that (cf. Hombert, Ohala and Ewan 1979, p. 41 with references). This line of reasoning, however, rather speaks against pitch perturbation as the origin of tone split (as in Thai), whereas it would be favourable to the explanation of how non-tonal languages may become tonal, cf. Modern Tibetan or the tonal dialect of Kammu (Khmu), to mention a Sino-Tibetan and an Austroasiatic case.

With these reservations in mind one must concede that numerous languages - even such that are already tonal - seem to strongly corroborate the idea that pitch perturbations caused by initials may develop into tonal contrasts (along with a loss of manner contrasts in the initials) so as to create the situation referred to by Brown (1975) as "voiced-low", viz. lower (or rising) tone after a formerly voiced initial, and higher (or falling) tone after a formerly voiceless initial, schematically: *baa > pàa as against *paa > páa.

II. THE SITUATION IN THAI

Also within the Tai family there are languages and dialects following the pattern outlined above. From a general linguistic point of view the dialects within Thai proper are particularly interesting in this context because the Siamese writing (dating back to the late 13th century) and old textbooks give very direct and interesting information about past stages. It seems rather well established (both from the spelling of Indic loanwords, from comparative studies within Tai languages, and from the classifications of the consonants in old textbooks) that there used to be a fourfold manner distinction in stops and at least a twofold distinction in other types of consonants, viz. (1) aspirated voiceless stops or continuants (called "high"), (2a) unaspirated voiceless stops, and (according to Li's reconstruction) (2b) glottalized voiced stops ((2a) and (2b) being both referred to as "mid"), and finally (3) plain voiced stops (called "low"; these are now aspirated voiceless stops, e.g. in Central Thai, having merged with category (1)). Thus, for exemplification: (1) *ph and *hm (or *m̥), (2a) *p, (2b) *ʔb, (3) *b and *m.

It is in itself a highly controversial issue what the terms "high", "mid", and "low" referred to originally, although it is off-hand tempting to relate them to a pitch-perturbating

effect (we expect fully voiced consonants to condition a lower pitch in the beginning of the following vowel than other consonants, and although it is not universally so that aspirated voiceless stops condition a higher pitch than other consonants, this may well have been the case in Old Thai). If one takes a look at the oldest textbook on Thai spelling and Thai verse (Phra Hora Athibodi's Cindamani or Cindamuni from the second half of the 17th century) one is indeed intrigued by the terminology. The letters that are still today referred to as "high" consonants are said to have a high "siang" (sound), all other consonants are said to have a mid (or medium) "siang" - this may be construed to have to do with Fo, but I presume that it rather refers to the presence versus absence of a high-pitched noise component. The consonants of mid "siang" are in turn divided into a set (the "low" consonants) that are said to have a low "kong", and the rest (the "mid" ones), which are said to have no "kong". The word "kong" is suggestive of a resonating, bell-like sound and must refer to the low-pitched full voicing which at one time was associated with "low" consonants; it is perhaps somewhat more surprising that no such voice is attributed to the "mid" stops reconstructed as voiced glottalized: *ʔb, etc. Maybe the voice really was less prominent here. - As I see it, it is not all that clear to what extent these pronunciations prescribed in a normative 17th century textbook reflect the everyday language of the area of that time, and to what extent they represent an artificial, learned recital of the individual letters of the alphabet based on the philological tradition associated with Indic loanwords in Thai. Thus the evidence can hardly be taken as proof that "low" *b etc. were still voiced stops in Thai proper at that time. Anyway, the terminology (for what it is worth without painstaking interpretation) suggests that pronunciations agreeing well with current reconstructions were at least known by learned people living some 300 years ago in the former capital of Sukhothai in Northern Thailand and taught by these to the administrators in the capital of Ayudhaya farther south. (Otherwise, the information - especially on pitch in the oldest Cindamani - is too controversial to be further dealt with here.)

As mentioned already: what we look for in tonogenesis having to do with initial consonants is evidence of a "voiced-low" development. Unfortunately, as demonstrated in the literature of the last decades, there are several Tai languages and dialects, including Modern Central Thai (Modern Siamese), that are not at all of the "voiced-low" type but rather show the opposite pattern. It cannot be stated quite briefly in what sense the pattern is opposite, since the actual tone contours depend on syllable structure and old prosodic distinctions. As for Thai, the old orthography distinguishes prosodic categories by tone marks such as "máj ẽek" and "máj thoo" (a third prosodic category being simply unmarked).

Now, to take Central Thai for illustration, the tone contour on a "live" syllable (i.e. an open syllable or a syllable ending in a resonant) without tone mark starts on a low pitch

(and rises) if the initial consonant is "high", but it starts on a higher (so-called mid) pitch after a "mid" or a "low" consonant. "Live" syllables with "máj ẽek", and likewise "dead" syllables (i.e. such that end in a stop), are low-pitched (so-called low tone) if the initial is "high" or "mid" but start on a high pitch (with or without a fall to low pitch, depending on syllable length) if the initial is "low". The relationship to initials is no more transparent in syllables with "máj thoo", since in Central Thai these all start on a high pitch but either stay high or exhibit a subsequent fall to low pitch, depending on whether the initial is "low" or belongs to one of the other categories (i.e. "high" or "mid"). One possible generalization is that in Central Thai "high" initials are accompanied by a feature of low pitch (somewhere in the syllable), whereas "low" initials are accompanied by a feature of mid or high pitch. ("Mid" initials side with "low" ones in tonally unmarked "live" syllables but with "high" ones elsewhere.) Although this pattern is pretty complex and requires a graphical representation in tone charts to make it stand out more clearly (see below), it certainly reveals a correlation of old voiced initials with a feature of high pitch, and reversely for old voiceless initials, i.e. dialects such as Central Thai are, in Brown's terminology, "voiced-high". - This I shall refer to as the tone split paradox.

Choosing Modern Central Thai rather than other dialects within Tai to illustrate the tone split paradox is well motivated. The acoustic phonetics of Central Thai has been studied so extensively, especially with regard to initial stop consonants and tones (by Abramson and others) that probably no other Southeast Asian tone language or dialect can be approached on the basis of existing phonetic knowledge of similar quality and detail. Furthermore, Modern Central Thai presents the paradox in its clearest possible form. However, in order to give some notion of the intriguing situation across dialects of Thai, I have entered the tonal reflexes of three different dialects (for some syllable types) in figure 1. The chart shows three types of syllable prosodies/terminations for Proto-Thai, viz. unmarked "live" syllable, "máj ẽek"-syllable, and "máj thoo"-syllable ("dead" syllables are left out for simplicity of exposition), and each of these is combined with four types of initials, each symbolized by a representative (for "high", "mid voiceless", "mid glottalized voiced", and "low" consonants, respectively - note that "low" b- is now ph- or p- with a voiceless consonant in all dialects, whereas nasals and liquids, such as m-, l-, are preserved and give direct testimony of the voiced status of old "low" consonants).

The dialects shown are limited to just three for which I have first-hand experience with tones and their phonetic characteristics, viz. the eastern type of Northern Thai (from the start a language of its own: Khammyang), Central Thai, and Southern Thai as spoken on the West coast in the extreme South. The tonal labels are chosen so as to give a reasonable fit both

Prosodies associated with "live" syllables in Old Siamese

Initial:		zero	1 ("máj èek")	2 ("máj thoo")
H	*ph-, *hm-, etc.	N: lo ris C: lo ris S: hi-mi ris	N: mi C: lo S: hi-mi ris	N: hi, gl C: hi fall, gl S: mi
M_1	*p- etc.	N: lo ris C: mi S: lo-mi ris*	N: mi C: lo S: lo-mi ris*	N: hi, gl C: hi fall, gl S: mi
M_2	*ʔb- etc.	N: mi ris C: mi S: lo-mi ris*	N: mi C: lo S: lo-mi ris*	N: hi, gl C: hi fall, gl S: mi
L	*b-, *m-, etc.	N: mi ris C: mi S: mi fall	N: mi fall C: hi fall, gl S: lo ris*	N: hi fall, gl C: hi peak, gl hi ris, gl S: lo

Figure 1

Some tonal reflexes in Thai dialects

N = Northern Thai as spoken in the Lampang area.
C = Central Thai as spoken in the Bangkok area.
S = Southern Thai as spoken in Satun.
"fall" = falling; "gl" = irregularity of glottal vibration; "hi" = high; "hi-mi" = higher mid; "lo" = low; "lo-mi" = lower mid; "mi" = mid; "peak" = peaked: with a (small) rise+fall; "ris" = rising.
Asterisk: analysis uncertain.

*Examples of the segmental reflexes (in Central Thai) of old initials (all examples given here having old "zero" prosody): H */phi:/ > /phi:/ 'ghost'; */hmi:/ > /mi:/ 'bear'; M_1 */pi:/ > /pi:/ 'year'; M_2 */ʔdi:/ > /di:/ 'good' (minimal pair with */ʔb/ missing here); L */bi:/ > /phi:/ 'to be fat'; */mi:/ > /mi:/ 'to have'. Note that word types such as /phi:/ 'ghost' and /phi:/ 'to be fat', or /mi:/ 'bear' and /mi:/ 'to have', contrast tonally in all three dialect areas.*

to my own impressionistic data and (especially with regard to Southern Thai, for which my own data is very meagre) to the tone charts in Brown (1965). I deviate from Brown on one point for Northern Thai: I hear (and observe on Fo-tracings) a high, strongly falling tone in "máj thoo"-syllables with initial "low" consonant, where Brown indicates a mid falling tone. (As for Southern Thai, I am not sure of the nature of the distinction Brown makes between two kinds of rather low rising tones in Satun Thai, and I simply use a somewhat arbitrary labelling: lower-mid rising vs. low rising to reflect it, but that is immaterial to the general point to be made with the use of this chart.) - The two options given for Central Thai "máj thoo"-syllables with a "low" initial reflect a difference between age groups (cf. Henderson 1982).

It will be apparent from figure 1 that it is, on the one hand, hard to set up common denominators for the tonal reflexes of "zero", "máj ēek", and "máj thoo" across dialects, and, on the other hand, equally difficult (if not more difficult) to set up common denominators for the tonal reflexes of the four types of initials across dialects. There is some suggestion in the chart that "máj ēek" has, relatively speaking, been associated with some kind of lowering compared to zero tone (this is in part contradicted by Southern Thai, but that has to do with a merger of zero- and "máj ēek"-prosodies in syllables with a non-low initial). There is definitely evidence for an association of "máj thoo" with relatively high tone followed by glottalization (but again this is contradicted by Southern Thai). - The Central Thai falling reflex of "máj ēek" (with a low initial) is clearly a coalescence with "máj thoo" (with a non-low initial).

As for the initials, there is some indication that the higher one gets on the scale from "low" to "high" consonant the lower is the tone (in some sense), but this does not hold for "máj thoo"-syllables in Northern and Central Thai, and in Southern Thai we find the opposite in "máj ēek"-syllables! There are (in part ingenious) explanations available for much of this, but off-hand the pattern does not seem particularly inviting from the point of view of deriving Modern Thai tone contours from the old syllable prosodies by any kind of simple algorithm, let alone a phonetically plausible one. Most of the chart runs counter to the expected direction of Fo perturbation. The only evidence in favour of "phonetically natural" Fo perturbation by initials is the relative lowness associated with "low" consonants in Southern Thai, but then again: why is the tone of zero-syllables falling with "low" but rising with "high" consonants?? If initials mechanically cause a perturbation of the initial part of the tonal contour we should expect the opposite. Thus, there must at least be some additional mechanism (be it articulatory or perceptual) at work.

III. EXPLANATIONS OF THE TONE SPLIT PARADOX IN THAI

There are several ways in which explanations of the tone split paradox can be attempted. The strongest hypothesis is that the speech production mechanism per se is solely responsible for the nature of the tone split via pitch perturbations stemming from the intrinsic properties of the initial consonants, but that these perturbations may - for physiologically explicable reasons - deviate from the a priori expected pattern in CV-sequences (voiced-low vs. voiceless-high) and may eventually lead to the opposite, viz. a voiced-high situation. An explanation along these lines was suggested some twenty years ago by Brown (1965), his assumption being that there will be a greater or lesser force stretching the vocal chords and the underlying thyro-arytenoid muscles depending on the behaviour of the arytenoid cartilages: the pull is strongest in consonants spoken with a wide open glottis such as aspirated stops, and weakest in fully voiced consonants. For the old "mid" initials in Thai Brown posits a laryngeal gesture which should give an intermediate degree of pull. If, now, the thyro-arytenoid muscles resist the pull by different degrees of contraction, and this contraction spills over into the following part of the syllable, the effect will be high, mid, and low pitch, respectively, with the old "high", "mid", and "low" initials in Thai. If, however, the muscles give in to the force operating on them, they must relax most for the articulation with open glottis and least for fully voiced articulation, the effect being low, mid, and high pitch, respectively.

Brown's explanation (which was apparently overlooked by Hombert, Ohala and Ewan 1979, cf. their references p. 38 bottom) is an ingenious one and interesting in the rigour and explicitness with which it makes appeal to laryngeal mechanisms in explaining prosodic sound changes. Nonetheless, it should be noted that we have no independent evidence for the contentions it makes about alternative configurations giving rise to different tones in different dialects. Moreover, these contentions are clearly interpetable only for aspirates, which do in fact seem to behave rather differently in different languages (although the underlying mechanisms are not yet well understood). As for fully voiced consonants, it is unclear to me why these would have a different degree of pull and hence different pitch-perturbating effects in - as it were - inverse relationship to the way the aspirates behave: why would the alleged mechanisms tend to place the aspirates as one extreme, and the fully voiced consonants as the other extreme along the scale of pitch perturbations?

Another category of possible explanations (to some extent already suggested for Tai) involves the assumption that the manner distinctions in initial consonants did not cause tone split directly via Fo-perturbations but that features of the consonants caused voice quality differences (cf. the so-called "register" differences in several Mon-Khmer languages) in the

vowels or in the whole remaining parts of the syllables, and that these derived features of articulation in turn conditioned pitch distinctions to arise. The resulting pitch contours might then be of quite a different kind than those derivable directly from the Fo-perturbating influence of initial consonants. This is a very interesting category of explanations; unfortunately, it is hard to provide positive evidence for the alleged intermediate steps, especially in languages which - like Thai - do not exhibit dialects or attested earlier stages with "register" conditioned by initials.

Then, thirdly, one may assume that the old consonant categories did in fact trigger the tone split by causing straightforward pitch perturbations (of the types expected) in all the dialects, so that these all had the potentials of becoming "voiced-low", but that other factors intervened and more or less changed the incipient pitch movements beyond recognition. Let us assume that the tone split always started with the phonetically expected pattern: relatively lower pitch after formerly voiced than after formerly voiceless initial, the resulting tones being at first direct reflexes of these relative pitches and only "changing place" later. It seems quite possible that, at the time when the initial Fo perturbations developed into a tone split, the tonal difference between the three prosodies marked by zero, "máj ēek" and "máj thoo" was, at least in part, a difference of level ("máj thoo" marking the highest tone). Now, as stated by Gandour (1974, p. 348), "A (...) plausible hypothesis, for the Tai language family at least, is that lexical contour tones develop from already existing level tones in order to maximize perceptual distance in the tone system." If, accordingly, we assume that the pitch perturbations first developed into further levels, these levels would multiply with the levels for the syllables with the three old prosodies, and the tonal space would end up being just over-crowded, thus necessitating the development of contours. In view of the complexity of the tonal pattern of Thai we would have to assume that such a differentiation of level tones into contours could happen in a variety of ways (being in part dialect specific). The resulting contours would undoubtedly in some cases cross each other (or cross a mid level tone), and as a consequence of this it might happen that certain syllable types in certain dialects came to exhibit a spurious association of relatively high pitch (somewhere in the course of the contour) with old "low" initials, and vice versa for old "high" initials. This, then, would qualify as the "voiced high" situation. - We do not know, however, that this is the way things happened (for one thing it is anything but certain that the tones from the old prosodies were essentially level tones). - Alternatively, as suggested by Nina Thorsen, personal communication) one might in principle imagine that the rising and falling slopes after different initials developed directly into corresponding contour tones, which later simplified into level tones in some instances. Thus, it would be logically possible for syllables with old voiced initial and "máj thoo" to first develop a strongly rising tone which later simplified into a high (peaking) tone. This line

of explanation may seem less plausible in the context of the entire Thai scenario, however.

Within this third category of explanations an alternative version says that even with initial pitch perturbations conforming to the voiced-low, voiceless-high scheme, the resulting tonal pattern may well deviate from this scheme from the very start of the perceptual reinterpretation procedure leading to phonemicization of tone ("tonogenesis"). Lea (1973, p. 64) speculates what might happen if the initial Fo slopes were translated into high or low tones: "Since Fo rises after voiced consonants, the following tone may appear high since the slope into it is upward. Similarly, the fall after an unvoiced consonant may lead to an interpretation of the following tone as low." This formulation differs significantly from the just mentioned suggestion by Nina Thorsen in that the perceptual restatement from falling to low, or from rising to high, is taken to coincide with the phonemicization rather than being a later reshaping of the tones. (Lea does not himself believe in this conjecture. He finds that "a more reasonable conjecture would be that the relative Fo values within the vowel (and not the preceding slope into the vowel) affects its interpretation as a high versus low tone", which may be construed to agree with the preponderance of low tone from voiced initial vs. high tone from voiceless initial among the cases of tonogenesis cited in the general literature. Gandour 1974, p. 348 finds the latter conjecture dubious on the basis of measurements of Fo in initials in syllables with different tones.)

Brown (1975) some ten years ago replaced his earlier physiologically based hypothesis by a hypothesis of this "third" category to account for the "voiced-high" dialects, although strictly speaking this is a hypothesis involving specific assumptions about both production and perception strategies. His first claim is that syllable initial pitch lowering after voiced initials tends to be compensated for by higher pitch in the last part of the syllable, compared to the tail of the pitch contour in syllables with voiceless initials. If, now, speakers listen to the whole pitch contour, the integrated pitch impression will be similar for both syllable types, but if the listener narrows his focus down to part of the syllable, the impression of pitch may be different depending on what part of the syllable is in focus. Thus the pitch will be perceived as either lower or higher with a voiced initial than with a voiceless initial, i.e., there is a possibility for one and the same type of pitch perturbation to lead to either the voiced-low or the voiced-high situation. Brown himself has characterized his second hypothesis as "sheer speculation", but more recently Strecker (1979) has spoken in favour of it from a general linguistic point of view (without taking a stand on its validity vis-à-vis current phonetic theory). - It should be added that both Brown and Strecker emphasize the importance of other driving forces in the tonal development, viz. such that have to do with the enhancement of contrasts within the tone system, or with alleged general tendencies for pitch contours within a tonal system to be shaped

in relation to each other (e.g. Strecker's principle that a high-falling tone falls more than a low-falling tone), and that they do this with a view to the complex developments in the various dialects for which they account so impressively.

As a phonetician one is certainly struck by the elegance and explanatory power of Brown's second suggestion, but one cannot fail to question its empirical and theoretical status vis-à-vis current general phonetic research. This applies to both of its components: production and perception. As for production, I do not know of solid evidence for the claim that the pitch rise after voiced initials tends to reach a level above that of the pitch after voiceless initials. Most studies indicate that Fo after voiced stops remains somewhat lower than Fo after voiceless stops throughout most or all of the vowel, though the Fo traces may cross in individual cases (cf. e.g. figure 2 in Hombert, Ohala and Ewan 1979, p. 40). Brown himself refers to graphs of English and Modern Standard Thai, but I think his interpretation of these is controversial (note also that Lea (1973), to take one source he quotes, does not draw the conclusions drawn by Brown on the basis of Lea's graphs). As for perception, general phonetic theory may not exclude the possibility of listeners narrowing their focus down to part of the pitch contour, but there is hardly any empirical evidence so far that this is the way dialects become tonally different, if such perceptual strategies are at all available in linguistic performance.

In short, like Brown's earlier hypothesis the more recent one attempts to account for developments in the past for which we have only indirect evidence, by reference to alleged phonetic mechanisms for which we have so far neither a comprehensive general theory, nor indisputable empirical evidence from directly accessible languages. This is obviously not an altogether desirable situation, although it should not be overlooked that if Brown's account is open to criticism it is exactly because he takes the demands for explicitness and rigour in the advancement of explanations more seriously than linguists often do. It is of paramount importance both for the progress of historical linguistics and for the progress of general phonetic theory that these disciplines be coupled together. Well attested cases of sound change provide an excellent testing ground for general phonetics, and conversely, historical linguistics should employ the most recent advances in phonetic theory (rather than some 19th century notions about phonetics) to ensure lasting progress. The tonogenesis issue is a good case in point, and seen in this general perspective Brown's use of highly sophisticated phonetic argumentation is indeed laudable, but it goes without saying that if both the data (viz. unrecorded changes taking place in the past) and the theory adduced to define and explain the data must be construed to fit each other, we are on shaky ground. I think it is important to remember Brown's very cautious remarks about the speculative nature of his own hypothesis.

In a more general sense Brown may be right in tying the explanation of the tone split paradox to the interplay between the production of perturbated pitch curves and the perception and categorization of these. To illustrate the possibility of advancing alternative hypotheses within this realm of thinking, I shall permit myself to refer to a suggestion which was ventured by John Ohala in discussions with me about the tone split paradox. His suggestion (which certainly was not meant as a serious theory about tonogenesis in Thai but just as a technically possible type of argumentation) hinges on the tone level contrasts already existing in the language. Assume that there is a certain frequency value which is the boundary (threshold) between what is categorized by the listener as a lower and a higher phonemic tone. A listener hearing a syllable whose pitch is only just below this boundary frequency may categorize it as having the lower tone, or he may construe it to be a higher-tone syllable if the initial consonant is of a type (voiced stop, for example) which might be responsible for a perturbation lowering the pitch across the boundary. Assuming that the syllable was actually intended as having the lower tone, we would have in the latter case a faulty categorization on the part of the listener which might eventually lead to tonal restatement for the lexical item in question. Something similar (mutatis mutandis) might conceivably happen with syllables spoken on pitches just above the boundary value. - Obviously, this kind of reasoning can only explain how some words of the language may switch tone due to a perceptual strategy at variance with the intentions of the speaker; it remains to be explained how such a tonal switch might be generalized to all words with all types of syllable structure and all the tonal categories of Old Siamese. The ultimate interest of this specific line of argumentation depends on how safely we can predict the occurrence of tonal recategorizations near the boundary values between contrastive tones; there is some slight evidence from data in a recent paper by Abramson and Erickson (1978) which may be construed to suggest that occasional perceptual switches in categorization having to do with the nature of the initial consonants do indeed happen near tonal boundaries in Modern Central Thai. However, as emphasized by Ohala himself, it is not the case that there is solid support for the specific hypothesis outlined above, which - as said already - was advanced only as an illustration of a certain category of putative explanations.

IV. THE TONE SPLIT PARADOX IN A WIDER PERSPECTIVE

As I see it, the essential thing about such proposals is that they make specific claims about phonetically and structurally defined situations in which a given type of sound change (e.g. voiced initial giving high pitch) may take place, i.e., they are predictive in this restricted sense (not, of course, in the sense that they predict whether the sound change will actually take place in a given language or dialect at a given time).

This means that they are open to empirical testing of a weak kind: if in the languages of the world there are many languages exhibiting the change under the specific conditions set by the hypothesis but few or no languages exhibiting it in which these conditions are not met, then obviously the typological findings corroborate (though not prove) the hypothesis, and vice versa. Now, if I understand both Brown's and Ohala's arguments correctly, it seems that they make different predictions: the latter definitely sets off tone split from other types of tonogenesis and accounts for the voiced-high situation only as a special case of tone split; Brown, however (both in his first and in his second hypothesis) posits phonetic mechanisms and strategies which in principle should be able to create the voiced-high situation irrespective of the previous status of the language as either tonal or non-tonal.

Hombert, Ohala and Ewan (1979, p. 54) speak of "the occasional correlation between higher tone and originally preceding voiced consonant (vs. lower tone and voiceless consonant), where just the reverse would be expected". The area notorious for such exceptions is Southeast Asia, and the languages or dialects in question typically belong to those language families which have had prosodies (tones or "terminals") long before the initials affected the tonal pattern, i.e., these are cases of tone split. Let us imagine first that a careful search for divergent tone patterns in remote dialects of newly tonal languages (e.g. of the Tibeto-Burmese and Mon-Khmer families) would reveal that a sizeable proportion of these dialects are in fact voiced-high just as is the case in Tai. That finding would seriously detract from the explanatory power of hypotheses which - like Ohala's suggestion - are crucially sensitive to the initial tonal conditions of the language undergoing tonogenesis. If, on the other hand, it is true that newly tonal languages in which tones reflect the previous voicing state of initials are (with few and marginal exceptions) voiced-low across the board, this spells trouble for the category of explanations represented by Brown's first and second hypotheses. It is not that these are shown to be wrong, but it takes a powerful, additional explanatory device to account for the typological difference observed in a wider array of languages.

Thus, quite apart from considerations of the inherent phonetic plausibility of one or another hypothesis, we have a statistico-typological approach to the general issue. As for the tone split paradox, one thing that might be done to get a more solid empirical foundation for the formation of general phonetic hypotheses about tonogenesis is to search carefully for tone languages or dialects all over the world which seem to have developed high-pitched tone from voiced initials, and to see whether these are practically all instances of tone split (except for cases which obviously have special explanations), or whether there are also clear-cut instances of non-tonal languages becoming tonal via voiced-high tonogenesis.

Unfortunately, the situation is not as simple as suggested above with regard to the possibility of revealing universal tendencies in tonogenesis under specified structural conditions. If we find a statistical preponderance of a certain phenomenon, or of a correlation among certain phenomena, that in itself does not tell us anything about universals of human languages, of course. We have to know to what extent the phenomenon or correlation in question is a truly universal tendency and as such triggers mutually independent spontaneous developments, and to what extent the various occurrences we have observed, are linked by some kind of interrelatedness among the languages or dialects under study. This is a rather trivial statement, to be true, but it may be appropriate to consider the problem in its general outline here since its importance is not always made explicit in connection with purely theoretical argumentation about the mechanism(s) of tonogenesis.

Languages may be interrelated by belonging to the same family or sub-family, i.e., by genetic relatedness, which in the strict sense implies that the correspondences among them are consistent with the assumption of a common parent language, though the scenario is often not all that transparent. As for South and Southeast mainland Asia, we know from authoritative comparative research that there are five well-defined families or sub-families which include tonal languages, viz. Chinese, Tibeto-Burmese, Tai (-Kadai), Mon-Khmer (Austroasiatic), and Indo-Aryan. There is no proof of genetic relatedness (so far) between families, except for Chinese and Tibeto-Burmese. In any case, genetic relatedness across the five groups enumerated here belongs to a very distant past and can hardly be directly relevant to recent tonogenesis taking place essentially within the last millennium. If, however, we look within each of the five families it may be rightfully claimed that tonogenesis in each particular language or dialect should first be viewed in a comparative perspective: as possibly derived from a common ancestral source together with tonogenesis in related languages or dialects. It may be that tonogenesis happened already in a parent language, or it may be that languages of a genetically defined group carry a structural and/or phonetic predisposition for developing in a specific direction, i.e., that we have an instance of "drift", whatever that exactly means. (Candidates for "drift" are, e.g., the tendency toward developing a binary tonal distinction caused by initials via a "register" difference, as in some Mon-Khmer languages, and the tendency toward developing tonal distinctions caused by "finals".)

This suggests, not surprisingly, that evidence for universals should include samples of languages belonging to several different families. It is equally essential, however, that there are samples from one and the same family if the languages in question have developed in significantly different ways so that they do not just appear as projections of an old, shared development or as manifestations of a common "drift" (cf. that Mon-Khmer comprises tonal and non-tonal dialects even within

one and the same language, and that this language family comprises tonal languages of the simplest possible type with a binary distinction as well as a language of a highly complex type with many contrasting contour tones: Annamese).

In addition to genetic relatedness we have the question of areal features. The area stretching from the southeastern coast of mainland Asia to the regions just north and south of the Himalayas, is today one vast tone language area. Why is there a tonal area of such extension, and why does it include several language families? In Southeast Asia there is apparently one layer of tonogenesis having to do with old finals, and another layer having to do with old initials. Why is the latter, rather than the former, manifested in recent cases of tonogenesis, and why is it more widespread geographically? Tonogenesis conditioned by initials is known outside this area, but to what extent is Southeast Asian tonogenesis from finals matched elsewhere? Is it somehow related in type to what happened in certain Scandinavian languages or dialects?

With such an areal feature as tonogenesis there may well be "drift" within each language family or sub-family, but this cannot possibly account for unrelated languages sharing such a development. How is the tendency to develop tones diffused across dialect or language boundaries? It is too far-fetched to assume that it is a coincidence that several unrelated languages of this area favour tonogenesis (cf. that Chinese, Annamese, and Tai languages share certain developments causing a high degree of tonal complexity which should be viewed in the light of the intimate and long-lasting contact known to have existed between these languages).

The general theory of language change thus faces a twofold task: to reveal and explain universal tendencies and to account for the ways in which innovations may be imported from one language into another. Phonetically (i.e. physiologically and/or perceptually) "natural" tonogenesis which occurs spontaneously and independently in many places is one thing; borrowing of tonal contours or perhaps even tonal distinctions is something very different. We do not know what it takes for a tonal pattern, or for a set of (possibly arbitrary) substitutions between segmental and tonal distinctions, to be copied from one language into another, and to what extent this is possible. "Crazy" types of tonogenesis might well have to do with language contact, e.g. borrowing of a tone contour which is introduced in certain items of the vocabulary (according to etymology) without any internal phonetic motivation and eventually changes the overall pattern.

V. CONCLUSION: THE COMPLEXITY OF TONAL DEVELOPMENTS

With our present knowledge it does not seem very likely that the tone split paradox can be explained solely on the basis of production and perception mechanisms and strategies operating on initial consonants and on pitches that are perturbated due to the presence of these consonants. The entire phonological system (the conditions to which system-internal forces respond) must be taken into due consideration when explaining the development of tones. Brown, Egerod, and Strecker attempt to do this (in various ways), but there are too many unknown factors involved e.g. with reference to the exact articulation of initials. Note, for example, that initial *b may have had a very different effect on Fo depending on whether it was at some chronological point just voiced or voiced aspirated ("breathy") or possibly implosive (Hombert, Ohala, and Ewan 1979 cite data according to which the onset Fo of the vowel after a breathy voiced consonant is markedly lower than after a plain voiced one, and other data according to which Fo after an implosive may be higher than after a plain voiced stop). - Gandour (1974) has actually suggested a reconstruction */bɦ/ at some chronological point, on the basis of evidence from Southern Thai (Egerod 1961), similarly Egerod (1971).

Moreover, there may be irretrievable differences in the relative chronology of events which have made otherwise closely related dialects develop rather differently on this point. In this context it is very important to take into consideration the role played by linguistic variation, in particular by coexisting pronunciation norms associated with different generations, different social groups, or different styles of speech. Unfortunately, there is but little available knowledge about such parameters of variation in the earlier stages of Thai, although the linguistic diversity (involving several dialects of Thai proper) is known for sure to be of considerable age in Thailand, so that bilingualism (in a wider sense) must also be considered as an essential factor in tonogenesis (possibly promoting seemingly ad-hoc system-internal readjustments such as tonal flip-flop?).

As for the dialects of Thai, these are mutually so closely related that it is rather tempting to account for their differences in tonal patterns as relatively late - though in part radical - deviations from a more or less common basis which may have included the phonetic prerequisites for development of a voiced-low tone split. But even under such an extreme hypothesis about linguistic homogeneity there are many degrees of freedom in the pattern one may posit for the old language. Take the tone marks of Old Siamese, for example: did these represent tone levels, rising or falling tone contours, phonation types involving features other than pitch, or what? The modern reflexes (cf. figure 1 above) very strongly suggest that the prosody indicated in Thai by "mâj thoo" had a component of syllable-final glottalization (or laryngealization)

all along. As for the prosody indicated by "máj ẽek", it may well at one point have had a component of final aspiration or breathiness (this is consistent with the tendency toward final low Fo in many modern reflexes of this syllable type), although Brown 1965, p. 38 & 52 just posits voicing finally in Ancient Thai. But it is hard to get beyond that.

As for the consonant categories called "high", "mid", and "low" the strong assumption of pitch perturbations corresponding to these terms does not tell us how these pitch perturbations eventually modified the existing tones (before the emergence of phonemic tone splits): did they result in pairs of allotones differing in overall level, for each tone, or was the result a split into allotones differing in the direction of the pitch movement? Did allotones differing first in pitch level later develop into contours because of influence from the rest of the tonal system? Even given a fully specified set of initial conditions we cannot fully predict what course the tonal developments will take, and of course we cannot predict whether tonogenesis will occur at all. General phonetic theory helps very much to understand what has happened, but we have to know more about the phonetics of past stages.

Basically, Marvin Brown and the other pioneers have been doing what must be done, viz. to combine scrutiny of the philological information available (from the spelling system, from old grammars, from ancient loanwords whose literary route is known, etc.) with the application of modern linguistic and phonetic theories. This line of research may be pursued further, however, especially with regard to an inclusion of the most recent work on speech physiology and perception and a reinterpretation of the evidence from old sources such as the Cindamani referred to earlier in this paper.

VI. FINAL REMARKS

Two further observations should be made concerning the relationship between initial Fo perturbation and tone: (1) It should not be overlooked that the "low" consonants of Thai include sonorants as well as (old voiced) stops. Sonorants are not supposed to lower Fo significantly but still those have exactly the same tonal reflexes as "low" stops. This suggests that the tone split should be explained with reference to the perturbating effect of non-low rather than low consonants.

(2) As mentioned earlier, there is some evidence that the duration of the Fo perturbations caused by the presence of preceding consonants is limited in languages in which the vowels carry lexical tones. Since there may nevertheless be considerable differences in vowel-initial Fo as conditioned by the preceding consonant, it seems as if the very start of the vowel is less crucial for the perception of lexical tone than the remainder. Now, if we look at the tonal patterns of modern Thai dialects (i.e. the patterns emerging from the combined effect of old final prosodies and old manner differ-

ences in initials), we may well venture the contention that on the whole it is the final part of the contours that is most important for the characterization of lexical tones in terms of meaningful features such as "rise", "fall (±glottalization)" etc. (I think this is consistent with the extensive research on the acoustics and the perception of Thai tones by Abramson, Gandour, and others). It would be interesting to know to what extent this is true generally. There is good evidence (Brown 1965, Egerod 1971) that tone in languages such as Thai is a property associated with the final part of the syllable and hence forming a joint system with such properties as checked vs. unchecked syllable, breathy vs. plain vs. glottalized vowel offglide, and long vs. short vowel. If there is in fact a clear-cut perceptual division between syllable initial properties being interpreted as non-prosodic irrespective of their phonetic nature, and syllable central and final properties being in part interpreted as prosodic, we may see the "phonologization" of initial Fo perturbations as a categorial shift from perceiving these Fo movements as part of (or coarticulation with) the syllable initial to perceiving them as part of the syllable remainder. (This formulation deliberately leaves it entirely open whether there is possibly a related categorial difference between properties associated with the vowel per se and properties associated with its termination, duration being interpretable as a "central" vowel feature in some languages and as a "terminal" feature in others, which might account for its varying status as a prosodic or non-prosodic property.)

REFERENCES

Abramson, A.S. and Erickson, D.M. 1978: "Diachronic tone splits and voicing shifts in Thai: some perceptual data", *Status Rep. Speech Res., Haskins Labs. 53.2*, p. 85-96

Brown, J. Marvin 1965: *From Ancient Thai to modern dialects* (Soc. Sc. Found. Press of Thailand, Bangkok; new edition Bangkok 1985)

Brown, J. Marvin 1975: "The great tone split: did it work in two opposite ways?", in: *Jimmy G. Harris and James R. Chamberlain (eds.): Studies in Tai linguistics in honor of William J. Gedney* (Office of State Universities, Bangkok), p. 33-48

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

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

Erickson, D.M. 1975: "Historical implications for an historical account of tonogenesis in Thai", in: *Jimmy G. Harris and James R. Chamberlain (eds.): Studies in Tai linguistics in honor of William J. Gedney* (Office of State Universities, Bangkok), p. 100-111

Gandour, J. 1974: "Consonant types and tone in Siamese", *J. Phonetics 2*, p. 337-350

Halle, M. and Stevens, K.N. 1971: "A note on laryngeal features", *MIT Quart. Prog. Rep. 101*, p. 198-213

Hombert, J.-M., Ohala, J.J., and Ewan, W.G. 1979: "Phonetic explanations for the development of tone", *Language 55*, p. 37-58

Kohler, K.J. 1985: "Fo in the perception of lenis and fortis plosives", *J. Acoust. Soc. Am. 78*, p. 21-32

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

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

Matisoff, J. 1973: "Tonogenesis in Southeast Asia", in: *Larry Hyman (ed.): Consonant types and tone* (Univ. of Southern California), p. 71-95

Petersen, N. Reinholt 1983: "The effect of consonant type on fundamental frequency and larynx height in Danish", *Ann. Rep. Inst. Phon., Univ. Copenhagen 17*, p. 55-86

Riordan, C.J. 1980: "Larynx height during English stop consonants", *J. Phonetics 8*, p. 353-360

Rischel, J. 1985: "An appraisal of research in the phonetics and phonology of Thai", *Ann. Rep. Inst. Phon., Univ. Copenhagen 19*, p. 43-93

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

INSTITUTE OF PHONETICS

JANUARY 1, 1985 - DECEMBER 31, 1985

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:
Jan Leon Katlev (from February 1st)
Peter Molbæk Hansen, cand.mag.
Peter Holtse, cand.phil. (from April 1, 1984)

TEACHING ASSISTANT:
Ole Nedergaard Thomsen

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 (until June 1st)
Professor John J. Ohala, Ph.d., University of California, Berkeley (until June 30th)
Professor Manjari Ohala, Ph.d., San José State University (until June 30th)
Amon Thavisak, M.A., Mahidol University, Bangkok.

II. PUBLICATIONS BY STAFF MEMBERS AND GUESTS

Eli Fischer-Jørgensen: "Some basic vowel features, their articulatory correlates, and their explanatory power in phonology", in: Phonetic Linguistics, Essays in Honor of Peter Ladefoged (ed.: V. Fromkin), (Academic Press, Orlando et al.) 1985, 79-99

Peter Molbæk Hansen: "Kanel, model, adel, sadel - og andre vanskeligheder ved at lære en maskine at læse højt", Teletek-nik 4, 1985.

Birgit Hutters: "Vocal fold adjustment in aspirated and un-aspirated stops in Danish", Phonetica 42, 1985, 1-24.

Birgit Hutters: Review of Pieter van Reenen: Phonetic Feature Definitions. Their Integration into Phonology and their Relation to Speech. A Case Study of the Feature NASAL. Foris Publications, Dordrecht 1982, 196 pp., Acta Linguisti-ca Hafniensia 19, 1985, 127-130.

Jørgen Rischel: "Was there a fourth vowel in Old Greenlandic?", IJAL 51, 4, 553-555.

Jørgen Rischel: "Resens tysk-grønlandske ordliste" (Robert Petersen & Jørgen Rischel ed. with introduction), in Grønland (theme: "Nyt fra 1600-tallet", New information about the 17th century), 1985, fasc. 5-6-7, 156-192.

Nina Grønnum Thorsen: "Intonation and text in Standard Danish", J. Acoust. Soc. Am. 77, 1985, 1205-1216.

Nina Grønnum Thorsen: Review of Nico Willems: English Into-nation from a Dutch Point of View. Netherlands Phonetic Ar-chives, Vol. 1, 1982, J. Phonetics 13, 1985, 357-358.

Nina Thorsen and Oluf Thorsen: Fonetik for Sprogstuderende, 3rd revised edition, 7th printing, Copenhagen 1985, 170 pp.

III. GUEST LECTURES AND SEMINARS

January 23 - John Jørgensen, Odense university: "Forskelle mellem børns og voksnes konsonantgrupper" (Differences between consonant clusters in children and adults)

April 17 - Jan Katlev: "Dansk allegro-fonologi" (Danish allegro-phonology)

May 17 - John Ohala and Jørgen Rischel: "The relations between phonetics and phonology"

October 4 - Professor Nelly I. Cubar (University of the Philip-pines): "Tagalog phonology: a generative approach"

October 25 - Inger Karlsson (Royal Institute of Technology, Stockholm): "Akustiska glottogram för friska och patologiska röster" (Acoustic glottograms of healthy and pathological voices)

November 1 - Professor Wilhelm Vieregge (Catholic University, Nijmegen): "Probleme beim Transkribieren normaler und pathologischer Sprachäusserungen" (Problems in transcribing normal and pathological utterances)

November 22 - Robert McAllister (Stockholm University): "A theory of speech motor control in the light of recent research in labial co-articulation"

December 13 - Anders Löfqvist (Lund University): "Laryngal artikulation - kinematik, kontrol, koordination" (Laryngeal articulation - kinematics, control, coordination)

IV. PARTICIPATION IN CONGRESSES, ETC.

Peter Molbæk Hansen and Jan Katlev participated in a conference on Spoken Danish under the auspices of the Danish Research Council at Hindsgavl Slot, April 13-15.

Peter Molbæk Hansen participated in two COST-209 meetings, one at the British TELECOM, Martlesham, United Kingdom, June 13-15, and one at Copenhagen University, November 13-14.

Peter Molbæk Hansen participated in a conference on Linguistics and Informatics under the auspices of the Danish Research Council, Middelfart, December 1-3.

Peter Holtse participated in the COST-209 meeting at Copenhagen University, November 13-14.

Peter Holtse gave an invited paper: "The phonetic basis of speech technology" at the Speech Technology Conference, held by the Danish Automation Society, November 6.

Peter Holtse gave a status report on "Speech Synthesis and text-to-speech conversion" to the Danish Society for the Blind, September 11, and read a paper on "Speech recognition and speech synthesis" (with P. Dalsgaard, Aalborg University), at a Danish Research Planning Council seminar, May 29.

Birgit Hutters participated in the International Conference on 'Speech Motor Dynamics in Stuttering', Nijmegen, June 13-15 and in the 5th International Congress on Cleft Palate and Related Craniofacial Anomalies, Monte Carlo, September 2-7, where she gave a paper (with Kirsten Brøndsted): "Analysis of the spontaneous speech and communication behaviour in Danish preschool children with cleft palate and cleft lip and palate".

Birgit Hutters gave a guest lecture at the Institute of Linguistics, Stockholm University: "Vocal fold adjustments in unvoiced obstruents - with special reference to aspiration and devoicing", March 6.

Jørgen Rischel participated in the International Conference on Sino-Tibetan Languages and Linguistics, Bangkok, 27-29 August, and gave a plenary session paper: "Tonogenesis in Thai: A phonetic paradox".

Nina Thorsen participated in a conference on Event Perception, Uppsala, June 26.

Nina Thorsen gave a guest lecture at Université de Genève: "Models of intonation", July 2.

V. INSTRUMENTAL EQUIPMENT OF THE LABORATORY

The following is a list of instruments that have been purchased or built during the period January 1 - December 31, 1985.

TAPE RECORDERS

1 cassette recorder, AIWA, type AD-WX 220.

EQUIPMENT FOR EDP

2 CTR terminal, WYSE, type 85.

MICROPHONES

1 microphone, Sennheiser, type ME 40
2 microphones, Sennheiser, type ME 20
2 microphones, Sennheiser, type MKE 2-3.