F 1

ARIPUC VOL. 5.

UDVALG FOR FOLKEMÅL'S
HÅNDBIBLIOTEK
-2. JUL. 1971

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1971

ARIPUC VOL. 5.

F 1

UDVALG FOR FOLKEMÅLS
HÅNDBIBLIOTEK

-2. JUL. 1971

ANNUAL REPORT

of the
Institute of Phonetics
University of Copenhagen

COPENHAGEN 1971

KBHVNS. UNIV. OFFSET AFD. 4832/71
ISBM NR. 87-505-0158-5.

CONTENTS

(continued)

PERSONNEL OF THE INSTITUTE OF PHONETICS

197o

Permanent Staff:

Eli Fischer-Jørgensen (professor, director of the Institute)
Jørgen Rischel (lecturer and amanuensis of general phonetics)
Hans Peter Jørgensen (amanuensis of general phonetics)
Oluf M. Thorsen (amanuensis of general and French phonetics)
Børge Frøkjær-Jensen (amanuensis of experimental phonetics)
Hans Basbøll (amanuensis of general phonetics)
Ole Kongsdal Jensen (amanuensis of general and French phonetics)
Karen Landschultz (amanuensis of general and French phonetics)
Carl Ludvigsen (engineer)
Poul Thorvaldsen (engineer)
Svend-Erik Lystlund (technician)
Inger Østergaard (secretary)
Inge Lindemark (secretary until 15/9 197o)
Kirsten Kirkebæk (secretary from 15/9 197o)

Part Time Teachers of General Phonetics:

Kirsten Gregersen (teaching assistant)
Steffen Heger (teaching assistant)
Mogens Baumann Larsen (amanuensis of Danish phonetics)
Nina G. Thorsen (teaching assistant)

Guest Research Workers:

Hideo Mase (Japan)

PUBLICATIONS BY STAFF MEMBERS 197o

Eli Fischer-Jørgensen, Introduction to H. J. Uldall, Outline of Glossematics I, 2nd ed. 1967 (197o), p. I-XXII.

Eli Fischer-Jørgensen, "Phonetic Analysis of Murmured (Breathy) Vowels in Gujarati", Indian Linguistics XXVIII (197o), p. 71-139.

Eli Fischer-Jørgensen, "Les occlusives françaises et danoises d'un sujet bilingue", Word 24, 1968 (197o), p. 112-153.

Jørgen Rischel, "Acoustic features of syllabicity in Danish", Proc. Sixth Intern. Congr. of Phon. Sc., Prag 1967 (197o), p. 767-77o.

Jørgen Rischel, "Consonant Gradation: A Problem in Danish Phonology and Morphology" in: Hreinn Benediktsson (ed.), The Nordic Languages and General Linguistics (Reykjavík 197o), p. 46o-48o.

LECTURES AND COURSES IN 197o

1. Elementary phonetics courses

One-semester courses (two hours a week) in elementary phonetics (intended for all students of foreign languages except French) were given by Hans Basbøll, Kirsten Gregersen, Steffen Heger, Hans Peter Jørgensen, Mogens Baumann Larsen, and Nina G. Thorsen. There were 2 parallel classes in the spring semester and 13 in the autumn semester (about 35o students in all).

Two-semester courses (two hours a week) in general and French phonetics (intended for all students of French) were given through 197o by Ole Kongsdal Jensen, Karen Landschultz, and Oluf M. Thorsen. There were 9 parallel classes in the spring semester and 7 in the autumn semester (on an average 175 per semester).

2. Practical training in sound perception and transcription

Courses for beginners as well as courses for more advanced students were given through 197o by Hans Peter Jørgensen, Jørgen Rischel, and Oluf M. Thorsen. (The courses form a cycle of three semesters with two hours a week.)

3. Instrumental phonetics

Courses for beginners as well as courses for more advanced students were given through 197o by Børge Frøkjær-Jensen. (The courses form a cycle of three semesters with two hours a week.)

4. Phonology

Courses for beginners as well as courses for more advanced students were given through 197o by Hans Basbøll and Jørgen Rischel. (The courses form a cycle of three semesters with two hours a week, the subject of the three semesters being elementary phonological method, trends in phonological theory, and generative phonology.)

5. Other courses

Carl Ludvigsen and Eli Fischer-Jørgensen lectured on psychological acoustics and auditory phonetics through the autumn semester.

Carl Ludvigsen gave a course in mathematics and electronics in the autumn semester.

Hans Basbøll gave a course in Danish phonology in the spring semester.

Hans Peter Jørgensen and Niels Davidsen-Nielsen gave courses in the phonetics of German and English, respectively, in the spring semester.

Oluf M. Thorsen gave a course in French phonetics in the autumn semester.

Moreover, the students followed Henning Spang-Hanssen's course in statistics for linguistic purposes through the spring semester.

6. Seminars

The following seminars were held in 1970:

Professor Eli Fischer-Jørgensen and amanuensis Hans Peter Jørgensen: Loose and close contact.

Professor I. H. Slis (Eindhoven): Durational and electromyographic measurements on force of articulation.

Dr. S. O. Siemssen and speech therapist A. Boberg presented a film showing tongue movements during speech articulation and swallowing, and a film showing speech rehabilitation after total paresis of the tongue.

Professor G. Ungeheuer (Bonn): Über den Begriff der phonetischen Information.

Nina G. Thorsen presented her study on voicing in English (see below p. 1ff).

Børge Frøkjær-Jensen demonstrated the glottographs of the Institute.

7. Participation in congresses and lectures at other institutions by members of the staff

Eli Fischer-Jørgensen and Hans P. Jørgensen lectured on "Loose and close contact" at the Laboratory of Speech Transmission of the Royal Technical High School in Stockholm.

Børge Frøkjær-Jensen visited the institutes of phonetics in Amsterdam, IPO Eindhoven, Bruxelles, Bonn, Köln, Hamburg, Essex, Leeds, Colchester, and Praha, Die technische Hochschule, Dresden, University College, London, and the Institute of African and Oriental Studies, London. The expenses were covered by a grant from the University of Copenhagen.

Børge Frøkjær-Jensen participated in the "Symposium on Intonology" in Praha.

Børge Frøkjær-Jensen lectured at a meeting held in Copenhagen by the Danish Society of Logopedics and Phoniatrics on "Comparative Phonetic-Acoustic Analysis Before and After Speech Therapy of Voices Suffering from Recurrens Paresis" (cf. ARIPUC 4/1969 (1970), p. 145ff).

Hans Basbøll and Jørgen Rischel participated in the Second Scandinavian Summer School of Linguistics in Stockholm.

INSTRUMENTAL EQUIPMENT OF THE LABORATORY

The following is a list of the instruments that have been purchased or built since January 1st, 197o.

1. Instrumentation for speech analysis

1 segmentator, type PT
1 electro-aerometer, type AM 5o8/4
1 audio frequency filter, type 445

2. Tape recorders

4 semi-professional recorders, Revox (stereo, speeds 3 3/4" and 7 1/2")

3. Loudspeakers

3 loudspeakers, Beovox, type 26oo

4. General-purpose electronic instrumentation

1 multi-generator, Exact, type 126 VCF
2 stabilized rectifiers, Danica, type TPS 1d
1 stabilized rectifier, Danica, type TPS 3c

5. Projectors

1 projector, Leitz, type Pradovit color 25o
2 overhead projectors, 3M company

REPORTS ON CONSTRUCTIONAL WORK

Jørgen Rischel and Poul Thorvaldsen

1. Improvements to the segmentator

In the past year the segmentator has been improved. A new gate with field-effect transistors working in the non-pinchoff region (variable impedance) has been constructed to replace the old one with a diode-bridge.

Instead of clipping the speech signal when using envelope shaping the new gate damps the signal. The damping function is approximately a Gaussian function for a trapezoidal input to the gate. At the same time the maximum damping is raised from 45 dB to 7o dB. (PT)

2. Construction of a four-channel electromyograph

A four-channel electromyograph suitable for use with needle or surface electrodes has recently been constructed. The signal is led to a high input-impedance pre-amplifier with a gain of 2o dB. The input mode can be either DC or AC (o.1 Hz). The pre-amplifier is followed by a general-purpose amplifier with gain adjustable from -6o dB to +4o dB. Thereafter the system allows for signal processing via high-pass and low-pass filtering, rectification and integration. Rectification is linear, and filter cut-off frequencies and integration time-constants are selectable. The bandwidth of the system is o - 2o kHz.

Further specifications relating to noise levels and CMRR are not given here, since we are at present conducting experiments to determine the effect of these values of coupling electrodes into the system. These experiments will be reported later. (PT)

3. Speech synthesizer

The synthesizer of the Institute of Phonetics has been modified so that it now includes 8 formant circuits for the synthesis of vowels and consonants. The control system (function generator) has been constructed according to the principles outlined in ARIPUC 3/1968 (1969), p. 17 ff. It is presently possible to control 12 varying parameters simultaneously. Specimens of speech have been synthesized with fairly good quality. (JR)

VOICING IN BRITISH ENGLISH t AND d IN CONTACT WITH OTHER CONSONANTS[1]

Nina G. Thorsen

1. Introduction

Surprisingly little has been published about voicing in British English (BE) plosives. Gimson (1966 p. 145-162 and 266-27o) does say something on this subject, but his treatment is incomplete. - For instance, nothing is said about the influence of stress distribution upon voicing in plosives.

A great deal of research has been done on American English stop consonants, but mostly when these stand initially before stressed vowels.

A study of voice assimilation in French obstruents has been done by Oluf M. Thorsen (1967). Eli Fischer-Jørgensen has investigated assimilation in Dutch (1952).

2. The material

The present investigation is based on an acoustic (mingographic) analysis of 76 sentences, recorded five times by six subjects, three males and three females, all speaking RP. In addition to this material, which consisted of examples of t and d, a much smaller material containing examples with b and g was recorded twice by two subjects

---

1) Thesis work for the cand.mag.-degree, completed in June 197o. I am grateful to Eli Fischer-Jørgensen for much good advice and very helpful criticism.

(two of the previously used males).

2.1. The main corpus of the material

t and d were investigated in the following two positions: (a) X+_V where + is a morpheme- or word boundary, and X is a pause, a vowel, or a vowel followed by p, b, s, z, or l, and (b) V_+X where + is a morpheme- or word boundary, and X is a pause, a vowel, or a vowel preceded by k, g, s, z, or l.

2.2. Additional material

In addition, examples of t and d preceded by k and g respectively, and of t and d followed by p and b respectively, were recorded twice by the aforementioned two male subjects.

b and g were investigated in the following two positions: (a) X+_V where + is a morpheme- or word boundary, and X is a pause, a vowel, or a vowel followed by t, d, s, z, or l, and (b) V_+X where + is a morpheme- or word boundary, and X is a pause or a vowel.

The examples with vowels and pause before/after the consonant under investigation were included for comparison.

2.3. The clusters under investigation

Thus the material contains the following 'clusters'

a) V+t, p+t, b+t, s+t, z+t, l+t,
t+V, t+k, t+g, t+s, t+z, t+l,
V+d, p+d, b+d, s+d, z+d, l+d,
d+V, d+k, d+g, d+s, d+z, d+l

There were 3o examples of each of these clusters (in each stress distribution, see below), in some cases slightly fewer, due to mistakes in the recordings.

b) k+t, g+t, k+d, g+d, t+b, d+p, d+b,
V+b, t+b, d+b, s+b, z+b, l+b,
V+g, t+g, d+g, s+g, z+g, l+g,
b+V, g+V

There were four examples of each of these clusters (in each stress distribution, see below), in some cases fewer, due to mistakes in the recordings.

2.4. Stress distributions

The clusters (and the examples with pause before/after the consonant to be examined) in the more comprehensive material were all investigated under two stress conditions:

first syllable stressed, second syllable unstressed (strong-weak)
first syllable unstressed, second syllable stressed (weak-strong)
and some of them furthermore under a third type of stress condition:
first syllable unstressed, second syllable unstressed (weak-weak).

The clusters in the more limited material were not so extensively examined, there being, for instance, no weak-weak examples, and there were also omissions among the strong-weak and the weak-strong distributions.

A few illustrations:

b+t, weak-strong: "Subtopians are happy"
strong-weak: "There's a cab to pay"
weak-weak: "It's a subterranean spring"
d+g, strong-weak: "The cad got hurt"
weak-strong: "Did Golding do it?"
weak-weak: "Did Golightly do it?"

3. The recordings and tracings

The recordings were made in the studio at the Institute of Phonetics with a condenser microphone, and a Lyrec Professional Recorder, 1/1 track, at 7 1/2"/sec, on Scotch Magnetic Tape, SP 1o2.

The recordings were traced on a mingograph, Elema 8oo. The traces were: one duplex oscillogram, one F_o curve (Trans Pitchmeter, Børge Frøkjær-Jensen), one logarithmic unfiltered intensity curve, one linear intensity curve, high-pass filtered at 5oo Hz (Intensity Meter, Børge Frøkjær-Jensen) and an oscillogram filtered so that essentially only the first harmonic was traced (Audio-Frequency Filter, Børge Frøkjær-Jensen).

The calculations on the data thus obtained were performed on a data processing machine.[2]

4. Results

4.1. General remarks

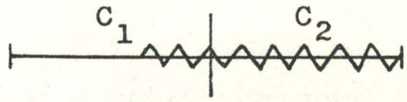
I was interested only in the duration of the voicing, whether audible or not. Audible voicing was of less interest in this study. I wanted to study the phenomenon of assimilation, the physiological mechanism itself, and the possible beginnings of a diachronic development which may be discerned.

It has previously been established that Grammont's description of the mechanism of regressive assimilation is wrong (Fischer-Jørgensen 1952 and Thorsen 1967), and it does not seem to be valid for BE either.

---

2) I am grateful to Børge Frøkjær-Jensen who was in charge of this part of the investigation.

Grammont (1939 p. 193f) maintained that a partial regressive assimilation always affected the last part of the preceding consonant. He may have imagined a situation like this:



which has not been found once in my material, nor in the Dutch and French material of Eli Fischer-Jørgensen and Oluf M. Thorsen, respectively.

There is normally some voicing to be found in a consonant immediately following a vowel - the vocal chord vibrations continue into the consonant until the pressure drop across the glottis is too small to allow vibrations. This voicing may be supported and prolonged by a following voiced consonant and the voicing, if not total, will typically be found to extend continuously through the first part of the cluster, which may end unvoiced, even when the first consonant is basically unvoiced (fortis) and the second consonant is basically voiced (lenis).[3]

4.2. Stylized display of voicing in consonant clusters

A stylized version of the typical situations found in my material (1475 clusters) is given on the following pages. - The clusters are given to the left. They are divided into groups according to the basic voicing characteristic of the first and second consonants. In Table Ia all clusters are included, and the number of examples

---

3) I shall use the terms fortis and lenis throughout about basically unvoiced and basically voiced consonants respectively - whether these be appropriate phonetic terms or not.

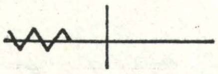
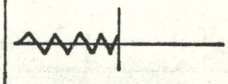
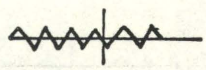
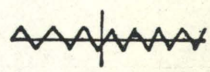
of each cluster is given. The figures in parentheses represent the total number of examples in each group (1a 242 examples, 2a 23 examples, etc., etc.).

Table Ia Voicing in consonant clusters

				1a		2a		3a	
p+t	1' C fortis	2' C fortis	47						
t+k			66		12		6		
k+t			4	(242)		(23)		(6)	
s+t			57						
t+s			68		11				
				1b		2b		3b	
d+p	1' C lenis				2		6		
b+t			38		18		33		
g+t			2				2		
d+k			29	(191)	23	(77)	33	(153)	
d+s			37		31		18		
z+t			85		3		1		
l+t							60		
			433		100		159		

continued on the
next page

Table Ia continued

									
p+d	1' C fortis	46	4a		5a		6a		7a
t+b		7				3		4	
k+d		4							
t+g		51		11		11		lo	
s+b		2	(197)		(13)		(15)		(24)
s+d		52				1			
s+g		5							
t+z	2' C lenis	29		2				6	
t+l		1						4	
d+b			4b		5b	5	6b	7	7b
b+d		8		4		32		43	
g+d		1				3			
d+g	1' C lenis	23		5		37		22	
z+b			(57)		(32)	5	(18o)	3	(265)
z+d		15		16		44		11	
z+g		1				4		3	
d+z		9		7		25		45	
d+l								81	
l+b						2		6	
l+d						21		38	
l+g						2		6	
		254		45		195		289	

In table Ib the number of examples of a particular cluster in each situation is given in percentages of the total number of examples of that cluster (included the atypical ones, see below), but the clusters from the smaller material have been omitted, since percentages for

Table Ib Voicing in consonant clusters (percentages)

2' C	1' C	Cluster								
				1a		2a		3a		
2' C fortis	1' C fortis	p+t	loo							
		t+k	79		14		7			
		s+t	loo	(89)		(9)		(2)		
		t+s	86		14					
				1b		2b		3b		
	1' C lenis	b+t	43		2o		37			
		d+k	34		27		39			
		d+s	43	(45)	36	(18)	21	(37)		
		z+t	96		3		1			
		l+t					loo			
2' C lenis	1' C fortis	p+d	96	4a		5a		6a		7a
		t+g	61		13		13		12	
		s+d	91	(59)		(4)	2	(4)		(7)
		t+z	51		3				lo	
		t+l	2						8	
				4b		5b		6b		7b
	1' C lenis	b+d	9		5		37		49	
		d+g	26		6		43		25	
		z+d	17	(lo)	18	(6)	49	(33)	12	(48)
		d+z	lo		8		29		52	
		d+l							9o	
		l+d					36		64	

such low figures are misleading. The figures in parentheses indicate the number of examples in each situation of a particular type of cluster in percentages of the total number of examples of that type of cluster, included those from the small material, since this makes very little difference, or none at all. There are four types: fortis + fortis (F+F), 271 examples, lenis + fortis (L+F), 421 examples, fortis + lenis (F+L), 249 examples, and lenis + lenis (L+L), 534 examples. Thus "(89)" in group 1a means that the examples in this group comprise 89% of the total number of F+F examples, etc., etc.

Note that in both tables clusters in different stress distributions have been pooled.

4.2.1. Clusters exhibiting unbroken partial or total voicing

The 1475 clusters in tables Ia and Ib follow the rule that the voicing in consonant clusters, if not total, lies unbroken in the beginning of the cluster:

It will be seen that an initial[4] fortis consonant (groups 1 - 3) is usually unvoiced (533 examples or 98% of the F+F and 63% of the L+F clusters), but it may be voiced in the beginning, if the preceding (final) consonant is fully voiced (159 examples or 2% of the F+F and 37% of the L+F clusters). Final fortis consonants (groups 1a - 7a) are usually voiced only in the beginning (439 examples or 89% of the F+F and 59% of the F+L clusters), but t and s may be fully voiced (81 examples or 11% of the F+F and 15% of the F+L clusters), though only when they are unstressed. Initial lenis consonants (groups 4 - 7)

---

4) Initial means initially in a word or morpheme, not initially in the cluster, and final means finally in a word or morpheme, not finally in the cluster.

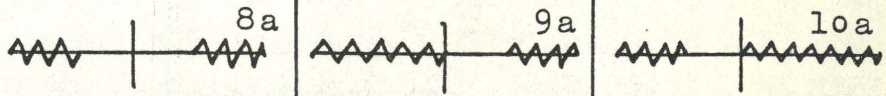
are unvoiced (299 examples or 63% of the F+L and 16% of the L+L clusters) or voiced in the beginning (195 examples or 4% of the F+L and 33% of the L+L clusters) like initial fortis consonants, but initial lenis consonants may be fully voiced, usually after a final lenis consonant (group 7b, 265 examples or 48% of the L+L clusters), but also sometimes after final fortis consonants when these are unstressed (group 7a, 24 examples or 7% of the F+L clusters). Final lenis consonants (groups 1b - 7b) are normally voiced throughout (7o7 examples or 87% of the L+L and 55% of the L+F clusters), but they may be voiced only in the beginning (248 examples or 45% of the L+F and 1o% of the L+L clusters).

4.2.2. Clusters exhibiting discontinuous voicing

A total of 98 clusters showed discontinuous voicing. They are displayed in tables IIa and IIb. In table IIa all clusters are included, in table IIb those from the small material have been omitted and the number of examples are given in percentages like in table Ib. In these tables the different stress distributions are considered. The figures in parentheses in table IIb indicate the number of examples of each type of cluster in each situation in percentages of the total number of the same type of cluster, for instance, 22% of the fortis + lenis clusters (group 8a) were found to have discontinuous voicing in the manner indicated at the top of the column.

Note that all of these 98 clusters have a lenis second consonant, and almost all (85 examples) have a fortis first consonant. 59 examples have been found in the weak-strong distribution.

Table IIa Voicing in consonant clusters



	8a w-st	8a st-w	8a w-w	9a w-st	9a st-w	9a w-w	loa w-st	loa st-w	loa w-w
1' C fortis, 2' C lenis									
t+b		2							
t+g	1								
p+d	2								
t+z	9	5		4			2		
t+l	25	21		1					
s+d	4								
s+b	1	3					2		
s+g	1						2		
	43 43	31 31		5 5			6 6		
1' C lenis, 2' C lenis			8b			9b			lob
z+d				3					
d+z	1								
d+l		3	4	1		1			
	1 1	3 3	4 4	4 4		1 1			
total:	44	34	4	9		1	6		

Table IIb
percentages

	8a w-st	8a st-w	8a w-w	9a w-st	9a st-w	9a w-w	loa w-st	loa st-w	loa w-w
			8a (22)			9a (2)			loa (2)
1' C fortis, 2' C lenis									
t+g	1								
p+d	4								
t+z	16	9		7			4		
t+l	48	4o		2					
s+d	-7								
1' C lenis, 2' C lenis			8b (2)			9b (1)			lob
z+d				4					
d+z	1								
d+l		3	5	1		1			

4.2.2.1. Tentative explanations

Various explanations offer themselves to the 85 examples with fortis first consonant. - The vibrations are discontinued in the first consonant (a) because the glottis opens and the vocal chords are too far apart for the Bernoulli-effect to be active or (b), if the glottis remains 'closed', because an intra-oral pressure is built up gradually and the pressure drop across the glottis is thus decreased. In the first case the air flows freely into the mouth cavity and the intra-oral pressure builds up rapidly. In the second case, air escapes into the mouth at each opening of the glottis during the vibratory cycle, and it will take longer for the pressure drop to decrease, but eventually it will be small enough to impede vibrations. - In any case vibrations can only resume if the intra-oral pressure is lowered, either by an active expansion of the mouth walls or by letting air out of the mouth - which is probably the explanation for the 67 examples (9+5+4+2 + 25+21+1) of t+z and t+l. In the 5 examples (2 + 1 + 2) of t+b, t+g, and p+d it may be that the mouth walls have been actively expanded, though this phenomenon is not generally assumed to take place in BE.

The 13 examples (4 + 1+3+2 + 1+2) of s+d, s+b, and s+g are rather hard to explain: it is difficult to see how vibrations can start after the buccal closure has been made and the intra-oral pressure thus has increased. But 10 of these examples have been found where the first syllable was unstressed and the second syllable stressed. It may be that a new and forceful innervation of the expiration muscles would increase the subglottal pressure sufficiently in relation to the supraglottal pressure for an air-stream to move through the glottis and set the vocal chords in motion, provided they are sufficiently close for the

Bernoulli-effect to be active. This same requirement must be met in the 5 examples of t+b, t+g, and p+d, and indeed it seems very likely that the vocal chords approach the 'closed' position at the beginning of the lenis consonant. In both cases some expansion of the mouth walls is probably called for.

Regarding the 13 clusters with lenis first consonant: 1o of these have z or l as second consonant and may be explained in the same way as the 67 examples mentioned above, except that it is fairly certain that the glottis remains 'closed' throughout the cluster.

Really strange are only 3 examples with z+d. z is fully voiced, and d is partially voiced - but in the end. I cannot explain why. Presumably the vibrations stop because of heightened intra-oral pressure. This coincides with the buccal closure, which makes for an even higher intra-oral pressure; however, vibrations start again before the explosion of the closure. This can only be explained by assuming an expansion of the mouth walls - but then: why does not this expansion happen earlier so as to produce continuous rather than interrupted voicing?

4.3. 'Rules' for the placement of voicing in consonant clusters

The above is, of course, purely speculative. In order to be certain about these phenomena glottograms should have been made. But whether the explanations are valid or not, all examples except those in groups 9 and 1o (i.e. 1557 examples in all) can be formulated in a rule:

Fortis consonants after consonants that end unvoiced are always unvoiced (433 examples, group 1).

Fortis consonants after consonants that end voiced are unvoiced (1oo examples, group 2) or voiced in the beginning (159 examples, group 3).

Lenis consonants after consonants ending unvoiced are unvoiced (254 examples, group 4) or they may be partially voiced in the end (82 examples, group 8).
Lenis consonants after consonants ending voiced may be unvoiced (45 examples, group 5), partially voiced in the beginning (195 examples, group 6), or fully voiced (289 examples, group 7).

4.4. Measurements (averages)

The next section contains some actual data on the degree of voicing found in the consonants investigated. The data represent averages of all subjects.

"Duration" at the top of columns in the tables always means "duration in msec of the consonant" (closure only, in plosives), "voicing" means "duration in msec of the voicing in the consonant" (closure, respectively), and "%" means "relative voicing in percentages". "+" denotes morpheme- or word boundary. The first of the four columns in each group gives the number of examples (in parentheses) that form the basis of the calculations. - The statistical significance of the differences in relative voicing in the consonants before/after a fortis and the corresponding lenis consonant has been calculated. "**" indicates statistical significance at the 99% level, "*" indicates statistical significance at the 95% level.

4.4.1. Final d

The d-closure is usually voiced finally, regardless of the following consonant (see Table III). The slight differences found in different stress distributions are not statistically significant.

Table III Voicing in the closure of final d:

stress-distri-bution	strong-weak				weak-strong				weak-weak			
		dura-tion	voic-ing	%		dura-tion	voic-ing	%		dura-tion	voic-ing	%
d+pause	(3o)	67	61	91	(3o)	79	61	77				
d+V	(3o)	32	32	loo	(19)	36	36	loo	(25)	37	36	98
d+k	(26)	64	62	97	(3o)	51	44	86	(3o)	56	5o	88
d+g	(28)	62	6o	97	(3o)	56	51	91	(3o)	53	47	87
d+s	(26)	63	59	92	(3o)	63	5o	79	(3o)	55	46	83
d+z	(3o)	71	71	loo	(3o)	64	62	96*	(29)	61	56	92
d+l	(3o)	75	74	98	(3o)	56	56	loo	(3o)	49	46	94

Table IV shows that there is a considerable difference in the degree of voicing found in k vs. g and in s vs. z after d: the lenis consonant is more voiced than the fortis counterpart. - Only in the case of z does the voicing show statistically significant dependance upon stress distribution: z is more voiced when it is stressed than when it is unstressed.

Table IV Voicing in the consonants following final d:

stress-distribution	strong-weak				weak-strong				weak-weak			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
d+k	(26)	86	17	20	(30)	79	13	17	(30)	70	9	13
d+g	(28)	71	30	42**	(30)	80	40	49**	(30)	63	23	36**
d+s	(26)	129	12	9	(30)	143	14	10	(30)	128	6	5
d+z	(30)	81	51	63**	(30)	118	99	84**	(29)	88	32	37**
d+l	(30)	62	60	97	(30)	88	87	97	(30)	67	61	91

4.4.2. Final t

The difference between the relative voicing in final and intervocalic t is statistically significant, but the difference in absolut voicing is small or non-existent (see table V). The relative differences appear, of course, because the intervocalic closure is comparatively shorter than the final closure. Otherwise, it seems that final t is slightly more voiced before a lenis consonant than before its fortis counterpart (though not before k vs. g when t is stressed). Unstressed t is slightly more voiced than stressed t.

Table VI shows that fortis consonants after t are completely unvoiced. g and z are unvoiced after stressed t but slightly voiced after unstressed t. l is partially voiced (in the end), more so when l is stressed.

Table V Voicing in the closure of *final t*:

stress-distribution	*strong*-weak				*weak*-strong				*weak*-weak			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
t+pause	(3o)	8o	22	27	(25)	63	26	42				
t+V	(22)	53	22	42*	(19)	25	22	89**	(3o)	29	23	79
t+k	(25)	77	18	23	(3o)	47	31	66	(3o)	41	31	75
t+g	(26)	68	17	25	(3o)	47	41	87	(3o)	47	34	73
t+s	(2o)	71	26	36	(29)	55	33	6o	(3o)	49	33	67
t+z	(27)	1oo	4o	4o	(3o)	78	64	82				
t+l	(22)	1o2	32	31	(29)	67	35	53				

Table VI Voicing in the consonants following final t:

stress-distribution	strong-*weak*				weak-*strong*				weak-*weak*			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
t+k	(25)	1o5	o	o	(3o)	8o	o	o	(3o)	69	o	o
t+g	(26)	91	o	o	(3o)	77	22	28**	(3o)	67	12	18*
t+s	(2o)	14o	o	o	(29)	14o	o	o	(3o)	112	o	o
t+z	(27)	11o	6	6	(3o)	137	74	54**				
t+l	(22)	77	42	54	(29)	85	61	71				

4.4.3. Initial d

Initial d is strongly influenced by the preceding consonants (see Table VII). Between vowels d is voiced (though not completely (91-92%) when unstressed). The two intervocalic d's do not exhibit marked differences. After fortis consonants d is almost completely devoiced. After lenis consonants d is partially voiced, more so when d is stressed.

Table VII Voicing in the closure of initial d:

stress-distribution	weak-strong				strong-weak				weak-weak			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
pause+d	(8)[5]		47[6]		(5)[5]		38[6]					
V+d	(30)	84	76	91	(30)	52	52	100	(30)	60	57	90
Vd[7]	(30)	83	77	92	(30)	44	44	100				
p+d	(29)	67	5	7	(19)	67	0	0				
b+d	(30)	60	50	83**	(27)	63	28	44**	(30)	32	17	53
s+d	(30)	84	9	11	(27)	57	0	0				
z+d	(30)	107	48	45**	(29)	73	12	17**	(30)	62	20	32
l+d	(29)	113	88	78	(30)	51	51	99				

---

5) There were 30/30 examples of stressed and unstressed initial d after pause, but only 8/5 that showed 'voicing lead'.

6) The closure duration could not be measured in initial d and t.

7) This example is one where there is no boundary between d and the surrounding vowels, included for comparison with the example just above.

Table VIII shows that lenis consonants before initial d are almost always fully voiced, independently of stress distribution (cf. Table III about final d). p and s are partially voiced in the beginning, more so when p and s are unstressed.

Table VIII Voicing in the consonants preceding initial d:

stress-distri-bution	weak-strong			strong-weak			weak-weak		
	duration	voicing	%	duration	voicing	%	duration	voicing	%
p+d	(29)100	27	27	(19)101	25	25			
b+d	(30) 88	88	100**	(27) 89	86	97**	(30) 79	77	97
s+d	(30) 84	36	42	(27)107	31	29			
z+d	(30) 70	69	99**	(29)101	95	94**	(30) 62	61	98
l+d	(29) 68	68	100	(30) 92	92	100			

Initial t after a pause is always unvoiced (see Table IX). Between vowels t is partially voiced, relatively more so when t is unstressed than when t is stressed. When initial t is unstressed after a stressed open syllable it is less voiced when no boundary precedes the t, but note that the differences are reversed (and statistically significant) when t is stressed. t is (almost) unvoiced after p, b, s, and z (z being partially devoiced before t, cf. Table X). t is partially voiced after l as it is between vowels.

Table IX Voicing in the closure of <u>initial t</u>:

stress-distribution	weak-<u>strong</u>				strong-<u>weak</u>				weak-<u>weak</u>			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
pause+t	(3o)		o		(3o)		o					
V+t	(3o)	98	33	34	(3o)	58	32	55	(29)	65	28	42
Vt[8]	(3o)	93	38	41**	(3o)	43	21	48				
p+t	(27)	58	o	o	(2o)	76	o	o				
b+t	(3o)	61	2	4	(29)	62	7	11**	(3o)	42	3	6
s+t	(3o)	53	o	o	(27)	62	o	o				
z+t	(3o)	96	o	o	(3o)	58	o	o	(29)	61	2	3
l+t	(3o)	loo	33	33	(3o)	51	32	63				

Fortis consonants before <u>t</u> are partially voiced in the beginning, more so when they are unstressed (see Table X). <u>b</u> is almost fully voiced. <u>z</u> has been partially devoiced, more so when <u>z</u> is unstressed. <u>l</u> is fully voiced. The difference between the relative voicing in <u>s</u> and <u>z</u> in the <u>weak</u>-strong distribution is significant, but this is hardly the case for the absolute values.

Table X Voicing in the consonants preceding initial t:

stress-distribution	<u>weak</u>-strong				<u>strong</u>-weak				<u>weak</u>-weak			
		duration	voicing	%		duration	voicing	%		duration	voicing	%
p+t	(27)	lo2	3o	29	(2o)	118	19	16				
b+t	(3o)	85	72	85**	(29)	95	85	89**	(3o)	85	78	9o
s+t	(3o)	113	35	31	(27)	121	28	23				
z+t	(3o)	77	39	5o**	(3o)	97	48	49**	(29)	67	43	64
l+t	(3o)	65	65	loo	(3o)	82	82	loo				

8) This example is one where there is no boundary between <u>d</u> and the surrounding vowels, included for comparison with the example just above.

The results in Tables III, V, VII, and IX concerning final and initial d and t, respectively, are displayed graphically in Figs. 1 and 2.

The figures do not give any further information, but they facilitate a direct comparison of d and t, and they help illustrate the concluding remarks to this section (but note that the figures only concern d and t).

4.4.5. Conclusions inferred from the measurements

I should mention first that a vowel is always longer before a lenis consonant than before the corresponding fortis consonant, no matter what the degree of voicing in the consonants may be.

Of the consonants it might be said generally that they preserve their basic voicing characteristic better when stressed than when unstressed, cf. that the 'empty-square' curve is usually the uppermost and the 'filled-square' curve usually the lowest one in the figures.

4.4.5.1. Single d and t

Before a pause after a vowel a d-closure is approximately fully voiced, relatively, but not absolutely, more so when stressed. The t-closure is partially voiced in the beginning, more so when unstressed than when stressed.

After a pause before a vowel the d-closure is usually unvoiced (22 examples of 30 in stressed initial d, 25 examples of 30 in unstressed initial d). It may be voiced and if so, the voicing-lead always has a certain duration, about 40 msec. The t-closure is always unvoiced after a pause.

Between vowels, d is always (almost) fully voiced. The t-closure is partially voiced, the voicing being to some degree dependent upon the placement of boundaries and stress (in the case of stress this is only valid for the relative voicing in final t, cf. Fig. 1b).

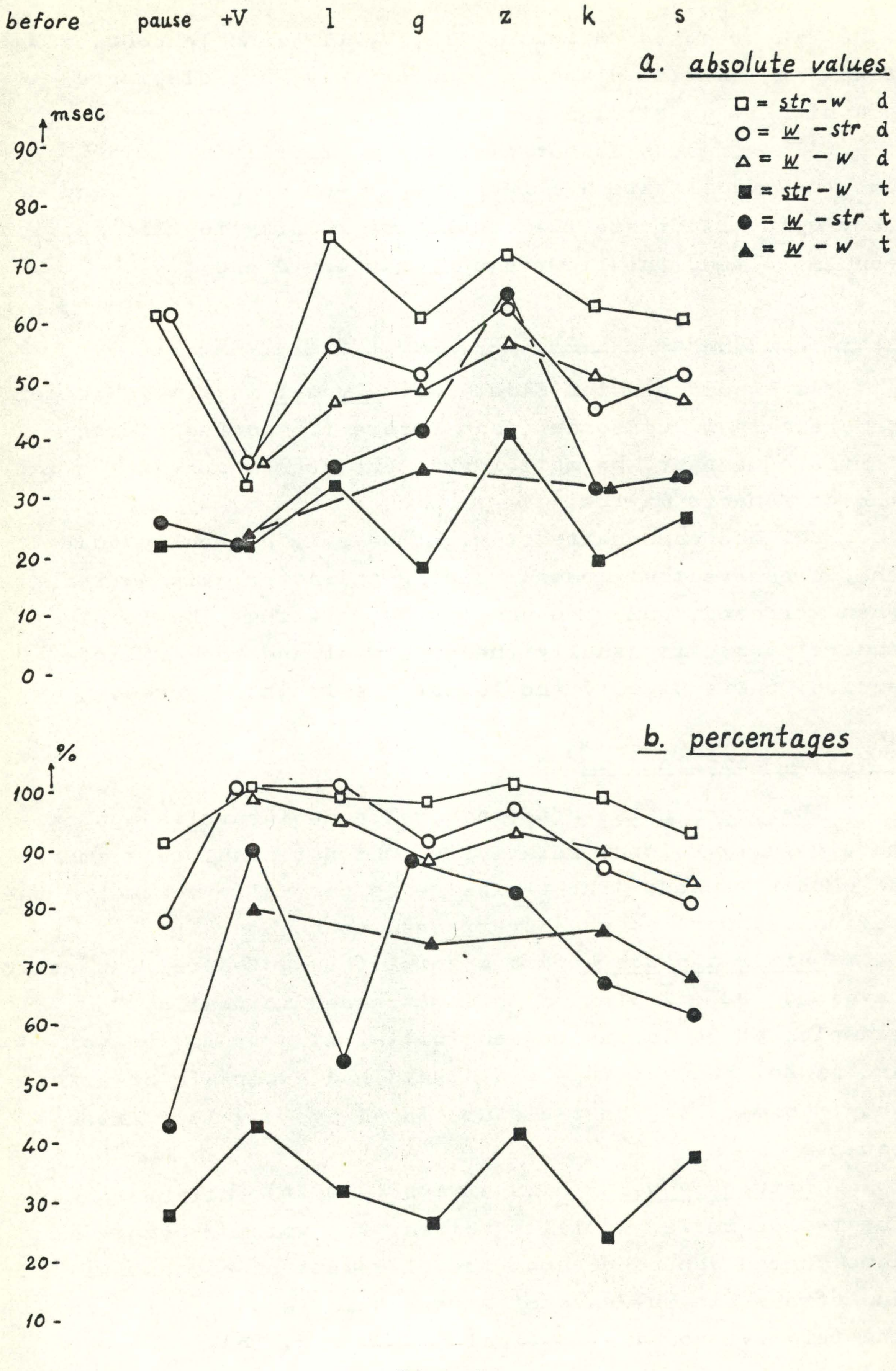


Fig. 1.

Voicing in the closure of final d and t.

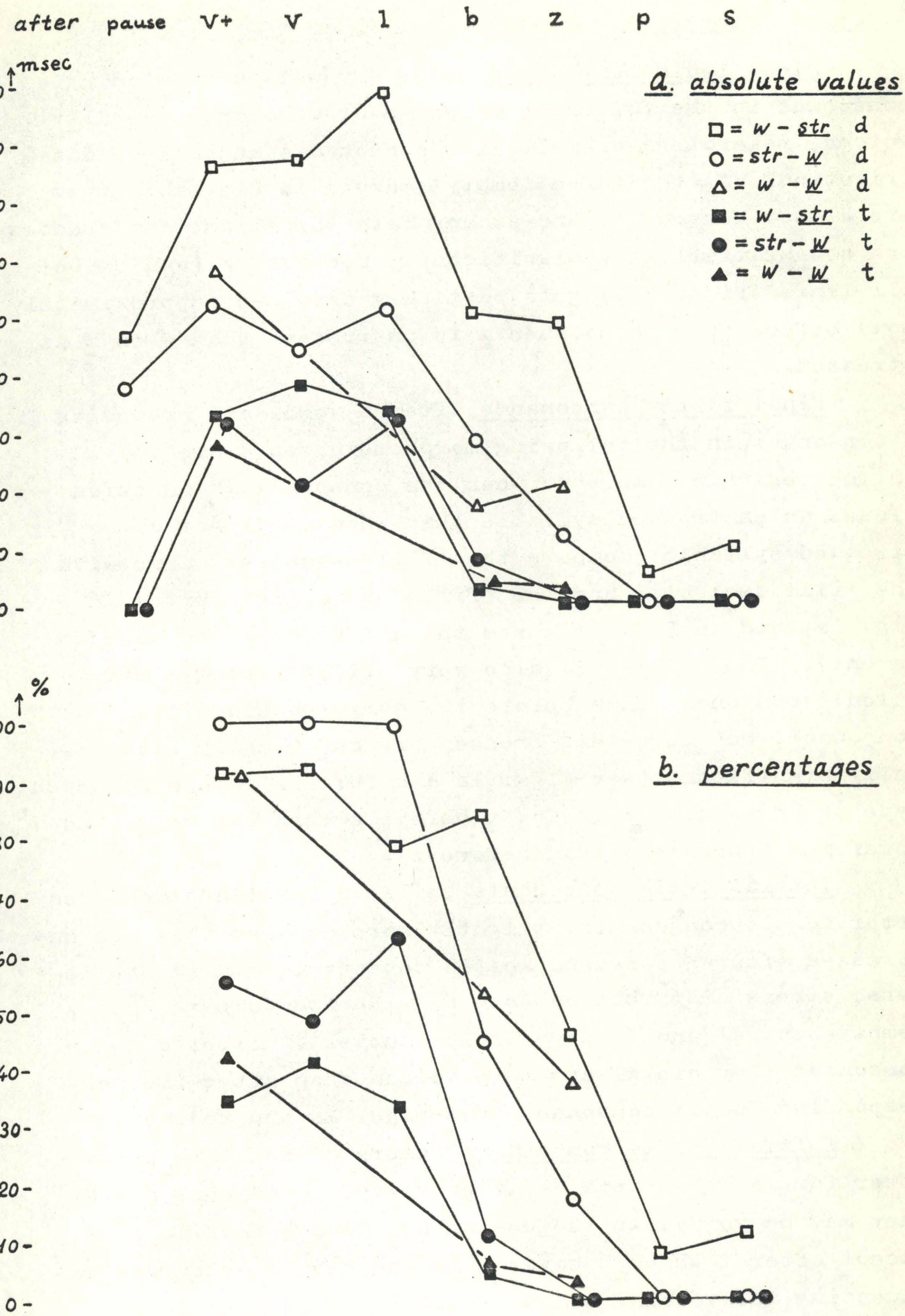


Fig. 2.

Voicing in the closure of initial d and t.

4.4.5.2. Clusters

Final lenis consonants after a vowel and preceding a consonant in the following morpheme or word are generally voiced, regardless of following consonants and stress distribution, cf. the three 'empty' curves in Fig. 1b: they are almost congruent, and at any rate the slight differences are not statistically significant. However, z (and maybe all lenis fricatives) gets partially devoiced (approximately 50%) before t, more so when z is unstressed than when z is stressed.

Final fortis consonants after a vowel and preceding a consonant in the following morpheme or word are voiced in the beginning, more so when the consonant in question closes an unstressed syllable than when it closes a stressed syllable (compare the 'filled-square' curve with the 'filled-circle' and the 'filled-triangle' curves in Fig. 1a, and in Fig. 1b where the differences are more apparent). The voicing is also very slightly longer before a lenis consonant than before the corresponding fortis consonant, but these differences are not usually statistically significant (see Figs. 1a and 1b; the values for each type of curve are lower for t before k than before g, and lower for t before s than before z).

Initial lenis consonants before a vowel are devoiced after fortis consonants, slightly more so when they are unstressed after a stressed fortis consonant than in the reverse stress distribution (see Fig. 2b and compare the 'empty-square' and 'empty-circle' curves). After a lenis consonant they are always more voiced than after the corresponding fortis consonant (see Figs. 2a and 2b).

Initial fortis consonants before a vowel are unvoiced after fortis consonants, - after voiced lenis consonants they may be voiced in the beginning, though this is rare, except after l which behaves like the vowels as far as influencing the voicing of the following consonant goes.

4.4.6. The additional material

The small material with b and g exhibited results that are in agreement with those obtained for d. It is interesting to note that b is relatively more voiced than d, and g is slightly less voiced than d. This is in accordance with what Eli Fischer-Jørgensen has found on many occasions (e.g. 1968). The fact that b is relatively more voiced than d and g supports the explanation given about the physiological conditions and mechanisms of voicing: vibrations will only occur when a sufficient air-stream passes the glottis, that is, as long as the supra-glottal pressure is not too high in relation to the subglottal pressure. As the cavities behind a labial closure is larger than behind a dental and velar closure it will take longer for the intra-oral pressure to build up to a level that prohibits voicing.

4.5. An assimilation rule

A tentative assimilation rule for clusters of two consonants, distributed over two neighbouring syllables, may be formulated:

> Voicing of fortis consonants, whether progressive or regressive, takes place only to a very small extent.
> Regressive devoicing affects only (z and presumably all) lenis fricatives, whereas progressive devoicing (more or less extensive) is frequent and affects plosives, fricatives, and l as well.
> Generally a consonant preserves its basic voicing characteristic better when stressed than when unstressed.

5. The effects of assimilation upon other properties of the consonants

Apart from voicing, Gimson (1966 p. 144-154) mentions the following features separating t and d: force of articulation, aspiration, length of preceding sounds, duration of the closure, and intensity of the burst. I have measured only three of these factors: aspiration, length of preceding vowel, and duration of the closure. Although I have not measured intra-oral pressure, which Gimson does not mention as a factor separating t and d, it is fairly certain that a difference in intra-oral pressure exists, and that it is assimilated completely with the voicing (cf. Fischer-Jørgensen 1969 and Thorsen 1967).

5.1. Duration of the preceding vowel

If duration of the vowel in -CV+ syllables ("+" denotes word- or morpheme boundary) were dependent upon the degree of voicing in the consonant one would expect the vowel to be longer the more voiced the following consonant is.

Devoicing of a final lenis consonant should shorten the preceding vowel, and voicing of a final fortis consonant should lengthen it.

In my material a vowel is always longer before syllable final d than before syllable final t, and in all cases, except two, the differences are statistically significant (at the 99% level (**) or at the 95% level (*)). - But it might be that the differences in vowel duration, although they be significant, are smaller in cases where the difference in voicing in d and t is small or eliminated, so that a partial assimilation can be said to have taken place.

5.1.1. Measurements (averages)

In Table XI are given data on the duration of vowels before d and t respectively and on the differences between the voicing (absolute and relative voicing, respectively) in d and t. - "C" stands for d or t, the lengthening marks apply only when C is d. Note that statistical significance has not been calculated for the difference in absolute voicing in d and t. About the number of examples: there were usually 3o or slightly fewer (interested readers are referred to Tables III and V for the exact figures).

5.1.1.1. The stressed vowels

[æ:] before d+pause is extra long, compared to [æ:] before d+k and d+g. The same tendency is observed in [æ] before t+pause, but not so strongly, which creates a large difference in vowel duration before d and t plus pause.

[u:] before d+z is also very long, but this is probably due to the choice of frame-word ('They booed Zacharia'). [ɛ] before d+l is rather short, I do not know why. The frameword was 'bled' in 'He bled like a pig'.

One more vowel difference is rather small, namely that between [ɛɪ] before d+ə and t+ɪ, respectively ('They would fade away', 'His fate is hard') and it is accompanied by a small difference in absolute voicing (1o msec). However, this is not the case in 'He bled like a pig' vs. 'He bet like a fool', so it is hardly significant.

5.1.1.2. The unstressed vowels

The unstressed vowels are more interesting. Firstly they are all the same, [ɪ], (in 'did' and 'it') and therefore can be compared 'vertically' as well as 'horizontally'. Secondly we find cases here where the difference in absolute voicing is small (less than, say 2o msec), and

Table XI Duration of vowels before d and t, - related to the differences in voicing in the consonants

	vowel duration before d	vowel duration before t	difference	difference in voicing in d and t in msec	difference in voicing (relative) in d and t
<u>strong</u>-weak					
æ(:)C+pause	288	164	124**	39	64**
ɛɪC+ə,ɪ	171	145	26**	1o	58**
æ(:)C+k	197	123	74**	44	74**
æ(:)C+g	192	119	73**	43	72**
ɔ(:)C+s	183	138	45**	33	56**
u(:)C+z	221	88	133**	31	6o**
ɛC+l	1oo	94	6	42	67**
weak-strong					
ɪC+pause	96	84	12	35	35**
ɪC+æ	6o	47	13*	14	11
ɪC+k	56	33	23**	13	2o
ɪC+g	56	38	18**	1o	4
ɪC+s	56	32	24**	17	19**
ɪC+z	59	45	14**	-2	14
ɪC+l	6o	41	19**	21	47**
weak-weak					
ɪC+æ	52	41	11**	13	19
ɪC+k	51	39	12**	19	13
ɪC+g	53	38	15**	13	14
ɪC+s	55	37	18**	13	16
ɪC+z	62				
ɪC+l	53				

the difference in relative voicing in the same cases is likewise small and not statistically significant (except in the case of IC+s in the weak-strong distribution) - we should thus expect the difference in vowel duration to be rather small.

5.1.2. Conclusion

The correlation between the difference in the duration of the vowel before d and t, respectively, and the voicing in these two consonants is displayed in Fig. 3. - The correlation, if any, does not seem to be a very simple one.[9] A hypothesis to the effect that voice assimilation in a (final) consonant affects the duration of the preceding vowel must probably be rejected. - But note that the assimilation concerns only voicing of final t, not, for instance, devoicing of final d.

The question of aspiration, or open interval, will be dealt with in the section about duration of the consonant.

5.2. Duration of the consonants

5.2.1. d and t

In Table XII are given the data on the durations of the d and t closures and their open intervals. "-" before a figure means that d is longer than t. "*" after a figure means that the difference is significantly different from zero at the 95% level, "**" indicates significance at the 99% level. "÷" denotes instances where the difference in relative voicing in d and t is not statistically signifi-

---

9) I have tried also to correlate the duration of the vowel before t with the voicing (absolute as well as relative) in t with no result, - there is no correlation to be found.

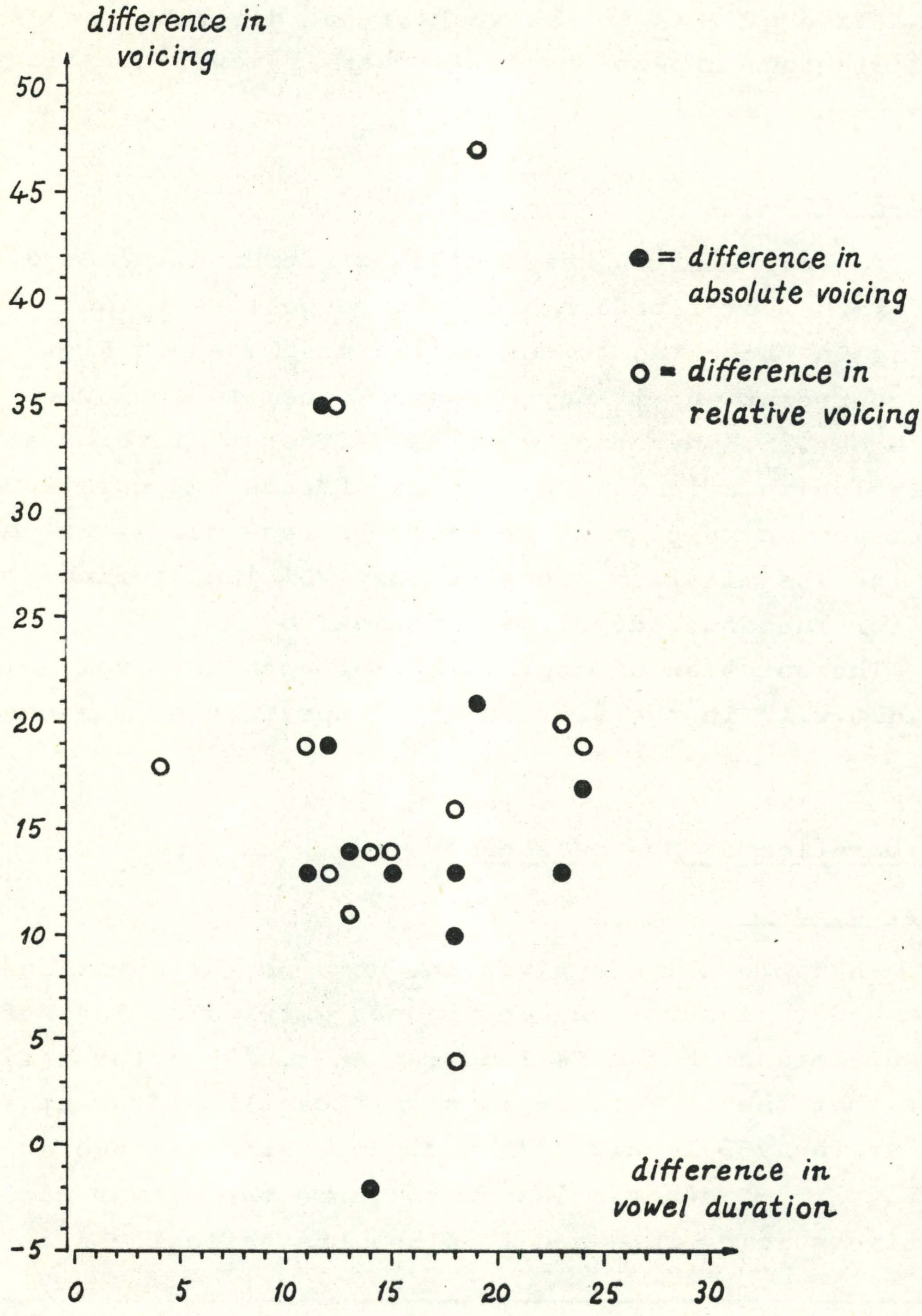

difference in voicing
50
45
40
35
30
25
20
15
10
5
0
-5
● = difference in absolute voicing
o = difference in relative voicing
difference in vowel duration
0
5
10
15
20
25
30

Fig. 3.

cant. "+" denotes instances where the difference in relative voicing in t and d is statistically significant, but less than 25 (%) (except in the leftmost column where it denotes word- or morpheme boundary). Figures that are not marked with "÷" or "+" are those where the difference in relative voicing in t and d is statistically significant and over 25 (%). "C" stands for d and t, respectively.

5.2.1.1. The open interval

Let me say with no further documentation that there is no correlation between the difference in open interval and the difference in voicing in d and t, i.e. the assimilation of voice does not affect the open interval.

It is evident that a consistent and significant difference in initial d and t before a vowel is to be found in the open interval. The difference manifests itself finally, too. But in this position the most effective cue to the distinction between d and t is probably not the open interval but rather the duration of the preceding vowel.

5.2.1.2. The closure

It is more difficult to discern a pattern in the closure durations. That is, the problem lies in establishing a pattern outside of cases of assimilation, for which the material is badly suited. - The only position that exhibits a stable difference in closure duration is finally in a stressed syllable after a vowel, where t is longer than d, as would be expected.

17 differences are negative, i.e. the d-closure is longer than the t-closure. These negative differences do tend to coincide with cases where the difference in voicing (absolute and relative) in d and t is small. -

Table XII Duration in msec of closure and open interval in t and d

	duration of t-closure	duration of d-closure	difference	open interval t	open interval d	difference
weak-strong						
pause+C				5o	11	39**
V+C	98	84	14**	47	11	36**
VC	93	83	1o	46	11	35**
p+C	58	67	- 9 ÷	4o	13	27**
b+C	61	6o	1	38		
s+C	53	84	-31**+	28	12	16**
z+C	96	1o7	-11	57	14	43**
l+C	1oo	113	-13	62	11	51**
strong-weak						
pause+C				39	13	26**
V+C	58	52	6*	4o	12	28**
VC	43	44	- 1	51	11	4o**
p+C	76	67	9 ÷	33	19	14**
b+C	62	63	- 1	32	17	15**
s+C	62	57	5	43	24	19**
z+C	58	73	-15* +	39	16	23**
l+C	51	51	o ÷	45	12	33**
weak-weak						
V+C	65	6o	5	38	13	25**
b+C	42	32	1o	42	15	27**
z+C	61	62	- 1	34	17	17**
strong-weak						
C+pause	8o	67	13			
C+V	53	32	21**	44	9	35**
C+k	77	64	13*	36	23	13**
C+g	68	62	6	28	16	12**
C+s	71	63	8			
C+z	1oo	71	29**			
C+l	1o2	75	27**			

continued on the next page

Table XII continued

	duration of t-closure	duration of d-closure	difference	open interval t	open interval d	difference
weak-strong						
C+pause	63	79	[-16]			
C+V	25	36	-11* ÷	25	13	12**
C+k	47	51	- 4 ÷	23	2o	3
C+g	47	56	- 9**÷	19	15	4
C+s	55	63	- 8* +			
C+z	78	64	[14**÷]			
C+l	67	56	11*			
weak-weak						
C+V	29	37	- 8* ÷	35	11	24**
C+k	41	56	-15**÷	23	18	5
C+g	47	53	- 6 ÷	22	14	8
C+s	49	55	- 6 ÷			

But there are other instances (framed in the table) where

(i) the d-closure is (considerably) longer than the t-closure and where the differences in absolute and relative voicing in d and t is rather large: 48 msec, 45 (%) in 'His doe was hurt'/'His toe was hurt' (closure difference "-11" msec), 55 msec, 45(%) in 'He'll dap again'/'He'll tap again, (closure difference "-13" msec), and 35 msec, 35 (%) in 'He'd looked at a cad'/'It looks like a cat' (closure difference "-16" msec);

(ii) the t-closure is longer than d-closure and where the difference in absolute and relative voicing in d and t is none or small: o msec, o (%) in 'The gap diminished'/

'There's a gap to fill' (closure difference 9 msec), and "-2" msec, 14 (%) in 'Did Zena do it?'/'It zipped at the back' (closure difference 14 msec).

Therefore I do not think that assimilation can be made responsible for the comparatively longer d-closures (or shorter t-closures).

A possible explanation might be that in syllable final position (where the unstressed d-closure usually is longer than the t-closure) the syllables chosen were 'it' and 'did'. 'it' may have been still more weakly stressed than 'did' and thus the t-closure shorter than the d-closure. But this explanation is not valid in the 5 (7) cases where the d-closure is longer than the t-closure initially.

5.2.1.3. Fortis-lenis

One thing may be deduced from this rather confused picture: If force and duration of the consonants go together in BE, i.e. if the consonant is longer the more 'fortis'[10] it is, and if length is thus a measure of 'fortisness', then the duration of the preceding vowel cannot be the result of the force of the following consonant, since I have examples of 'long' vowels followed by 'long' consonants. (Delattre (1940) explains the vowel length in French in this way, i.e. a vowel is shorter the more fortis (and the longer) the following consonant is.)

5.2.2. Consonants other than d and t

As regards the other consonants under investigation (p b k g s z): these behave according to expectations and not like t and d. The vowel preceding these consonants is

---

10) From now on 'fortis' and 'lenis' are used in their ordinary sense about force of articulation.

always longer before b g z than before p k s, and p k s are always longer than their basically voiced counterparts, Also where the difference in voicing has been eliminated. k vs. g and s vs. z exhibit differences in consonant duration that are independent of their phonetic voicing (I cannot say whether this would be true also of p vs b, since these two consonants have only been investigated finally in a syllable after a vowel, where their voicing does not vary much).

5.2.2.1. Duration of the consonants and their voicing

In Table XIII (which is only a sample) I have looked at some consonants after d and t, namely k and z after d and t, respectively, and s before d and t (these consonants are in otherwise identical positions).

Table XIII Voicing in and duration of k/z after d/t and of s before d/t

	voicing in msec	voicing in percentages	duration of the consonant (closure)
d+k/t+k	17/o	2o/o	86/1o5
d+z/t+z	51/6	63/6	81/11o
s+d/s+t	31/28	29/23	1o7/121

It will be seen that a higher degree of voicing is correlated with a shorter duration of the consonant (closure). Thus it seems that the assimilation of voicing causes a partial assimilation of the duration of the consonants, though not to such an extent that a basically unvoiced consonant becomes as short as its basically voiced counterpart.

5.2.2.2. d and t

Again d and t are atypical (see Table XIV).

Table XIV Voicing in and duration of d after s/z and before k/g and of t before s/z

	voicing in msec	voicing in percentages	duration of the closure
s+d/z+d	9/48	11/45	84/1o7
d+k/d+g	44/51	86/91	51/56
t+s/t+z	26/4o	36/4o	71/1oo

In about half of my examples (of which those above are only a sample) that closure is longer which is also the most voiced. This is true of both initial and final d's and t's. Thus it seems that d and t have behaved differently from the other consonants. I am not sure why this is so.

It is possible that the t-closure is shorter than the d-closure because of the longer open interval of t, i.e. the strong aspiration and affrication may be due to a rather loose closure which tends to be exploded rather early. The loose closure would explain why we find shorter t-closures also in such cases where there is no proper open interval (finally in a syllable before a consonant in the following syllable). But since not all t-closures are shorter than the d-closures, the loose closure cannot be operating always and may therefore be only a part of the explanation, the rest of which I am unable to offer.

5.3. Final remarks about vowel duration before voiced and unvoiced obstruents

Chomsky and Halle (1968 p. 3o1) propose vocal chord adjustment rate as an explanation for the lengthening of vowels before voiced obstruents, i.e. the movement to the open position for unvoiced consonants should be more rapid and cut off the vowel earlier than the supposedly more complicated movement to the vocal chord position for voiced consonants. This explanation is probably not generally valid: the vowel before final voiced t in my material is shorter than the vowel before voiced d (but, admittedly, it may be that the vocal chord position is different in voiced t from that in voiced d). In these cases t occurs in 'it' which is very short and weak, and therefore the vowel is particularly short. But there are also examples where the vowel before voiced p is shorter than before voiced b, and in these cases the choice of frame-word cannot be made responsible for the relatively shorter vowel before p.

6. Concluding remarks

Gimson (1966 p. 267-270) writes about assimilation, partly under the heading 'Allophonic Variations', partly under the heading 'Fortis/Lenis Variations'.

Among allophonic variations at word boundaries he mentions:

(i) Devoicing of l in 'at last' and other close-knit entities. A similar sequence in my material ('It looks like a cat') exhibited 71% voicing in l.

(ii) Devoicing of word final lenis fricatives and plosives. d in my material is only devoiced to a very small extent (91% voiced in stressed syllables, 77% voiced in unstressed syllables).

(iii) Lenis fricatives followed by fortis consonants get devoiced. z before t in my material was never less than 49% voiced, slightly dependent upon stress placement.

(iv) Initial lenis fricatives and plosives after a pause are unvoiced. In my material were found only a few instances with voicing lead.

Among fortis/lenis variations Gimson mentions:

(v) Word final lenis fricatives followed by a word initial fortis consonant may be realised as the corresponding fortis fricative, and the phonemic change may be complete in that a preceding long vowel may be realized in the reduced form appropriate to a syllable closed by a fortis consonant. This does not hold for my material. All subjects have more voicing in z than in s finally, all have s longer than z, and all have the vowel longer before z than before s.

(vi) b d g finally are not usually influenced in the same way as final lenis fricatives by a following fortis consonant. This is confirmed by my material, where final d after a vowel regardless of stress placement and regardless of the following consonant was voiced.

(vii) Word or morpheme final fortis consonants rarely show tendencies to assimilate to their lenis counterparts. This is confirmed by my material.

Gimson does not mention the progressive devoicing from final fortis to initial lenis consonants found in my material.

References

Chomsky, Noam and Morris Halle 1968: The Sound Pattern of English (New York).

Delattre, Pierre 1940: "Anticipation in the Sequence: Vowel and Consonant Group", The French Review XIII.

Fischer-Jørgensen, Eli 1952: "Om Stemthedsassimilation", _Festskrift til L.L.Hammerich_, p. 611-617.

- 1968: "Les occlusives françaises et danoises d'un sujet bilingue", _Word_ 24.

- 1969: "Voicing, Tenseness and Aspiration in Stop Consonants, with Special Reference to French and Danish", _ARIPUC_ 3/1968, p. 63-115.

Gimson, A.C. 1966: _An Introduction to the Pronunciation of English_ (London).

Grammont, Maurice 1939: _Traité de Phonétique_ (Paris), p. 185-228.

Thorsen, Oluf M. 1967: "Voice Assimilation of Stop Consonants and Fricatives in French and Its Relation to Sound Duration and Intraoral Air-Pressure", _ARIPUC_ 1/1966, p. 67-76.

A STUDY OF INTERNAL OPEN JUNCTURE IN BRITISH ENGLISH

Kirsten Gregersen

1. Introduction

The aim of the present paper is to investigate certain acoustic properties of internal open juncture[1] in the dialect of British English called Received Pronunciation.[2] More specifically duration of segments before and after internal open juncture are considered, and, to a lesser degree, the intensity of certain consonants. Other aspects, e.g. formant transitions and voice assimilation, were excluded.

Generally speaking internal open juncture may be expected to occur where there is spacing of letters in a written or printed text, but this is not true of all cases. In the phrase 'not at all' the /t/ in 'at' is pronounced as if it belongs to the following word, see Gimson (1962 p. 275); in 'Plato' some American phoneticians state that the /t/ is pronounced as if word-initial, see Gleason (1955 p. 43) and Hockett (1955 p. 172).

In handbooks of British English phonetics the manifestation of internal open juncture has been described impressionistically from Sweet (189o) to Gimson (1962), and, to my knowledge, only one phonetic experiment on internal open juncture has been published, viz. O'Connor

---

1) By different phoneticians termed internal open juncture, open juncture, internal juncture, and open transition.

2) For a description see Jones (1967 p. xviii).

and Tooley (1964). They made an auditory investigation with paired phrases only differing in respect to placement of internal open juncture which occurred either between vowel and consonant or between consonant and vowel. All tape-recordings with glottal stop were rejected.[3] Voiceless stops were found to be the most efficient juncture-markers whereas the results for voiceless fricatives, voiced fricatives and stops, and voiced continuants are rather poor.

American phoneticians have, on the other hand, delved deeper into the problem both from a phonological and a phonetic point of view. Since Trager and Bloch (1941) coined the term juncture it has been an integral part of most American linguistic analyses, including generative grammar, but the status of internal open juncture has been different in the various linguistic schools.[4] The first large-scale phonetic study was made by Lehiste (1960), and later several papers on the acoustic and auditory properties of internal open juncture have been published. Lehiste (1960) found that post-junctural allophones of almost all phonemes are considerably longer than either pre-junctural or medial allophones; vowels are excluded from this rule: they are longer finally than initially and initial allophones may start with laryngalization. Hoard (1966) states that pre- and post-junctural consonants show systematic difference of duration particularly when there is no glottal stop, only in one contrastive pair did he find significant difference in vowel duration (longer pre-junctural vowel in 'Nye trait' than medial in 'nitrate'); like Lehiste, Hoard found that post-junctural consonants are longer than pre-junctural.

---

3) This was done because the authors believe that glottal stop is an unequivocal boundary marker.

4) Lehiste (1960) gives an exhaustive treatment of the problem.

From the above may be concluded that an acoustic investigation of internal open juncture in RP was needed, so I decided to undertake such an experiment to determine whether its manifestations in RP could be compared to those found for American English. I also wished to see if the status of final and initial /n/ is still so precarious as it once was, e.g. the words 'umpire' and 'nickname' were in Middle English metanalyzed from 'a noumpire' and 'an ekename'.

2. Experiment

2.1. Phonetic notation

The notation follows that used by Jones (1967) with the divergencies mentioned below. "+" indicates possible placement of internal open juncture; "?" stands for glottal stop, creaky voice or laryngalization (not separated and henceforward termed glottal stop); "V" for vowel; and "C" for consonant. Square brackets are used for phonetic and slants for phonemic notations.

2.2. Material

When reading articles on juncture I collected a vast number of contrastive pairs only differing in respect to placement of internal open juncture, e.g. of the type 'gay mate', 'game eight'. To limit the material I selected those pairs where internal open juncture occurs either between vowel and consonant or between consonant and vowel, i.e. /..V+CV../ and /..VC+V../. Two pairs with internal open juncture contrasting with close juncture[5] were also included, as well as one phrase exemplifying metanalysis

---

5) Close juncture is generally signalled by absence of spacing in a written text.

in modern English and one with /n/ in close juncture. Table I shows the phoneme combinations and Tables II and III the actual phrases.

Three lists were made incorporating the phrases in random order. To ensure uniform stress- and intonation-pattern I thought it necessary to use the same frame for all phrases. As the phrases did not all belong to the same substitution class they were inserted into the quotation frames "The password is '...'" (list A) and "He said, '...'" (list B); in the third list the phrases were inserted as contrastive pairs with no frame, each phrase occurring twice, the second time in reverse order of the first.

2.3. Subjects

For the experiment I chose six subjects who had all spent their formative years in the Home Counties and belonged to the socio-linguistic group that speaks RP. They were from 25 to 46 years old, four male subjects were from the staff of Hull University, one female and one male were from the staff of the University of Copenhagen.

2.4. Recordings

The subjects were first asked to read aloud lists A and B, each list four times; then they were shown the list with contrastive pairs (list C) which they only read once; hereby I got ten recordings of each phrase for each subject. For lists A and B the subjects were asked to read them with a uniform stress- and intonation-pattern, they themselves chose their tempo, only for list C were they asked not to exaggerate the difference between the contrastive pairs.

That list C was given them after reading the two other lists was done to prevent the subjects knowing what was under investigation. I did not succeed in this; they all discov-

TABLE I

Phoneme-combinations included in the material

Post-junctural or post-consonantal vowel

Pre-consonantal or pre-junctural vowel	i:	i	æ	ɔ:	u:	ə	ei	ai	əu	au
i:		z	d		z		m	l		
ɔ:										t
u:	d									
ə							n		(l)	
ei		k		k		(t)	m	t,d		
ai							l			
əu							m			

All consonants in the table occur medially. Those in brackets may be preceded, but not followed, by internal open juncture, the others may occur both with preceding and following internal open juncture.

	V cs	%	S	clos. cs	%	S	op.int. cs	%	S	C-V cs	? cs	%?
shore tower	19.2	72	*	11.7	5o	*	8.6	62	*			
short hour	13.4			5.6			5.4			5.7	1.2	75
grey tie	23.3	72	*	9.9	66	*	9.3	66				
great eye	16.1			6.1			6.o			6.8	1.2	75
may kill	15.9	72	*	12.7	64	**	9.o	65				
make ill	11.1			7.9			6.o			5.7	2.2	7o
may call	15.5	76	*	11.5	68	*	9.4	66				
make all	11.2			7.4			6.3			4.3	2.2	63
average		73	***		62	***		65	***	5.6	1.6	71
New Deal	19.7	1o7		8.8	5o	***	2.9	143				
nude eel	21.2			4.3			3.8			5.7	3.1	76
free Danny	18.6	1oo		9.o	47	*	2.5	2o8				
freed Annie	18.5			3.9			4.2			4.7	2.1	75
grey dye	26.1	1o5		9.9	46	*	2.5	2oo				
greyed eye	27.1			4.1			4.3			5.6	2.2	77
average		1o4			48	***		184	**	5.3	2.5	76

TABLE II - Averages for all subjects

V: vowel duration, clos.: closure of stop, op.int.: open interval of stop, V-C: interval between stop and following vowel or glottal stop, ?: glottal stop, %?: percentual occurrence of glottal stop, %: /..V+CV../ in percentage of /..VC+V../, S: level of significance, * indicates significant difference at the 5% level, ** at the 1% level, and *** at the o.1% level.

TABLE III

Averages for all subjects

As Table II, except that C stands for consonant duration.

	V cs	%	S	C cs	%	S	C-V cs	? cs	%?
see zinc	19.4	103		11.5	87				
seize ink	19.7			9.8			5.7	2.6	70
see zoos	21.1	103		12.4	84	*			
seize ooze	21.7			10.4			5.7	1.9	72
average		103			85	**	5.7	2.2	71
hoe-maker	16.2	87		9.2	84				
home-acre	13.5			7.7			3.6	6.5	38
gay mate	21.6	83	*	9.2	110				
game eight	17.3			10.0			5.7	4.0	77
see Mabel	14.8	87	*	8.7	101				
seem able	12.7			8.9			4.0	5.4	63
average		85	**		98		4.4	5.3	59
a nation	6.7	94		9.6	60	***			
an Asian	6.3			5.8			2.8	4.5	47
a name	6.5	106		10.1	66	***			
an aim	6.9			6.7			4.7	3.3	63
average		100			63	***	3.7	3.9	55
I laid	17.6	110		10.2	88				
I'll aid	19.9			8.9			3.9	5.1	67
see lying	18.4	90		8.6	115				
seal eyeing	16.4			10.1			7.2	3.4	71
average		99			103		5.6	4.1	69

ered at least some of the contrastive pairs. Two things may account for this: (1) they all had a degree in modern languages (five in English and one in German) and may therefore have some knowledge of English phonetics, and (2) some of the phrases were far-fetched and perhaps suspicious.

The tape-recordings were made with professional equipment in the language laboratory of Hull University and in the Institute of Phonetics in Copenhagen.

A mingographic recording containing a duplex-oscillogram, a fundamental frequency curve, and two intensity curves (one linear without filtering and one logarithmic with high-pass filtering at 500 cps) were made.

3. Results for /..V+CV../ and /..VC+V../

The averages for all subjects are here considered; the level of significance has been calculated by means of a pair test based on the average for each subject (see Croxton 1953 p. 228f).

3.1. Vowels

The duration of vowels (termed "V" on tables and figures) before pre-junctural consonant and before internal open juncture was measured. Vowels are considered beginning where the intensity curves have reached the 'vowel-level', i.e. when they have stopped rising sharply (this is perhaps not quite correct, as the beginning of the vowel is excluded, but the absolute difference between two sets of measurements is not affected by this). As it was impossible to separate /r/ and /j/ from following vowel, /rV/ and /jV/ were measured instead (it is generally supposed that there is no difference in the duration of consonants that are separated from a following internal open juncture by one or two segments, so including them in the measures for vowels should not influence the differences in vowel duration).

The vowels are considered terminating where the intensity curves start falling sharply before stops, continuants, and pause. Before /z/ it was difficult to determine this point, because the intensity of /z/ was often higher than that of the vowel which, furthermore, did not fall; the only point common to all cases was where the high-pass filtered intensity curve and the duplex oscillogram started their respective sharp rise and fall. This point was therefore chosen to ensure comparable measurements, in spite of the inconsistency which then arises in the vowel durations, since the closing movement before stops is considered as part of the closure, whereas before /z/ it is included in the vowel. It was likewise difficult to separate /l/ from the preceding vowel, so these measurements are not reliable.

Vowels are significantly shorter before /tkm/ than before /+/, see Tables II and III, and Figs. 1 and 2. There is no significant durational difference before /dznl/ and before /+/.

3.2. Consonants

3.2.1. Obstruents

What is termed "clos." on Table II (and V) is roughly corresponding to the closure of the stops, it is measured from the point where the vowel ends to the release of the stop. For /tkd/ the closure is significantly longer pre- than post-juncturally, see Figs.3 and 4. To ascertain whether the closure measured for post-junctural stops was closure and not closure plus pause (pause and closure look alike on the mingograms) sonagrams were taken of a few phrases with particularly long closures.[6] From these it seems safe

6) This investigation was suggested by professor Eli Fischer-Jørgensen.

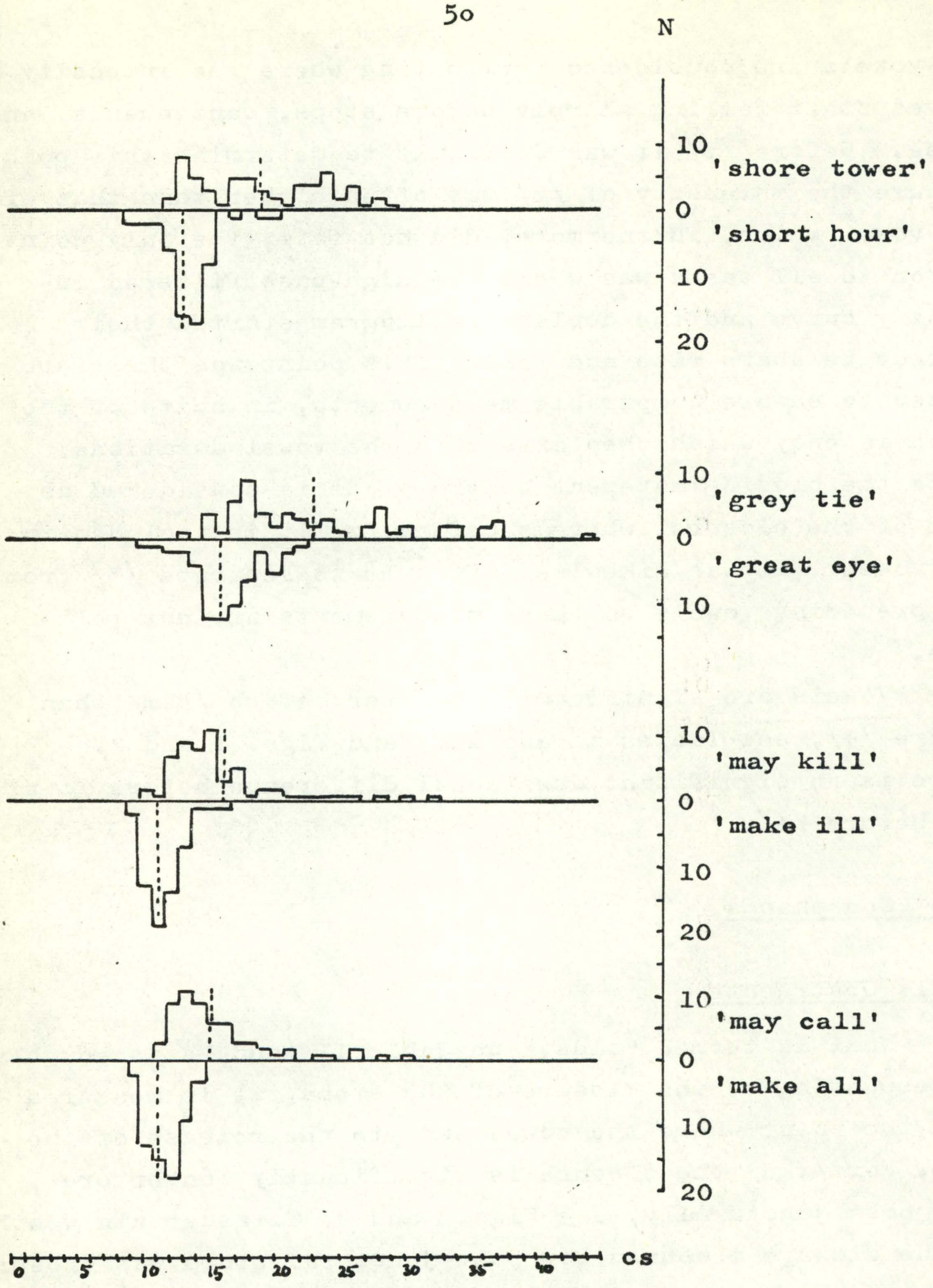


Figure 1.
Dispersion of vowel duration for all subjects.
Dotted line indicates average.

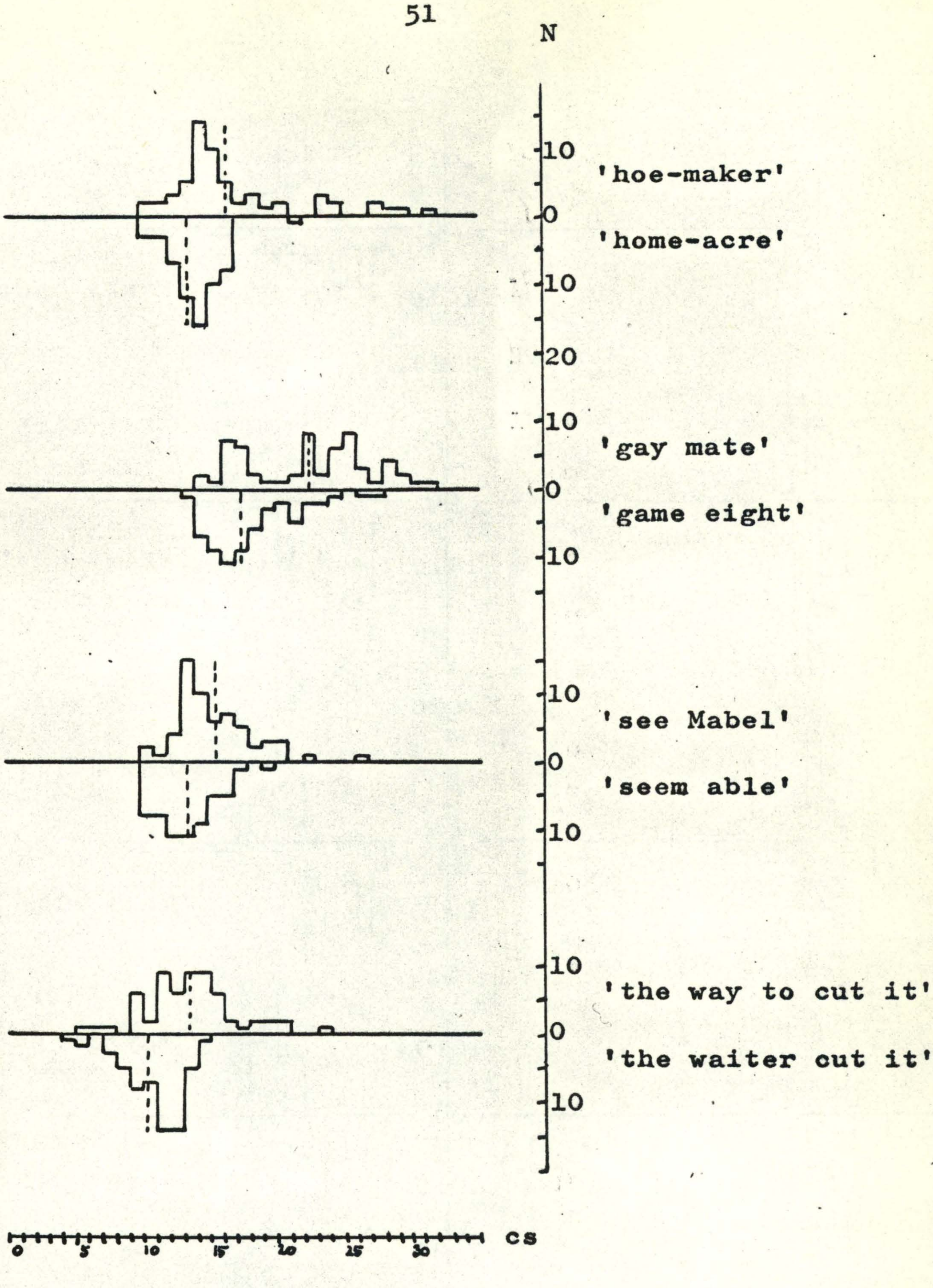


Figure 2.
Dispersion of vowel-duration.

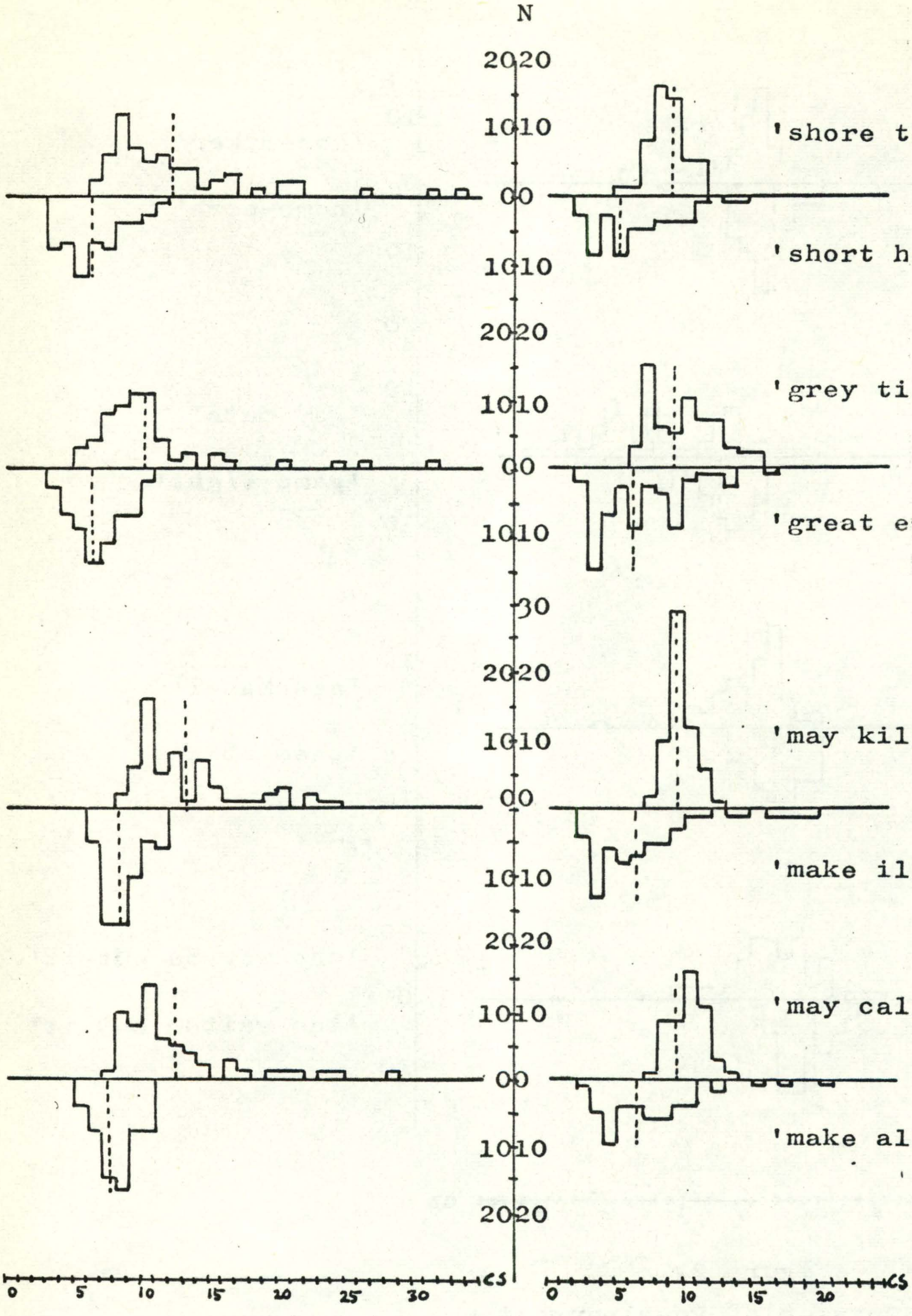


Figure 3.

Dispersion of duration of closure (to the left) and of open interval (to the right) for unvoiced stops.

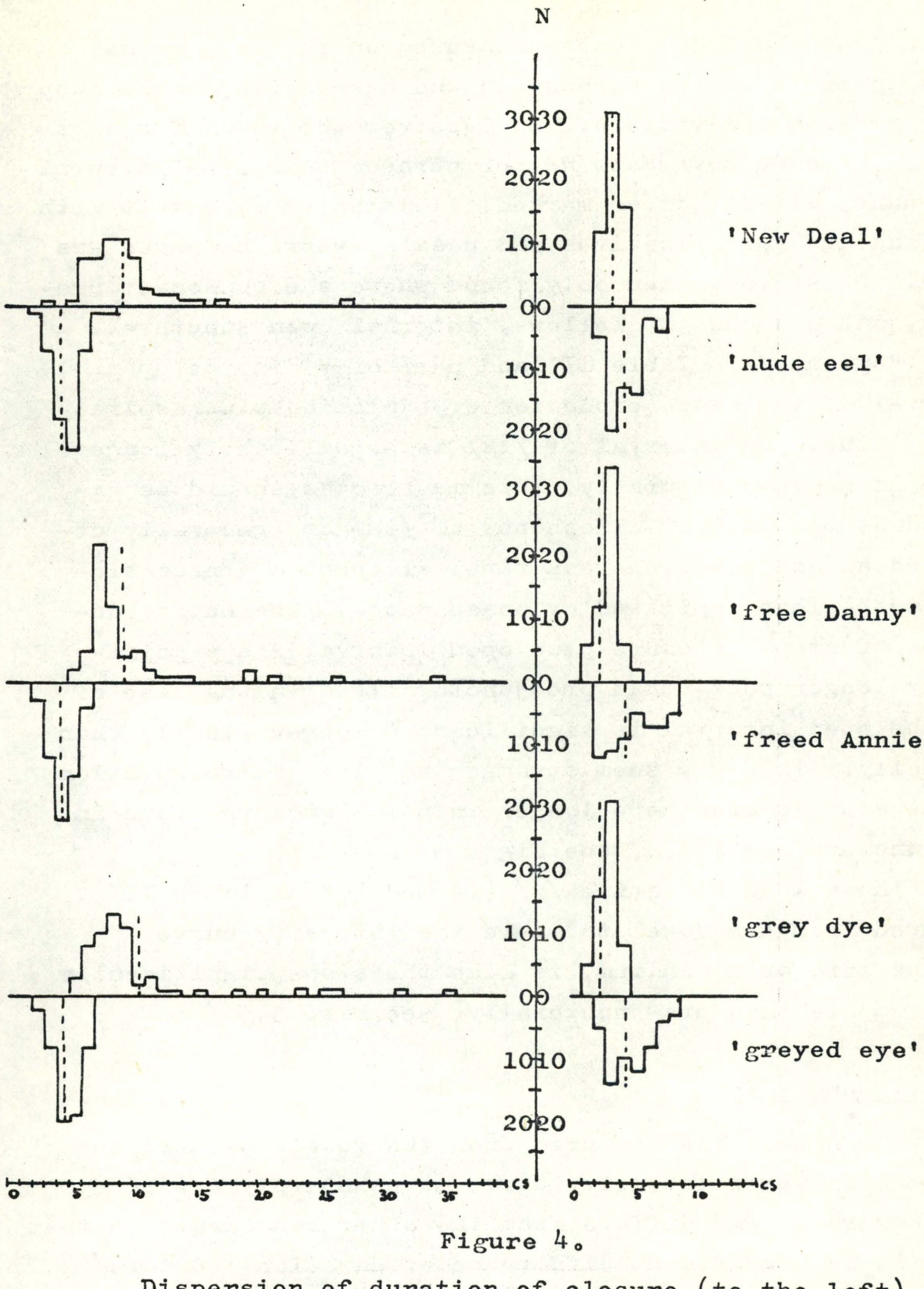


Figure 4.

Dispersion of duration of closure (to the left) and open interval (to the right) for /d/.

to conclude that the closure measured on the mingograms never covers a pause, because in the cases with dental stop the implosion was visible, and moreover the vowel formants showed the same movements as for phrases with pre-junctural consonant, albeit not so marked. This is in agreement with the findings for fricatives and nasals, where no pause was found. Pauses are then only found where the consonant precedes, and not when it follows, internal open juncture.

"op.int." on Table II (and part of V) is the open interval of the stop (explosion or explosion plus aspiration). The open interval of /tk/ is significantly longer post- than pre-juncturally, this is also as should be expected as the initial allophones of /tk/ are generally described as aspirated whereas final allophones are considered to be less aspirated or unaspirated. The entire duration of stops (closure plus open interval) is significantly longer post- than pre-juncturally. In the case of /d/ the open interval is significantly longer finally than initially. This may seem strange, but the entire duration of the stop is also here longer in post-junctural than in pre-junctural position, see Figs. 3 and 4.

The voiced fricative /z/ (termed "C" on Table III), measured from the vowel to where the intensity curve reaches zero or a minimum, is like the stops significantly longer post- than pre-juncturally, see Fig. 5.

3.2.2. Sonorants

The nasals are measured from the vowel to where the high-pass filtered intensity curve moves sharply upwards or downwards. /m/ differs from the other consonants in that there is no significant difference of duration. But /n/ shows exactly the same durational difference as that found for obstruents, see Fig. 5. For /l/ no difference was found, but the segmentation was not reliable.

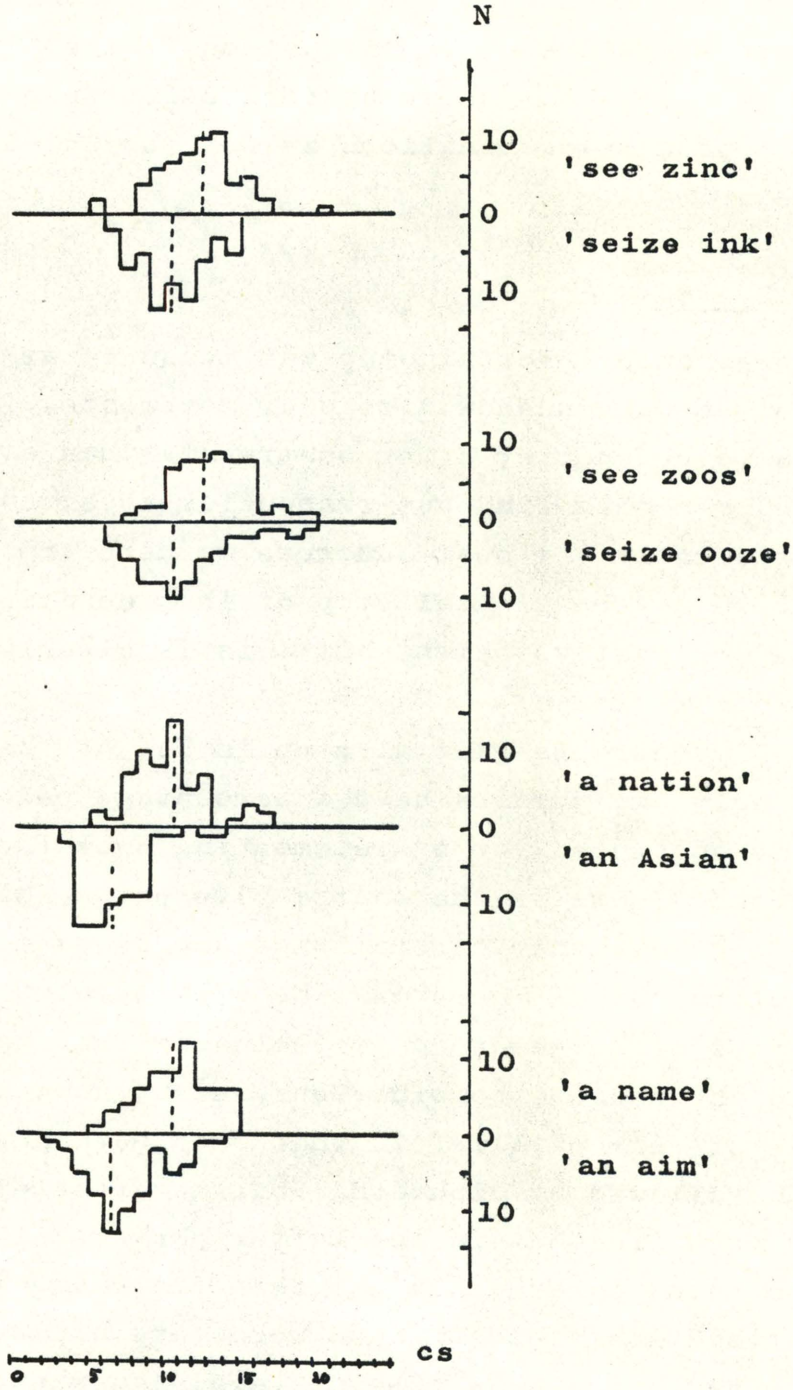


Figure 5.
Dispersion of consonant duration.

3.3. Interval from /C/ to [?] or /V/

This was measured from the end of the consonant to where the curves again show amplitude, on Tables II, III and V it is termed"C-V". The durational difference between the different phoneme-combinations is not very great and may be disregarded.

3.4. Glottal stop

The duration of glottal stop was measured as the interval where the curves show irregular movements. This interval seems to be shorter after obstruents than after sonorants. I can only find one reason for this: unvoiced stops are apparently the best juncture markers and therefore there need not be glottal stop of long duration to separate the contrastive pairs, but this is clearly not true for either /d/ or /z/.

Generally there is glottal stop in 70% of the phrases /..VC+V../, only two phrases have a percentage below 50. The 70% may be accounted for by remembering that the subjects had noticed some of the contrastive pairs, moreover some of them spoke rather carefully as one tends to do in front of a microphone. There must be a particular reason for the phrases with less than 50%; clearly this may not be sought in the phoneme-combinations, e.g. 'home-acre' and 'game eight' are parallel in that they both have nasal preceded and followed by diphthong, but nevertheless glottal stop is more frequent in the latter phrase. The cause must then be found on other levels than the phonemic. In most grammatical analyses compound words are closer connected than, say, noun plus numeral (e.g. 'game eight'). It is therefore reasonable to suppose that this closer grammatical connection has been signalized by fewer instances of the strong juncture-marker glottal stop.

Considering the other results for the pair /həu(+)m(+)eikə/ ('hoe-maker':'home-acre') there are no significant differences in the averages for all subjects; all subjects do, however, have significant durational differences: two have both longer pre- than post-junctural vowel and shorter pre- than post-junctural consonant, one has longer pre- than post-junctural vowel, and three have shorter pre- than post-junctural consonant. It may then be concluded that there is here no stable feature signalizing internal open juncture.

It is curious to note that there are fewer instances of glottal stop in 'an Asian' than in 'an aim' though they have the same grammatical structure. One would also here suppose that indefinite article plus noun would be rather closely connected grammatically. The discrepancy between the two phrases may perhaps be accounted for by the number of times 'an aim':'a name' has been mentioned as a contrastive pair in phonetic descriptions of English; the subjects may have recognized it and been on their guard against pronouncing the phrases alike and therefore introduced glottal stop. The reason 'an Asian':'a nation' is not mentioned as often is that it is not always a contrastive pair, 'Asian' being pronounced /eiʃən/, /eiʃjən/, /eisjən/, eizjən/ or /eiʒən/ (all subjects pronounced the phrases with identical segmental phonemes, though).

3.5. Maximum intensity of stops

The maximum intensity of the open interval was measured on the linear intensity curves, for male subjects the integration time was 10 ms and for the female subject 5 ms. Both integration times are rather long for the open interval of stops, i.e. the curve may not reach its maximum.

The maximum intensity was significantly higher for post- than for pre-junctural unvoiced stops, see Table IV and Fig. 6; but as long open interval is correlated with

TABLE IV

Averages for all subjects

As for Table II, except that op.int. stands for maximum intensity.

	op.int. dB	%	S
shore tower short hour	29 26	89	*
grey tie great eye	31 27	87	
may kill make ill	25 21	83	*
may call make all	25 21	83	**
average		86	***
New Deal nude eel	26 23	88	*
free Danny freed Annie	23 24	99	
grey dye greyed eye	25 25	1o1	
average		96	
the way to cut it the waiter cut it	29 3o	1o4	
not at all	28		

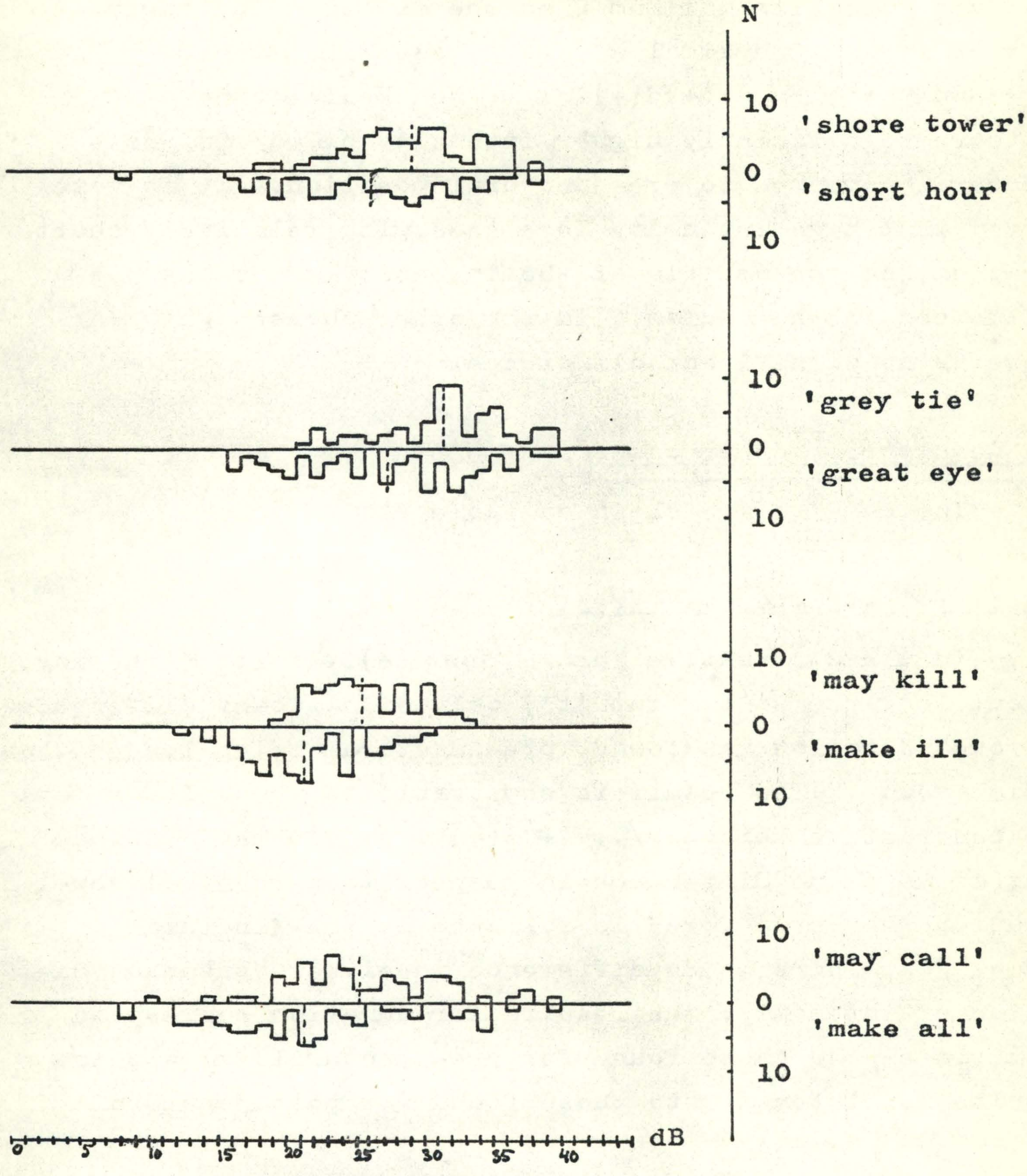


Figure 6.
Dispersion of maximum intensity of unvoiced stops.

high intensity one cannot be sure whether the intensity really reaches its maximum when the duration is long or whether it is a physical artifact, see Fig. 7.

Only for /nju:(+)d(+)i:l/ ('New Deal':'nude eel') is there significantly higher intensity of /d/ in post-junctural compared to pre-junctural position. It is noteworthy that high intensity here goes with relatively short duration, so the inertia of the intensity-meter has not influenced these results. In the other phrases with /d/ there is no significant difference.

4. Results for other cases

The results are shown on Table V.

4.1./..V+CV../ and /..VCV../

In the contrastive phrase /ðəwei(+)təkʌtit/ ('the way to cut it':'the waiter cut it') only significant difference in vowel duration was found, pre-junctural being longer than medial vowel. The result is comparable to those found for the contrastive phrases /..V(+)t(+)V../; one may conclude that it makes no difference in respect to percentual vowel duration whether a vowel is followed by pre-junctural or medial /t/. There is no difference in either consonant duration or intensity; the results for duration correspond most closely to those found for pre-junctural /t/ and the results for intensity to those found for post-junctural /t/.

The results for /ə(+)ləun/ ('a loan':'alone') are unreliable because of the vocalic quality of /l/, which makes the segmentation difficult.

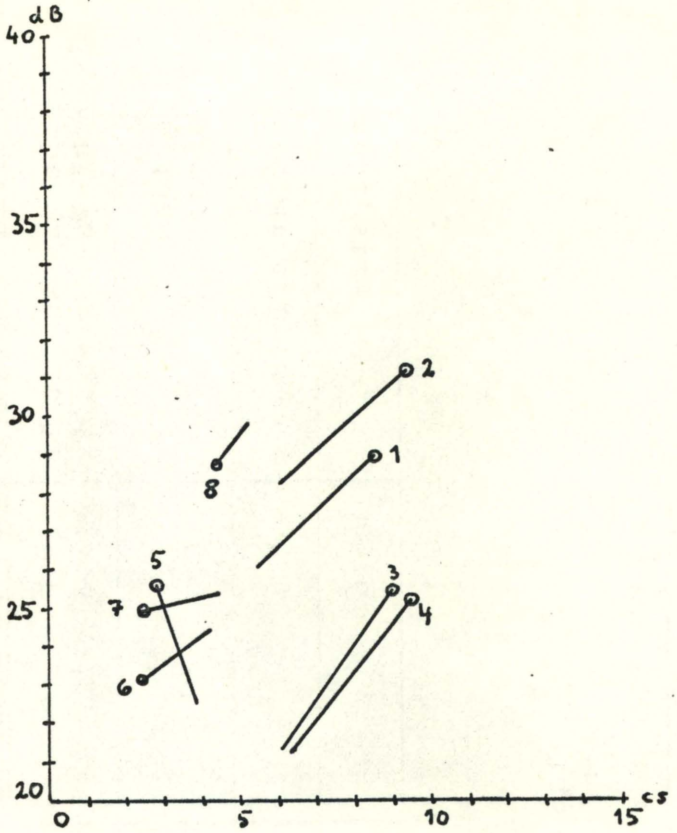


Figure 7.

Maximum intensity of stops compared to duration of open interval.

Circle indicates result for post-junctural and end of line result for pre-junctural or medial stop.

1. /ʃ ɔ:(+)t(+)auə/ 5. /nju:(+)d(+)i:l/
2. /grei(+)t(+)ai/ 6. /fri:(+)d(+)æni/
3. /mei(+)k(+)il/ 7. /grei(+)d(+)ai/
4. /mei(+)k(+)ɔ:l/ 8. /ðəwei(+)təkʌtit/

	V cs	%	S	clos. cs	%	S	op.int. cs	%	S	C cs	%	S
the way to cut it the waiter cut it	12.8 1o.o	79	*	5.9 5.8	97		4.3 5.2	121				
a loan alone	7.4 6.9	93								1o.3 9.1	91	
not at all	3.4			6.5			6.5					
analysis	5.6									7.o		

TABLE V

Averages for all subjects

As for Tables II and III.

4.2. Unpaired phrases

4.2.1. 'not at all'

After repeated listenings of the tapes it was found that all subjects used the initial allophone of /t/ in 'at'. Because the vowel in 'at' is short and unaccented its duration cannot be compared to that of the contrastive pairs with /t/. The duration of /t/ is closest to that found for pre-junctural /t/, the intensity does not belong clearly to either group. It then seems to be the case that for perceiving initial /t/ it is sufficient that the distance C-V is zero.

4.2.2. 'analysis'

In this word the first vowel is slightly shorter than that in the contrastive pairs /..V(+)n(+)V../, the duration of /n/ is closest to that of pre-junctural /n/.

5. Results for two subjects

Two subjects had results that differed slightly from those of the others; they are, however, included in the averages, but it also seemed worth-while to consider them separately. The results of the percentual duration of segments are shown on Figs. 8 and 9.

5.1. Female subject IN

The most evident divergency from the other subjects is in the open interval of unvoiced stops, it being longer pre- than post-juncturally. There may be several reasons for this. IN is bilingual, her father being English and her mother Danish.[7] In English there are in final position six stop-phonemes, in Danish, on the other hand, we have only three stop-phonemes in this position, generally pronounced

7) She spent her childhood and youth in England, and for the last 2o years she has lived in Denmark.

[b̥ d̥ g̊], so it is common for Danes not to make any distinction between voiced and unvoiced final stops when speaking English. Therefore she may exaggerate the difference between the two sets of stops by aspirating the unvoiced ones to make the Danes perceive the difference. It could, of course, equally well be an individual variation. The entire consonant duration is still longest for post-junctural unvoiced stop.

The open interval of /d/ is also abnormally long, but it shows the same tendency as for the other subjects. But the open interval is so long pre-juncturally that there is practically no difference between the entire duration of post- and pre-junctural /d/.

The duration of /z/ is longer pre- than post-juncturally; the opposite was the case for the other subjects. In /geim+eit/ ('game eight') and /si:m+eibəl/ ('seem able') the /m/ was likewise abnormally long. It is curious to note that in /həum+eikə/ ('home-acre'), where there is only one instance of glottal stop, the /m/ is significantly shorter in this position than post-juncturally.

She used glottal stop in most post-junctural vowels.

It is doubtful whether IN's results should be included in the averages or not. The following appears to happen to levels of significance if her results were excluded: for duration of open interval of /tk/, and for duration of /z/ they would improve, whereas the opposite is the case for vowel duration before /tk/ and for open interval of /d/.

5.2. Male subject PW

This subject read the lists most naturally and rapidly. This may account for his very sparing use of glottal stop.

He is the only subject who does not have a significant difference in vowel duration before unvoiced stops. The open interval of /d/ is shorter pre- than post-juncturally

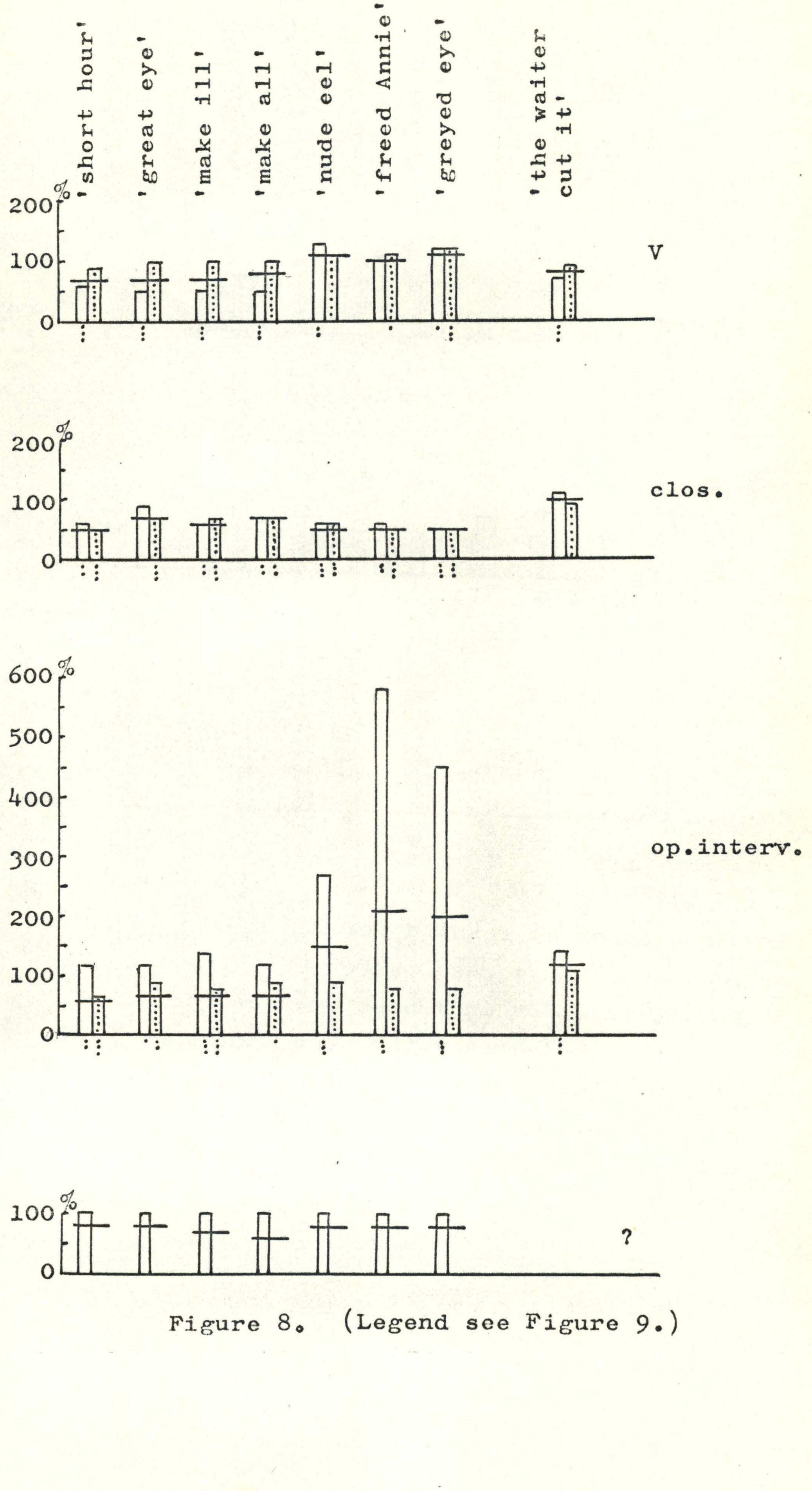


Figure 8. (Legend see Figure 9.)

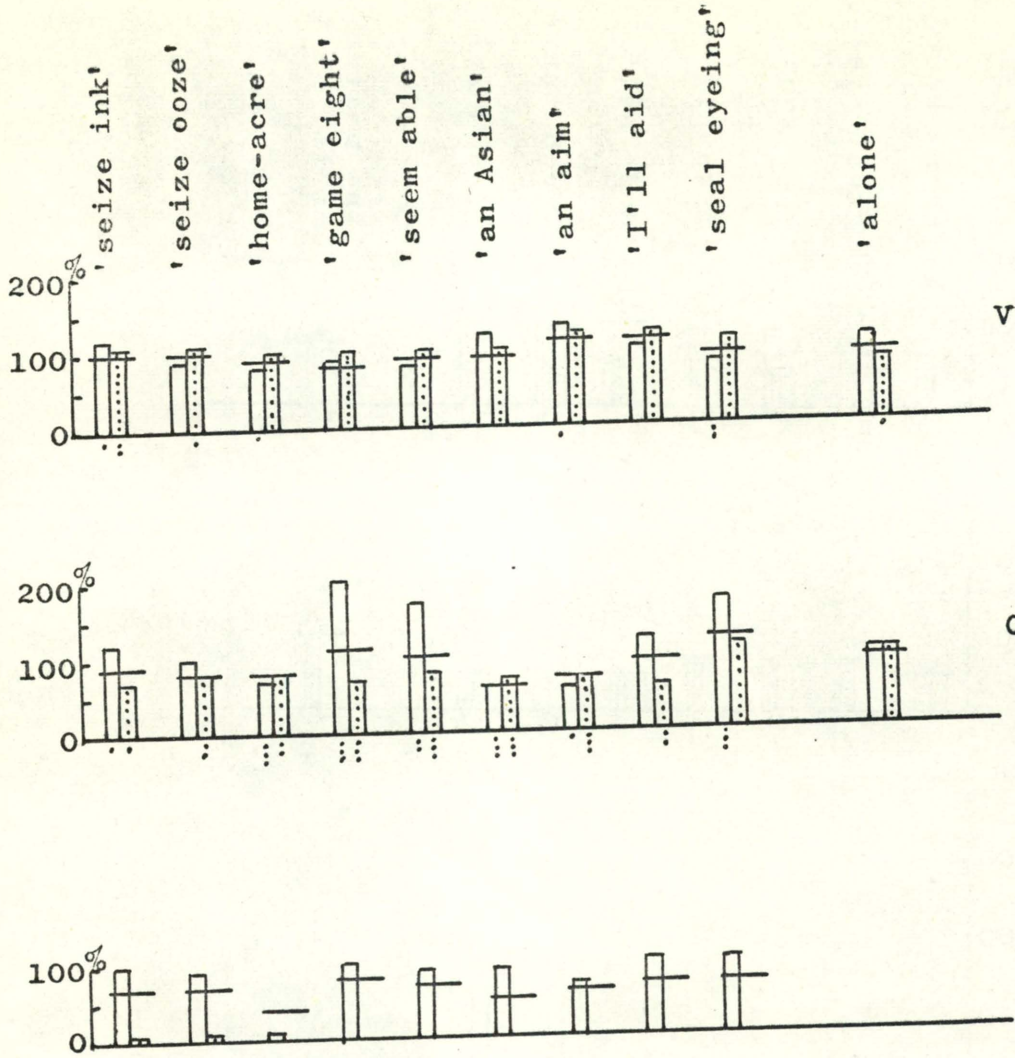


Figure 9.

White columns: results for IN.
Dotted columns: results for PW.
Horizontal line indicates average for all subjects. Dots below the zero line show level of significance , one dot at the 5%, two dots at the 1%, and three dots at the 0.1% level. Other symbols as for Tables II and III.

(the opposite was the case for all other subjects), so the entire duration of /d/ is markedly longer post-juncturally. The duration of /m/ was shorter in pre-junctural position in all three contrastive pairs, so it may be supposed that the phonetic difference between close and less close grammatical connection disappears with increasing speed of utterance.

If PW's results were excluded from averages it seems to improve the levels of significance for vowel duration before /tk/ and open interval of /d/.

6. Conclusion

The results may be summarized thus (see Figs. 10-12): vowels are shorter before unvoiced stops and, to a lesser degree, /m/, than before voiced stops, voiced fricatives, /n/, and internal open juncture. Consonants other than /m/ are shorter before internal open juncture than after internal open juncture; medial /n/ has the same duration as final /n/ whereas medial /t/ corresponds to initial /t/. Furthermore post-junctural vowels may be preceded by glottal stop and pause (no pauses were found before post-junctural consonant). These two things seem, however, to disappear with increasing speed of utterance and closer grammatical connection, as also do differences in vowel duration before unvoiced stops with or without internal open juncture.

The results found for RP then correspond closely to those found for American English by Lehiste (1960) and Hoard (1966). Lehiste states that pre-junctural allophones of consonants may sometimes be drawled, this explains the long pre-junctural consonants of IN. Hoard's results agree completely with those found for PW. It appears that there is no difference between the manifestation of internal open juncture in American and British English.

As for the examples without glottal stop it is possible that my material would also yield better results

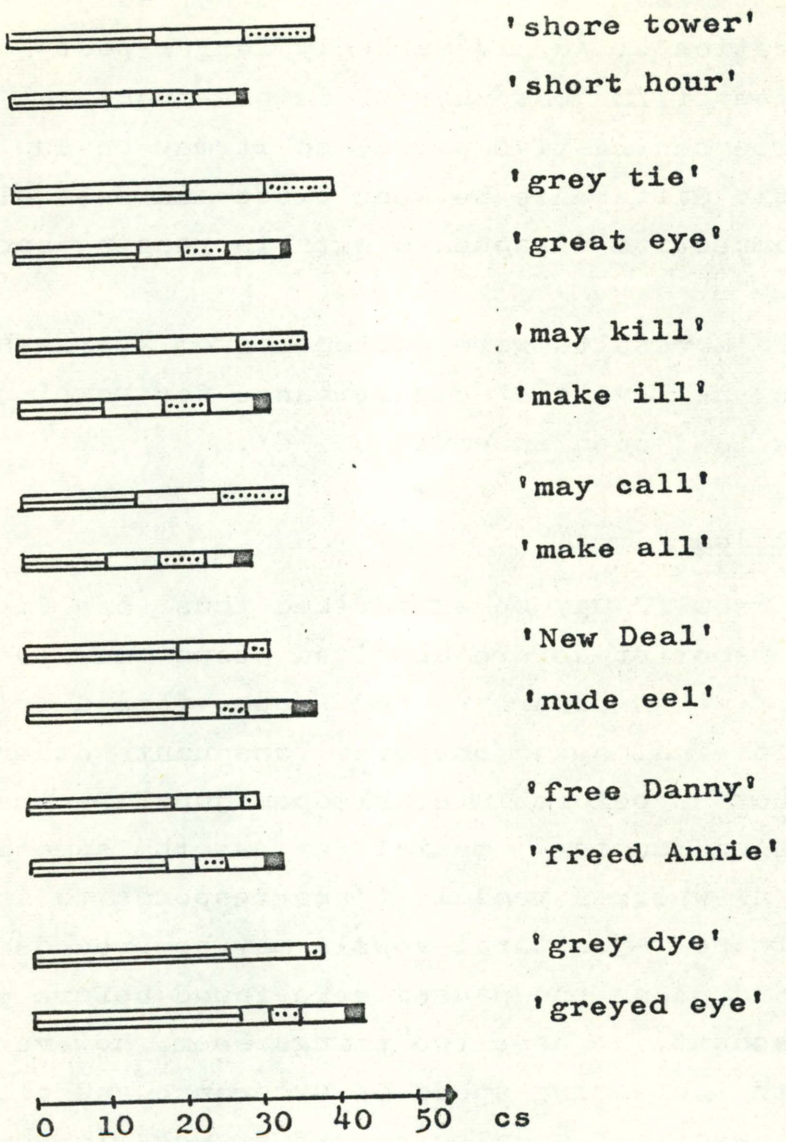


Figure 10.

Average duration of segments.

vowel duration

duration of closure or interval C-V

duration of open interval

duration of glottal stop.

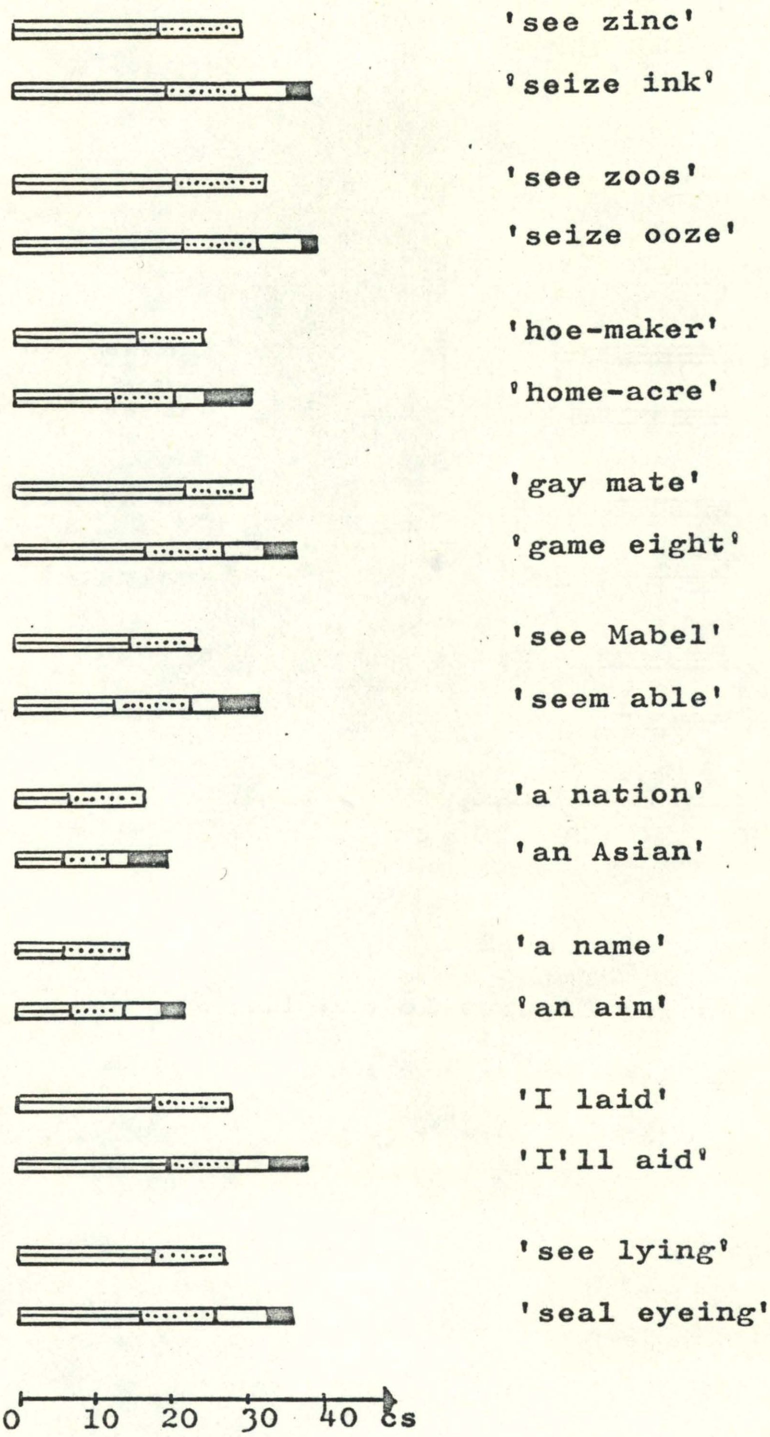


Figure 11.

As for Figure 10 except that

signifies consonant duration.

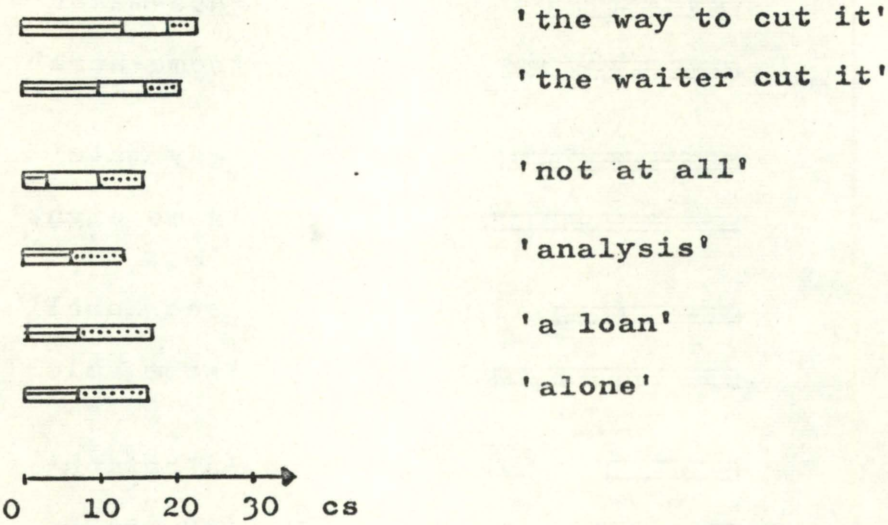


Figure 12.

As for Figures 1o and 11.

for unvoiced stops than for any other consonants in an auditory test in agreement with the findings of O'Connor and Tooley (1964). This is true not only of English; in an analysis of Swedish Gårding (1967) found that unvoiced stops and vowels were the best juncture-markers (in her material phrases with glottal stop were not excluded).

From the acoustic results it would not seem probable to mistake final for initial /n/, but it is not possible to verify until auditory tests have been made.

References

Croxton, Frederick E. 1953: Elementary Statistics (New York).

Gimson, A. C. 1962: An Introduction to the Pronunciation of English (London).

Gleason, H. A. 1955: An Introduction to Descriptive Linguistics (New York, rev.ed. 1961).

Gårding, Eva 1967: Internal Juncture in Swedish (Lund).

Hoard, J. E. 1966: "Juncture and syllable structure in English", Phonetica 15, p. 96-1o9.

Hockett, Charles F. 1955: A Manual of Phonology (Baltimore).

Jones, Daniel 1967: English Pronouncing Dictionary (London, 13th ed.).

Lehiste, Ilse 196o: "An Acoustic-Phonetic Study of Internal Open Juncture", Phonetica suppl.ad vol. 5, p. 1-54.

O'Connor, J.D. and Olive M. Tooley 1964: "The perceptibility of certain word-boundaries", In Honour of Daniel Jones (London), p. 171-176.

Sweet, Henry 189o: A Primer of Phonetics (Oxford, 3rd ed. 19o6).

Trager, G. L. and B. Bloch 1941: "The Syllabic Phonemes of English", Language 17, p. 223-246.

A COMMENT ON LEXICAL INSERTION

Jørgen Rischel

1. Introductory remarks

In working with the generative phonology of a highly inflectional or derivational language one repeatedly faces the problem how to distinguish processes that are genuinely phonological (synchronically speaking) from what is more appropriately considered "morphological alternation", suppletion, and the like. Now, any attempt to define the limits of morphology involves an hypothesis about the nature of lexical items, including their representation in terms of marked boundaries, phonologically specified idiosyncracies, and non-phonological idiosyncracies. It goes without saying that it is difficult to discuss these matters in a meaningful way unless we know where the lexical items come from, and at which point in grammar lexical insertion takes place. Thus it is that contemporary work in syntax <u>and</u> in phonology has to face one and the same crucial problem: how does "lexicon" fit into a transformational grammar?

The reflexions of which some elements are presented below, were provoked as conceptual prerequisites by research on the interrelationship between morphology and phonology in languages exhibiting a high degree of morphophonemic complexity (viz. Danish and West Greenlandic Eskimo). However, the specific questions of lexical insertion on which the present paper concentrates are of a very general nature, and the subject-matter has indeed received a good deal of attention in the last few years in studies of well-known languages such as English. I have for obvious reasons (such as the advantage of using familiar quotations for illustration) preferred to refer to examples in English although my intuitions about the idioms cited are often insufficient for a serious analysis.

I have made no attempt to survey the literature on the subject. It must be mentioned in particular that it has not been possible to include a discussion of the general literature on morphological and lexicological theory in the present paper which is to be understood only as a comment to the discussion of lexical insertion in some recent papers by McCawley (1968) and other advocates of the generative semantics trend of tranformational grammer. The validity of generatice semantics is not at issue here;[1] it is the sole aim of the paper to present some, mainly morphological and phonological, evidence bearing upon the place of lexical insertion in a type of grammer in which underlying representations are assumed to be of a semantic kind.

2. The introduction of lexical items in grammar

2.1. Deep versus late insertion

According to the now classic theory of transformational grammar (see e.g. Chomsky 1965) the base component of grammar produces strings consisting of chunks of morpheme-size, which are supplied with a phonological specification. The syntactic transformations operate on these items, and in the course of these processes other phonological specifications may be introduced. That is, lexical insertion takes place essentially <u>before the transformational component</u> of grammar, but some of the

---

1) The adoption of this kind of framework does not imply a dogmatic belief in it as the only possible type of grammar. However, as long as generative semantics (in essentially the form this theory has at present) has not been shown to be inadequate in dealing with problems of syntax, it seems fruitful to explore its implications for morphology and phonology.

transformations require additional lexical information to be introduced later.

According to generative semantic theory the rules of syntax operate on an underlying semantic structure. Apparently a major part of the syntactic component is pre-lexical (in European structuralist terminology: syntax is concerned with content rather than expression). McCawley posits a whole hierarchy of sentences underlying a surface sentence like John killed Bill, the surface verb kill being derived - by successive applications of a cyclical rule called PREDICATE RAISING or PREDICATE LIFTING - from such predicates as CAUSE, BECOME, NOT, and BE ALIVE. This rule lifts the verb of the lowest sentence, i.e. BE ALIVE, to the next sentence, so that the lifted verb comes to stand adjacent to NOT, and on the next cycle it lifts the whole verb complex to the sentence above it, so that it comes to stand adjacent to BECOME of the higher sentence. By lexical insertion at this point we get die (BECOME NOT ALIVE). If, however, the rule applies once more, the verb resulting from earlier applications is lifted to the highest sentence so that it comes to stand adjacent to CAUSE, and by lexical insertion we get kill (CAUSE BECOME NOT ALIVE).

It is implied by this analysis that there is not necessarily any syntactic difference between the derivational relationship of kill to die and the derivational relationship of, say, break in he broke the glass to break in the glass broke. Similarly, since it is obvious that the complex predicate NOT ALIVE underlies the surface adjective dead (if McCawley's analysis is at all correct in detail[2]), it follows that the derivational

---

2) The validity of this particular analysis is immaterial to the general principles.

relationship of die to dead corresponds to that of darken to dark, etc. In each case there is a difference between arbitrary substitution of lexical material (die - kill, dead - die) and a more regular type of derivation (break - break, dark - darken), but this reduces to merely a matter of morphological irregularity: die - kill is nothing but an instance of suppletion, if we wish to formulate the relationship within the framework of derivational morphology.

The analysis of kill which I have sketchily referred to above, implies that lexical insertion must occur after cyclical rules like PREDICATE RAISING. On the other hand, it is claimed that lexical insertion occurs before some post-cyclical transformations that move constituents around, since these rules may depend on specific lexical items (I shall revert to this kind of argument later).

2.2. Cyclic versus non-cyclic insertion of lexical items

It is supposed that lexical insertion takes place either last in the transformational cycle or at the very beginning of the post-cyclic rules of the syntactic component.[3] - As far as I can see the latter solution may mean that a form like kill directly replaces a lexical entry containing the semantic material CAUSE, BECOME, NOT, BE ALIVE in some configuration, after it has been brought together by PREDICATE RAISING. There may be no connection between the lexical items die and kill except that the lexical entries for these items share semantic material. - The situation is quite different if lexical

3) These alternatives were presented by McCawley at the First Scandinavian Summer School, Stockholm 1969.

insertion is cyclic. The lexical item _die_ should, according to that conception, be inserted on some cycle, and on the next cycle PREDICATE RAISING applies to this item in its phonological form so that the higher sentence comes to contain a complex consisting of CAUSE and _die_. When lexical insertion applies again this hybrid configuration is taken as a lexical entry, and the item _kill_ emerges. Lexical items like _die_ and _kill_ are thus very directly connected since the causative form is transformationally derived from the phonological representation of the verb _die_.

Obviously lexical insertion is somehow cyclic if it can be shown that this process is sensitive to the phonological shape of some of the material that makes up a lexical entry.

Lakoff (1970 p. 78-79) refers to the _come_-_bring_ situation as evidence in favour of the hypothesis of lexical decomposition: "The ordinary sense of "come" is related to the ordinary sense of "bring" by a predicate of direct causation (...) In addition, there are many idiomatic expressions containing the phonological form _come_, whose corresponding causative has the phonological form _bring_ (...) There are enough of such cases to require that a rule be stated relating the cases with "come" and the cases with "bring" (though there will, of course, be exceptions to any such rule). In the lexical decomposition framework, the rule of predicate-lifting will create complex predicates such as "CAUSE - _come_". The regularity is that "bring" substitutes for such a complex predicate."

Phenomena like this one do not prove that lexical insertion is cyclic. However, the relationship between _come_ and _bring_ is most easily accounted for if we assume some kind of ordering so that the various readings of come are first replaced by a common representation, after which the causative bring is derived by a lexical rule (which is more or less insensitive to the derivational history of come). The cyclic interpretation of lexical insertion sketched above is one approach to this kind of solution.

There are other types of evidence bearing upon the question, but we shall leave this aside here.

2.3. Is cyclicality of lexical insertion phonologically plausible?

According to the cyclicality hypothesis a morphologically indivisible chunk like bring is inserted in two steps: first the form come is specified in its phonological shape (i.e. specified to the same extent as in those contexts where it appears as a surface verb), and afterwards the form bring is derived from the form come plus the element CAUSATIVE. This conception of the process seems to me inescapable if lexical insertion is to be cyclic.

There can hardly be any formal objection against the postulation of transformations that takes some phonological material plus some semantic material and replaces it all by some totally new phonological material. But it is not a particularly convincing hypothesis about what normally goes on in language. It seems awkward to introduce phonological feature matrices in syntax with the sole purpose of having a set of such matrices trigger the insertion of a new set of totally unrelated feature matrices.

This is an entirely general problem concerning suppletive morphology. In what sense can e.g. the phonologically specified stem be be said to "underlie" the forms am, is, was? I find it questionable whether the analysis of competence can be correlated with a reasonable model of performance if we claim that non-productive, more or less atomic formations like bring or was take more derivational machinery than regular formations like (causative) break or (pret.) drown+ed. It would indeed be inexplicable why some of the most central words in language should be represented by forms exhibiting the longest paths of derivation. - It seems intuitively more satisfactory to assume that the aforementioned idiosyncratic forms are costly only in terms of lexical representation. It would not matter if they took a good deal of "space" in lexicon since these very forms are highly frequently used items. - I shall revert to this crux of morphology in section 3.1. below.

As for the relationahip between specifically come - bring and die - kill it seems reasonable to assume that bring is equivalent to come plus the element of CAUSATIVE; the parallelism of die - kill is less convincing, but it is possible to see the connection. As mentioned earlier in this paper the relationship between bring and come can be considered as suppletion, since we have numerous cases where causative verb and the item from which it is derived, turn up as identical surface verbs: break, turn, etc. In other cases the item from which the causative is derived turns up as a surface adjective, cf. dry, clean (or the underlying predicate appears both as surface verb and adjective with more or less identical phonological specifications, cf. fit). It is obviously a lexical rule of English that causative verbs can be derived with no phonological change from items that occur as non-causative verbs or adjectives (the derivation come ⇒ bring, and the blocking of causative *come, must be stated as an idiosyncracy, but that poses no problem in the cyclic framework). Now, why is _kill_ derived from _die_ rather than from _dead_ (if it is at all derived from any of these)? The adjective _dead_ obviously reflects a stative predicate, from which _die_ is derived as a perfective verb. Thus, the above derivation involves a postulate: causative verbs are derived from inchoative rather than stative predicates (cf. the derivation of _bring_ from COME rather than from BE THERE, or the like). If we follow this principle, the transitive i.e. causative, verb _clean_ is not directly derived from the stative predicate reflected by the adjective _clean_, but it is derived from an intermediate derivation which means something like BECOME CLEAN.

How should lexical insertion be imagined if there is no surface item reflecting such intermediate derivations?[4] In the case of _clean_ from CAUSE TO BECOME CLEAN the obvious solution is to insert an intermediate form which is later blocked if no further derivation occurs: (i) lexical insertion provides the entry BE CLEAN with a phonological form: _clean_, (ii) predicate raising applies, and lexical insertion for BECOME + _clean_ gives an unaltered phonological specification ("zero" derivation), (iii) predicate raising applies again, and lexical insertion for CAUSE + _clean_ gives an unaltered phonological specification ("zero" derivation once more). Output constraints block the lexical output _clean_ in the stative sense unless it appears as an adjective, and they also block this output in the inchoative, "medial" sense, whereas the output is permitted in the causative sense. - This solution presupposes _either_ that lexical output constraints are global rules ("remembering" the derivation) _or_ that there is a diacritic marking of lexical items indicating what they are derived from, in order to distinguish permitted and non-permitted uses of _clean_.

---

4) Cf. Postal (1970) p. 87-88, footnote 37.

The approach outlined above provides us with a very powerful tool, and it must be seriously asked whether it is not too powerful. In the case of kill, what prevents us from positing a "zero" derivation from an item *kill that is more or less synonymous with die? Similarly, instead of deriving bring from come, we might claim that come has a near synonym *bring, which,however, is used only if causative derivation has applied to it. This would mean that lexical insertion in the case of COME may either give come or *bring (possibly as an optional choice in some cases but not in others) and that lexical output conditions block the causative derivation if the string contains the form come, whereas they require this very derivation if the string contains the form bring. The obvious advantage of this kind of solution is that a phonological specification of the verb stem occurs only once in the above-mentioned set of derivations, and that the form of lexical entries becomes easier to state.

On the other hand, if we give up the idea that lexical insertion of bring involves the phonological specification of come, we do not really capture the regularity mentioned earlier, viz. that a great many expressions with come also occur with causative bring. It is interesting whether a generalization to this effect can be made without involving lexical transformations like come ⇒ bring. I shall attempt to approach this problem by considering the place of "idioms" in general in grammar.

3. The identity of lexical entries

3.1. Are idiomatic expressions lexical items?

In a grammar with a deep structure idiomatic expressions like to kick the bucket (i.e. 'to die')

to pull someone's leg

to go west (quoted as British army slang 'to die')

can be generated like other grammatical strings, and their specific use as metaphors becomes a matter of semantic interpretation (cf. the interesting discussion of "meaning rules" in Kiparsky 1970 p. 277).

In a generative semantic theory of syntax, on the other hand, the underlying structures of such expressions must reflect their metaphorical sense. It may, therefore, be assumed that the lexical entries for kick the bucket or go west are essentially similar to that for die, and that the lexical entry for pull someone's leg at least shares some properties with that for fool. If that is true, there may be no connection between the constituent structure of the string to which lexical insertion applies and the apparent surface structure of the output. A string like kick the bucket will be a lexical item just like the stem die.

It is, however, obvious that idioms like the above are not "morphemes". A sentence like

he kicked the bucket

is formed by inserting the past tense affix -ed inside the string kick the bucket, i.e. if the latter is a lexical item there must be a transformation applying after the insertion of the idiom with the effect:

he [kick the bucket] -ed ⇒ he kicked the bucket

This can be accounted for in a decently simple fashion if the idiom chunk has a lexical phrase marker associated with it, so that kick is marked as a verb and the bucket as a noun phrase. Similarly, the sentence

he pulled Tom's leg

contains an idiom whose invariant part is pull ...'s leg, with an associated marking of phrase structure so that the tense affix and the object noun can be put into their proper places by transformations with the combined effect:

he [pull ...'s leg] -ed Tom ⇒ he pulled Tom's leg

However, morphological idiosyncracies complicate the matter. Compare the sentences:

he pulled my leg
he went west

As regards "portmanteau morphs" like my and went, there are two reasonable hypotheses: (i) that they are derived by a morphological process after insertion of the "stem"; (ii) that they are inserted directly as lexical items.

(i) If we assume that forms like my, went are produced in a separate, late subcomponent, i.e. morphology, we may set up derivations like

he [pull ...'s leg] -ed me ⇒ he pulled me's leg
⇒ he pulled my leg
he [go west] -ed ⇒ he go-ed west ⇒ he went west

However, if we derive my from me+s, went from go+ed, we are back in the problem of come - bring: it is unsatisfactory to introduce a phonologically specified item with the sole purpose of replacing it by something else. An alternative conception of the relationship between lexicon and morphology is that lexical insertion rules generate sets of truly synonymous alternants such as

GO ⇒ { go / *wend / (etc.) }   PAST TENSE ⇒ { -ed / -t / (etc.) }   (for simplicity standard orthography is used

whose members may combine with each other to produce forms like goes=*wends, *goed=went, etc., whereas morphology comprises a set of conditions on such forms, i.e. functions as a filter that blocks unpermitted forms. The problem with this latter conception is, as mentioned in section 2.3. above, that it is too powerful a device in grammar theory. In fact, the linguist is in danger of being involved in a jungle of ad hoc decisions such as a choice between *wend (cf. send) or *wean (i.e. wēn, cf. mean) as the underlying "stem" of went.

(ii) As suggested in section 2.3. we may instead take "portmanteau morphs" as specified in lexicon, so that much of what is traditionally called "morphology" comes under lexical insertion proper. This means that there are lexical entries of the form

GO+PAST

I+POSS.

and that lexical insertion, when applying to these complexes, produces the items went and my directly.

A general evaluation of the relative merits of "morphological" and "lexical" approaches to the specification of such idiosyncracies as suppleting alternants cannot be undertaken in this paper.[5] If, however, the latter approach is considered from the point of view of idiomatic phrases it has obvious shortcomings within the framework of the current generative semantic conception of lexical insertion. Consider the literal and metaphorical readings of the sentence he went west. On the literal reading went would be inserted directly as a lexical item, but what about the metaphor went west? If we set up go west and went west as separate lexical items, we burden the lexicon with an enormous lot of items. If, on the other hand, we have just one lexical item for the metaphorical expression, we shall have to derive irregular forms anyway after postlexical transformations.

This would fail to capture the important generalization that the pattern go-went in the idiom is the same as the inflection of the verb go in its ordinary sense.

---

5) It will be apparent that I do not quite agree with Kiefer (1970 p. 5) that "we may safely conclude that to treat inflectional morphology in the lexicon is completely inadequate". (I hope to expound on this elsewhere.)

This situation is unacceptable. Lexical insertion must treat the verb go in such idioms as go west, go to hell, etc. as the same item as the ordinary verb go, otherwise we shall have to state the same morphological facts several times though there can be no reasonable doubt that go in the different uses is mastered as the "same" verb by speakers of English.[6] It is misleading to treat such occurrences of a verb in different idiomatic expressions as "homonymy" on a par with, say, the homonymy between race meaning 'species' and race meaning a kind of contest, but this is exactly what we should be forced to do.

3.2. "Idiom" formation

The evidence considered above suggests that "idioms" may not be a matter of lexical insertion only. The idiom go west contains the verb go, which is a lexical item, and the adverb west, which also is a lexical item. There is nothing lexically strange whatsoever about this phrase, and that conditions its effect as a metaphor (unconventional reference, cf. Reddy (1969)). Though the real nature of metaphor formation is enigmatic, I see no fatal consequences of the working hypothesis that the ambiguity of go west, kick the bucket, etc. is established "before" the insertion of phonologically specified lexical items takes place. I suggest, therefore, that there is a subset of pre-lexical rules in syntax which take inputs like DIE and produce outputs like KICK THE BUCKET. The output from such a rule will be a syntactic tree of the same kind, but its internal organization may be slightly or highly deviant from the tree structure of the input. It is to the derived tree that lexical insertion applies. There would of course be many such rules, so that we may define a whole subcomponent of syntax which functions as an IDIOM GENERATOR.

---

6) Morphology may decide on semantically doubtful cases. Cf. Weinreich (1966 p. 466): "of the two homophonous verbs ring (1. 'sound', 2. 'encircle'), it is to the former that the ring of ring the changes 'exhaust the variations' is related (cf. rang the changes, not ringed the changes)."

In the case of an idiom like kick the bucket the semantic input and output of the rule are quite different (although they share the property of being predicates). Pre-lexcial idiom formation in grammar must mean that the "generator" takes a (derived) semantic representation and replaces it by another semantic representation.[7] If this is to make sense, the latter must also be a derived representation which essentially fulfills conditions imposed upon semantic trees by the syntactic rules that precede idiom formation. In the particular idiom kick the bucket the representation inserted by the rule may have to be marked as metaphorical (if some of the transformations that are blocked in metaphors are later), but otherwise it will be identical with the representation of the phrase KICK THE BUCKET in its literal sense. The funny effect of many such idioms is due to the fact that they can be understood both as derived directly from the underlying representation of their literal sense and, via idiom formation, from a quite different underlying representation. If this were merely a consequence of arbitrary homonomy of lexical entries it would be hard to explain why so many metaphors are syntactically perfectly well-formed on a literal reading.

If we now return to the surface verb come in its various idiomatic uses: come about, come up (for discussion), etc., these undoubtedly have more or less different underlying representations, but the assumption would be that the idiom generator replaces them by one common representation, viz. that of the "ordinary" verb come. If this representation forms a surface verb, lexical insertion gives come, but if the element CAUSE has been lifted and included in the verb, lexical insertion gives bring. There will be one lexical entry for each of these surface verbs, which explains the regularity of the

---

7) Semantically this will be a mirror image of the process suggested by Weinreich (1966 p. 453); also cf. Kiparsky (197o p. 277).

alternation come - bring.

The treatment of the different idiomatic uses of come will thus differ from the treatment of homonyms like race 'species' and race 'competition'. In the former case a common representation is inserted by idiom formation, in the latter case two lexical entries happen to yield identical phonological representations. It goes without saying that a distinction between these two types of derivational coalescence is extremely difficult to make in actual practice, but that is a problem which has concerned lexicologists for a long time (within more or less different frameworks of description). I have nothing useful to say about this at present, but another example may help to illustrate the difference:

In Danish there is a surface verb $træde_1$ 'tread', which also occurs in many idiomatic expressions such as træde i spinaten 'put one's foot in it' (literally: 'step on the spinach'). There is another verb $træde_2$ 'thread (a needle)'. The two verbs $træde_1$ and $træde_2$ have more or less similar inflections, whereas the various idiomatic occurrences of $træde_1$ have absolutely identical inflections. This suggests that $træde_1$ is mastered as one verb listed at one place in the lexicon, whereas $træde_2$ is a different verb, which just happens to be homonymous with $træde_1$ in some of its forms. But of course the internalized lexicon may be organized differently depending on the input data one has received. I remember as a child hearing $træde_2$ in the infinitive in expressions like jeg skal lige træde nålen 'I just have to thread the needle'. My reaction was that this must be some funny idiom containing $træde_1$, and it was not until later that I realized that it is a different verb derived from tråd 'thread'. According to the model outlined above, this later statement would mean that some information was moved from the "idiom generator" to the lexicon proper, partly because of the acquisition of data which must belong in the latter place.

4. The bipartite lexicon

According to the hypothesis outlined above a surface verb like come in such expressions as come about is inserted via two kinds of rules: (i) idiom formation, which changes the underlying predicate to the verb COME, and (ii) lexical insertion, which provides this verb with a phonological form. Both kinds of rules belong to the subject-matter that is traditionally conceived as "lexicon" in broad sense. Thus, what the hypothesis actually implies is that lexicon consists of two parts, one of which operates on derived semantic representations to produce structures that qualify as lexical entries,[8] whereas the other performs the insertion of phonological material. For brevity I shall refer to the former part as "$lexicon_1$", and to the latter as "$lexicon_2$".

Now, at what point in grammar does lexical insertion occur? Grosu (1971 p. 42) mentions the behaviour of eject and throw out in favour of lexical insertion preceding certain post-cyclic transformations: "throw out and eject could probably replace the same semantic configuration, but only the former can be affected by the particle movement transformation". If, however, there is a distinction between $lexicon_1$ and $lexicon_2$ this argument may be interpreted as valid for $lexicon_1$ if we define this subcomponent in a wider sense so that its rules do not only replace semantic material but also perform purely structural restatements of syntactic trees. It may be supposed, for example, that the rules of $lexicon_1$ take as their input some representation of the meaning THROW OUT and produce as their output either a verb $[_V \text{THROW OUT}]_V$ or verb plus particle $[_V \text{THROW}]_V$+OUT.

---

8) The relationship of these alleged processes to the word-building and semantic extension types of rules presented in Kiparsky (197o p. 266f) cannot be treated in this paper.

If the former representation occurs as input to the rules of $lexicon_2$ we may get the phonological specification of eject; if the latter representation occurs we have two appropriate lexical entries, and we get the phonological specification of throw plus that of out. I.e. throw out will not be phonologically specified as a lexical item per se, which is exactly the kind of economy that would be desirable. (This presentation is, of course, grossly oversimplified. Firstly, eject and throw out are only synonymous on one reading; secondly, eject may perhaps have a complex structure for some speakers of English, cp. that its parts occur elsewhere: emit, inject.)

Under this assumption the rules of $lexicon_1$ are of essentially the same kind as other transformational rules of the syntactic component,[9] and they may not form a well-defined subcomponent at all.

The particle movement argument obviously vanishes as far as $lexicon_2$ is concerned, but if it is at all valid it certainly shows that the rules of $lexicon_1$ must apply before certain post-cyclic rules. The lexical insertion rules of $lexicon_2$, on the other hand, may well apply at a very late point in grammar, perhaps even at the very end of the syntactic component. However, further consideration of these problems requires a discussion of the internal organization of lexical rules, which is outside the scope of the present paper.

---

9) Much of $lexicon_1$ will, of course, be highly idiosyncratic, and it may seem awkward to include such material amongst other syntactic rules. Green (197o p. 81) speaks directly against the approach outlined here: "Unpleasant implications follow from the claim that there is, for example, only one lexical entry refuse, and that the other meanings are derived by syntactic processes of incorporation or deletion of semantic constituents."- I agree that many types of polysemy may be profitably treated in terms of related lexical entries (rather than derivational "translation" into one entry), but I do not see at the moment how this copes with the problems presented above.

References

Chomsky, Noam 1965: Aspects of the Theory of Syntax.

Green, Georgia M. 1969: "On the Notion 'Related Lexical Entry'", Papers from the Fifth Regional Meeting, Chicago Linguistic Society.

Grosu, Alexander 1971: "On Coreferentiality Constraints and Equi-NP-Deletion in English" (ch. 2: "Pre-lexical Transformations and Semantic Primes"), Working Papers in Linguistics No. 7 (Computer & Information Science Research Center, The Ohio State University).

Kiefer, Ferenc 197o: Morphology in Generative Grammar (= Interim Report No. 29, Research Group for Quantitative Linguistics, Stockholm).

Kiparsky, Paul 197o: "Semantic Rules in Grammar" in: Hreinn Benediktsson (ed.), The Nordic Languages and Modern Linguistics, p. 262-28o.

Lakoff, George 197o: Linguistics and Natural Logic (= Studies in Generative Semantics No. 1, The University of Michigan).

McCawley, James D. 1968: "Lexical Insertion in a Grammar without Deep Structure", Papers from the Fourth Regional Meeting, Chicago Linguistic Society.

Postal, Paul M. 197o: "On the Surface Verb 'Remind'", Linguistic Inquiry 1, p. 37-12o.

Reddy, Michael J. 1969: "A Semantic Approach to Metaphor", Papers from the Fifth Regional Meeting, Chicago Linguistic Society.

Weinreich, Uriel 1966: "Explorations in Semantic Theory", Current Trends in Linguistics III (ed. Th. A. Sebeok), p. 395-477.

A STUDY OF EMG PEAK AMPLITUDE MEASURES.

P.Mansell[1]

1. Introduction

In Mansell (1970b) an explanation was sought for the phemonenon of non-systematic variation in EMG waveforms associated with speech articulatory gestures. The waveforms under consideration were derived from surface electrodes on the upper lip and had undergone amplification, rectification and integration. It was noted that this variation was to be found on all the parameters conventionally employed to quantify these waveforms.

It was argued that the view we take of the significance of these variations to models of speech production is dependent upon our view of the relationship between the articulatory movement and the derived EMG signal. If we see the EMG as being directly related to the movement in each utterence of a token of a linguistic type, then we must develop production models which allow for variable outputs in response to a constant input. It is to be noted that this argument holds whether we take the EMG waveform to be the result of a set of "context-free" commands, with "coarticulation" being a purely mechanical effect weighting the relationship between the EMG and movement, or whether the EMG is taken to be the result of sets of commands which are themselves contect-sensitive. If, on the other hand, we are willing to see the relationship between EMG and movement as more complex, then alternative explanations for the significance of the variation can be derived. We might suppose, for example,

1) University of Essex.

that not all the muscle activity represented by the EMG waveform is of relevance to the articulatory gesture of concern - only a certain level of activity, say, may be required for such a gesture, the higher levels of activity recorded being representative of overshoot phenomena. Continuing this hypothetical line of argument, we might further suppose that amplitude and timing measures within this overshoot region (and peak amplitude measures for a large number of tokens would necessarily fall in such a region) would be liable to variation and of little significance in investigation of the organization of speech gestures.

In Mansell (1970a, 1970b) some implications of the former approach were explored. Basic to both approaches, however, is a more satisfactory description of the nature of the variation. In Mansell (1970b) the correlation between a limited number of the conventionally used parameters, including EMG peak amplitude, was investigated, and the resulting very low correlations regarded as providing some evidence of the inappropriateness of the parameters concerned to the description of relevant aspects of the EMG signal. The nature of the variation is not directly approached in this paper (but see section 5.2. below): rather, I am concerned to ask whether the events occurring at any electrode site are typical of lip activity as a whole, as measured simultaneously at other sites. This measure I have chosen to call the reliability of EMG peak amplitude measures. To demonstrate the reliability at a suitable level of significance will not decide anything about the causes of variation at particular sites. But if demonstrated this would resolve any doubts about the contribution of purely local factors to the variation.

2. Experimental method

The utterances for this experiment, all occurring monosyllables of English, and listed in Table 1, were

Table 1: Utterances used.

$C_n V_r C_n$	$C_l V_n C_n$	$C_n V_n C_l$
Cool	pack	cap
	back	cab
	mac	cam

where n = neutral with regard to lip activity
l = labial
r = rounded

chosen to provide a number of situations typical of those studied in the EMG speech literature, and at the same time provide reasonably simple derived waveforms for amplitude measurement. The articulator of interest is the upper lip, and the utterances allowed for the study of the contrast between voiced, voiceless and nasal labial stops, in syllable-initial and syllable-final positions, as well as the contrast between activity for labial consonants and a rounded vowel.

Four channels of EMG were utilized. The subject's upper lip was cleaned with detergent and the electrodes (sanded silver cup, approx. 5 mm. in diameter) were deployed in a Common Reference Derivation (see Cooper, Ossleton & Shaw 1970). That is to say, four active electrodes were taped to the upper lip as symmetrically as possible with respect to the midline, and all connected to a single reference electrode taped to the nose. Connecting all the active electrodes to a common reference electrode enables us to derive signals which are not only synchronous in time but which also have strictly comparable amplitudes in that one of the inputs to each of the four differential amplifiers is a constant. It was felt that such an electrode system was particularly suited to this study, since it enabled comparisons to be made among a number of amplitude measures during the same utterance.[2]

The centres of the electrodes labelled Right Centre and Left Centre in the presentation of results were sited approximately 1 cm. from the midline of the lips. The electrodes labelled Right Corner and Left Corner were placed with centres approx. 2 cm. lateral to the centre electrodes (but see section 5.1.2.). Thus they were not placed sufficiently laterally to pick up interference from other muscle groups which might have been active for spreading gestures.

The items analyzed in this paper were elicited as part of a larger corpus involving, in addition to the utterances reported here, more complex interactions between

---

2) It will be seen that the techniques of handling the data used in this paper are of such a nature as to cancel out the effect of constant variations between electrode sites. Such constant differences are not likely to affect correlation measures, and there is the check of bilateral symmetry to guard against such effects on amplitude measures. The possibility remains, of course, that uncorrelated variations may occur at each electrode site as a result of peripheral processes, which phenomena would be confounded with data relating to lip activity. Such a possible source of error is allowed to stand here, since the examination is essentially of the feasibility of current practice of using amplitude measures as discriminants of EMG activity: and current practice does not include any means of controlling for peripheral variations.

the labial stops and the rounded vowel, together with items involving a labial fricative. The items, in orthographic form, were arranged in lo randomized lists, which were read by the subject at a steady pace. The frame for each monosyllable was:

I'd like to say a __________ again.

As noted, the signals from the electrodes were led to four differential amplifiers (Fonema), and were recorded together with the audio signal on an Ampex SP 300 FM recorder running at 15 ips. The signals were then HP filtered to eliminate low-frequency movements artefacts due to changes in skin-electrode contact. They were finally rectified and integrated (5o msec. time constant) before being displayed on a 4-channel Mingograf ink-jet recorder. Two runs of the data were made, each run containing two channels of EMG, audio oscillogram and a rectified integrated version of the audio.

The nature of the design required that amplification be identical on all EMG channels. The system was calibrated and adjusted before the experiment and checked again after the experiment was over. The amplification and time relationships of the Mingograf channel were checked repeatedly throughout run-out of the data.

Measurements were carried out manually from the Mingograf traces. Peak amplitudes were measured in arbitrary linear units from an arbitrary baseline, set slightly higher than the amplitude of the inter-utterance noise. It was observed that the baseline remained stable on all channels throughout the experiment. lo tokens were measured for each type. In the case of the rounded vowel in cool, however, it was observed that on a number of occasions the derived amplitude curve contained two peaks. These were treated as separate items of data for the purposes of this experiment. The number of measurements taken for cool, 15, is therefore higher than in the case of stops. First and second peaks for any token of cool are distinguished in Tables 2-7 as (a) and (b) after the number of the token.

3. Objectives and hypotheses

The specific objectives and hypotheses of this experiment can be summed up under the following heads:

3.1. Dependence of Peak Amplitude on electrode placement

The basis for deriving hypotheses here should be an account of how response can be graded along the length of the upper lip. Such an account appears impossible to give. Certainly, the variety of labial articulations— "a speech articulation system in miniature" (Öhman, Leanderson & Persson 1965)— would suggest that the sphincteric function commonly ascribed to the orbicularis oris muscle cannot be the only mode of operation of this muscle in speech. From the anatomical point of view, the picture seems only generally established. According to Hollinshead (1968, p. 336) it is not known to what extent true sphincteric fibres exist. Sicker, DuBrul & Lloyd (1970 p. 143-144) note that the fibres of the muscle can be divided into an upper and a lower group which cross each other lateral to the corner of the mouth. This division appears to correspond to the observation in speech research that EMG signals from upper and lower lips in bilabial stops reveal a different mode of operation, and in labial fricatives a different degree of involvement. In addition, however, it is noted by the above authors that the majority of upper and lower fibres are confined to one side only,

TABLE 2

Raw amplitude scores (arbitrary linear units)

2.1 PACK

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	34.5	27	37.5	23
2.	5o	34	6o	4o
3.	38	25	33	28
4.	42	32	4o	31
5.	45.5	29	49	32
6.	39	25.5	51.5	28
7.	4o	28.5	45	36.5
8.	32.5	23	34.5	26.5
9.	35	21	37	21.5
1o.	41	2o	46	33

2.2 CAP

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	34	27	33.5	28
2.	43	32	4o	29
3.	37.5	31.5	42	32
4.	36.5	23	32	21
5.	38	37	42.5	32.5
6.	39	38	44	25
7.	5o	21	39	33
8.	31	19.5	4o.5	31
9.	4o.5	32	4o	27
1o.	25	14	27.5	14.5

TABLE 2 - continued

2.3 BACK

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	43.5	22	43.5	28
2.	46	31	50.5	35
3.	42.5	22	37.5	28.5
4.	52.5	32	38.5	37.5
5.	37.5	19.5	45	30
6.	43	37.5	44	28
7.	41.5	29	39	36
8.	31.5	22.5	40.5	32
9.	42	32	49.5	36
10.	29	21	30	16

2.4 CAB

1.	32.5	19.5	39.5	24.5
2.	39.5	30	42	35.5
3.	45.5	34.5	48.5	39.5
4.	36	17	46.5	22.5
5.	38	33.5	32	18.5
6.	42.5	33.5	39.5	22.5
7.	50	37.5	54.5	33
8.	20.5	15	35	18.5
9.	40	20.5	47.5	30
10.	29	15	31	16.5

TABLE 2 - continued

2.5 MACK

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	38	21	45.5	25.5
2.	38.5	31	49	31
3.	37.5	24.5	48.5	33
4.	35.5	22	37	25
5.	37.5	27.5	39.5	25
6.	31.5	28.5	37.5	23
7.	51.5	30	41	41
8.	35	22.5	40.5	29.5
9.	37.5	24.5	33.5	23
10.	38.5	25	31.5	26.5

2.6 CAM

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	43.5	30	40	34
2.	42.5	17	47	28.5
3.	30.5	22	38.5	24
4.	33	22.5	39	20.5
5.	58	35.5	43	36
6.	47	32.5	50	33
7.	43	38	41.5	31.5
8.	41.5	29.5	50	30
9.	35	23	42	29
10.	53	29	48	31.5

TABLE 2 - continued

2.7 COOL

No.	R. Centre	R. Corner	L. Centre	L. Corner
1.	2o	18	24	23
2a.	15.5	14	16	17
2b.	13	13	16	19
3.	32	2o.5	36	28.5
4.	25.5	2o	3o.5	29.5
5a.	38.5	17	45	25.5
5b.	27	24	32.5	2o.5
6a.	37	23.5	39.5	31.5
6b.	19	14	18.5	24
7a.	31.5	13.5	36.5	26
7b.	18	12	16	25
8a.	28	17	33.5	32.5
8b.	26.5	15	24	2o.5
9.	17	11	15.5	14.5
1o.	27	18	22	24.5

TABLE 3

Tests of difference among parameters of same type: Mann-Whitney U.

Type	A-B		B-D		C-D		A-C	
	U	Sign	U	Sign	U	Sign	U	Sign
Pack	3	.001	33	NS	7	.001	42	NS
Cap	16.5	.01	50	NS	8.5	.001	42	NS
Back	8	.001	34.5	NS	6	.001	46	NS
Cab	15.5	.01	46.5	NS	9	.001	38	NS
Mac	0	Abs.	33.5	NS	7.5	.001	36	NS
Cam	7	.001	41	NS	0	Abs.	46.5	NS
Cool	40	.001	27.5	.001	98.5	NS	102.5	NS

where: A = Right Centre C = Left Centre

B = Right Corner D = Left Corner

All significance tests one-tailed.

TABLE 4

Kendall Coefficient of Concordance: W.

Type	N	k	$s(=\sum(R_j - \frac{\sum R_j}{N})^2)$	$W(=\frac{s}{\frac{1}{12}k^2(N^3 - N)})$	$X^2(=k(N-1)W)$	Sign. (df=N-1)
Pack	10	4	973.5	.738	26.568	.01
Cap	10	4	756	.572	20.592	.02
Back	10	4	689	.522	18.792	.05
Cab	10	4	1019	.772	27.936	.001
Mac	10	4	703.5	.533	19.188	.05
Cam	10	4	910	.689	24.8	.01
Cool	15	4	3408	.761	42.616	.001

TABLE 5

Spearman Rank Correlation Coefficient: r_s.

Type	A-B		B-D		C-D		A-C	
	r_s	sign.	r_s	sign.	r_s	sign.	r_s	sign.
Pack	.646	.01	.51	NS	.676	.05	.758	.01
Cap	.549	NS	.118	NS	.445	NS	.33	NS
Back	.564	.05	.491	NS	.273	NS	.273	NS
Cab	.92	.0005	.639	.05	.855	.005	.743	.05
Mac	.388	NS	.316	NS	.655	.05	.288	NS
Cam	.661	.05	.828	.0005	.388	NS	.541	NS
Cool	.617	.01	.46	.05	.737	.005	.945	.0005

where: A = Right Centre C = Left Centre

B = Right Corner D = Left Corner

All significance tests 1-tailed.

TABLE 6

Tests of differences across types.
Mann-Whitney U.

Types	A-A		B-B		C-C		D-D	
	U	Sign.	U	Sign.	U	Sign.	U	Sign.
PACK/CAP	38.5	NS	45.5	NS	34	NS	41	NS
BACK/CAB	36	NS	44	NS	50	NS	33.5	NS
MAC/CAM	31.5	NS	38.5	NS	30	NS	36.5	NS
PACK/BACK	39.5	NS	50	NS	48	NS	42.5	NS
PACK/MAC	36	NS	43.5	NS	41.5	NS	38.5	NS
BACK/MAC	31	NS	46.5	NS	42	NS	31	NS
CAP/CAB	48	NS	44.5	NS	39	NS	44.5	NS
CAP/CAM	31.5	NS	49	NS	25	.05	37	NS
CAB/CAM	32.5	NS	42	NS	39.5	NS	35	NS

where: A = Right Centre; C = Left Centre

B = Right Corner; D = Left Corner

All significance tests one-tailed.

TABLE 7

(See text, section 5.2)

Type	Nx, where $x > 32$	% contrib. of error score to s	Nx, where $y < 12$	% contrib. of error score to s
Pack	2	34.7	1	33.3
Cap	1	42.9	1	16
Back	1	42	1	17.6
Cab	2	39.9	2	47.2
Mac	1	15.7	2	57.7
Cam	2	49.7	1	18.6

interlacing at the midline with the fibres of the other side. The deepest fibres of the muscle, however, cross from side to side without interruption. We should expect bilateral operation in speech, and hence can attach little functional significance to this arrangement. Some authors are content to observe that the bundles of the orbicularis oris corresponding to the red zone of the lips are finer and more densely arranged than are those in the periphery, while Hamilton (1958 p. 165) suggests a further functional division. He writes:

> "The upper and lower lips are provided with a series of small muscular fasciculi which act as labial tractors, i.e. they tighten the lips...."

Such an observation does not receive any support that I am aware of in the literature, and seems an uncertain basis on which to build an account of graded response.

In the absence of such an account, therefore, a specific hypothesis about the nature of the relationship between electrode siting and amplitude of derived waveform was not possible. The experience of many researchers that the highest amplitude traces were to be expected near to the mid-line of the lip led to the expectation that the more laterally placed electrodes would exhibit lower amplitude traces.

3.2. The reliability of peak amplitude measures

It was suggested that the following two conditions should be taken as measures of reliability:

(i) That even if the absolute amplitudes of the peaks from the various positions should vary, there should nonetheless be a high overall correlation among the amplitude scores paralleled by high correlations between pairs of electrodes. Failure of correlation measures to reach acceptable standards of significance would be interpreted as suggesting that events at the electrode sites were to a degree independent, and that hence the measure as a whole was unreliable.

(ii) That each output channel should show the same measure of discrimination on peak amplitude measures as between different linguistic types. Thus, if it is possible to discriminate between a syllable-initial and syllable-final stop by means of the signal from one electrode site, then, if the measures are to be held reliable, this discrimination should also be possible at all other electrode sites. It was predicted that a set of utterances which could be shown to be deviant on this measure would either exhibit such deviance through absolute unreliability, or because different parts of the lip were being used only partially differently in the two linguistic types. That is, the behaviour of the centre of the lips in two linguistic types may be different while the behaviour of the corners of the lips may be the same. In order that the latter explanation be entertained it is suggested that the deviance should be exhibited bilaterally in any given case.

4. Results

Two factors led to the choice of non-parametric tests for the statistical analysis of data. In the first place, experience with the use of parametric tests had led to the conclusion that the very large standard deviations encountered suggested that the assumptions about the distribution of the data required by these tests were not being met. Secondly, in the present design, the multiplication of electrode sites required that the number of utterances be kept fairly small. The choice does not appear to limit the validity of the results, however, since, according to Siegel (1956) the use of non-parametric tests increases the generality of the findings.

The tests employed here are the Mann-Whitney test for differences between two independent samples, the Kendall Coefficient of Concordance, a measure of agreement between results where the parameters are greater in number than two, and the Spearman Rank Correlation Coefficient. For the case of r_s, significance was determined by using a table of critical values, with extreme and doubtful cases calculated according to

$$t = r_s \sqrt{\frac{N-2}{1-r_s^2}}$$

The raw amplitude scores, expressed in arbitrary linear units, appear in Tables 2.1 - 2.7 The results obtained from statistical tests on these scores are given in Tables 3 - 7.

4.1. Dependence of amplitude scores on electrode placement

The first set of results relate to the dependence of the amplitude scores on electrode placement. This was investigated by looking for significant differences between the scores from amplitude sites within types. The results are presented in Table 3. H_o in this case was that the scores from different sites did not differ. Taking the stop series first, it will be seen that in every case H_o could not be rejected at a satisfactory level of significance for the comparisons between amplitude scores for centre sites, and between amplitude scores for corner sites. On the other hand, it will be seen that for each stop also, H_o can be rejected at a high level of confidence (above .01 in every case) for the comparisons between right center and right corner, and between left centre and left corner. Indeed, it will be seen that U, a measure which can vary from 0 to

50 in this case, achieves absolute significance in one case for each of these sets of comparisons.

In the case of the rounded vowel, the case is rather different. Here there are no grounds for rejecting H_o in the comparison between left centre and left corner, and between right centre and left centre, whereas we can reject H_o at a high level of significance for the comparisons between right centre and right corner, and right corner and left corner.

4.2. Results relating to first criterion of reliability

Tables 4 and 5 give results relating to the first criterion of reliability, given in 3.2. (i) above. The results for the Kendall Coefficient of Concordance (Table 4) show significant correlations (above .05 level) for all electrode positions for all types. Table 5, results of the Spearman Rank Correlation test, shows the nature of these correlations in greater detail.

There is a distinct lack of agreement between these two measures in the case of <u>cap</u>, which according to the Kendall test has a W significant at the .02 level, whereas it demonstrates no significant correlations among its electrode positions on the Spearman tests. This discrepancy appears to be explained by suggesting that for the Kendall test the value of <u>s</u> was inflated by the fact that the amplitude scores on all electrode positions for <u>cap 10</u> (see Table 2.2.) were uniformly ranked lowest at each site. The conjunction of lowest rankings gave rise to a large difference score.

From the stop sequence there appear to be few regularities to be extracted on the Spearman tests. The exception is the case of <u>cap</u> vs. <u>cab</u>: while the <u>cab</u> scores provide grounds for rejecting H_o at a significant level for all four sets of comparisons, in the case of cap there are no grounds for rejecting H_o for any of the sets of comparisons.

Pack and back share a lack of correlation between contralateral corner electrode sites, and a significant correlation between right centre and right corner sites. They are distinguished by the fact that back, unlike pack, lacks a significant correlation between the two left sites, and between the two centre sites.

Cam, which might be expected to follow cab, is distinguished from it by having an extremely significant correlation between the scores on corner sites, and not showing a significant correlation between centre sites, or between sites on the left side. By the same token, mac is distinguished from back by showing a correlation between left-hand sites whereas it does not show a correlation between right-hand sites where back does.

Cap and pack share a lack of correlation on the comparison between corner sites. On the other comparisons cap shows no significant correlation while pack shows varying degrees of significance.

Cab and back share a correlation between right-hand sites, although the significance is higher in the case of cab. On other comparisons, however, they are distinguished by degrees of significance on the part of cab and lack of significance on the part of back.

Cam and mac share a lack of significant correlation between centre sites. They are distinguished, however, on the comparison between left sites by the fact that cam shows no significant correlation while mac exhibits a correlation significant at the .05 level. On other comparisons, mac shows no significance, while cam does.

In the case of the rounded vowel in cool, on the other hand, there is consistently significant correlation among the sites, with the highest being between the centre sites and the lowest being the corner sites.

4.3. Results relating to second criterion of reliability

The second criterion of reliability, given in 3.2. (ii), was investigated by comparing scores from identical electrode sites across linguistic types, utilizing once again Mann-Whitney U tests. In all cases except one no significant differences could be found across types. The single exception occurs at the left centre site for the comparison between cap and cam. Here H_o could be rejected at the .05 level of confidence.

5. Discussion

The results will be discussed in the order in which they were presented.

5.1. Dependence on electrode position

The results relating to the dependence of amplitude scores on electrode position are initially very clear. During each of the stops the lips appear to operate symmetrically, since there is no means of distinguishing between sets of signals from right centre and left centre, and between right corner and left corner. On the other hand, in every case for the stop sequence the amplitudes from the centre sites were found to be significantly greater than those from the corner sites. There are two possible explanations for this latter finding:

(i) That the signals on the corner sites are related to those on the centre sites by volume conduction of potentials from the centre sites to the corner sites. (See Dedo & Dunker, 1966).

(ii) That evidence has indeed been given for the operation of a mechanism which grades response along the length of the upper lip, with maximum amplitude concentrated at the centre of the lips for labial stop consonants. It is maintained that there is sufficient evidence contained in the present experimental design to enable us to choose the latter explanation with some degree of confidence. In the first place it should be noted that the volume conduction hypothesis includes a more extreme form of the grading hypothesis - namely that there is little or no activity at the corner sites, the derived signal having been transmitted from other sites. Further, in the case of the volume conduction hypothesis, we should expect there to be a consistently high degree of correlation between centre and corner sites. It is evident from Table 5 that this is not in fact the case.

Acceptance of explanation (ii), however, must involve leaving open the question of the nature of the mechanism involved. The functional basis for the mechanism seems clear - the centre of the lips has further to travel to achieve closure than do the corners. The physiological basis, though, as implied by Section 3.1., is extremely problematical. I am not aware of any work relating to the existence and nature of an "innervation zone" in the lips such as is found in skeletal muscles and has been reported by Sonesson to be the case in the <u>vocalis</u> muscle (Sonesson 1960 p. 24). Hence the possibility of selective involvement of motor units in different types of operation cannot be discussed.

5.1.1. The possibility of artefacts

It appears possible to explain the asymmetries in the results in Table 3 for cool relatively simply as an artefact of electrode placement. This explanation further has the merit of pointing to a possible difference in active lip region between stop consonants and rounded vowels.

It will be remembered that while on the right side the significant difference between centre and corner sites familiar from the stop cases was also found for cool, this difference was not observed on the left side. Further, while a lack of difference was found between the centre sites, again following the stop case, a difference was found at a high level of significance between the two corner sites. Now, suppose that the left corner electrode had been placed further towards the midline than the right corner electrode. Let us further suppose, following Section 5.1.1., that the amplitude of the signal is graded in some way from highest at the centre to lowest at the corner, then the asymmetries in the data would be explained, in that the amplitudes of left corner and left centre sites would become more similar, while at the same time the amplitudes from the left corner site would become dissimilar to the amplitude scores recorded contralaterally.

This explanation, however, begs the question of why such an artefact of electrode placement does not show up in the stop data. To resolve this question, it is possible to make the speculation, given that there are, crudely, "regions of high activity" and "regions of low activity" along the length of the upper lip, that the region of high activity is wider for the rounded vowel than for the stop consonants, and hence included the asymmetrical left corner electrode. This electrode, in common with the right corner electrode would be supposed under this hypothesis to be outside the high activity region in the case of stops. The present study shows no direct evidence for or against such a hypothesis. As a first approximation to distinguishing between the form of the labial articulations in the two cases of closure and rounding, however, the hypothesis has a certain plausibility.

With the possible exception, then, of the rounded vowel, the results have shown that the peak amplitude of derived EMG waveforms _is_ dependent upon electrode position. The nearer the midline the electrode is placed, the larger the signal can be expected to be. This relationship is found on both sides of the lips, and confirms the observations of researchers seeking optimum positions for recording lip EMG. Further, as expected, the lips act in a symmetrical fashion, there being for the stop group no significant differences between the sets of amplitude scores from centre sites, or from corner sites.

5.2. Correlation measures and the overshoot hypothesis

As will be apparent from the presentation of results in Section 4.2., the Spearman Rank Tests appear to give a more detailed account of the manner in which the reliability criterion of 3.2.1. is met than do the Kendall tests. The experience with the discrepancy between _cap_ on the two tests reported in 4.2., however, suggests a means by which the Kendall tests in this case can be applied directly to the problem of the explanation of EMG variation as stated in Section 1 of this paper.

Computation of _W_ in such tests involves ranking the scores for each electrode site separately within a linguistic type, and then, for each token, computing the sum of the ranks given the scores on each of the sites. The mean of these sums of ranks for all the tokens of a type is computed, and the difference between the total of ranks for each token and the mean is found. These quantities are squared and the squares summed to give s. Generally speaking, the magnitude of W, and hence the significance of W, grows as s increases in size for a constant number of parameters and number of tokens within the parameters.

The largest contributions to s come from cases where, to illustrate from the case in hand, the four electrode sites are agreed in ranking scores for a particular token either high or low. In both these cases the difference between the sum of ranks and the mean of the sums of ranks will be great, and the contribution to s large.

A theory of EMG variation which supposed that there was a region of overshoot, within which amplitude scores were variable, would presumably predict that individual sites would more often agree in ranking amplitude scores for a particular token low than high, since the highest scores must be within the region of overshoot, whereas the lowest scores almost by definition will not. Therefore the "overshoot" hypothesis predicts greater variability at the high end of the sets of amplitude scores, and greater homogeneity at the lower end.

In order to test whether this prediction is borne out by the facts, the following crude test was devised. Conventionally, ranks are assigned with the highest number for the lowest rank. In the case that all four electrode sites were agreed in assigning the amplitude scores for a particular token the lowest rank, the sum of ranks will be 40. If, on the other hand, they assign them the highest rank the sum of ranks will be 4. The difference between these scores and the mean (22, i.e.

$$\frac{4\sum_{i=1}^{10} i}{10}$$

in our case) will be +18 and -18 respectively. The sign difference will disappear when the difference scores are squared and the same contribution will be made to $\underline{s}$ in each case. If there are differences between the tendency for high amplitudes and low amplitudes to exhibit greater and less variance, then this should be revealed by:

(i) the number of sums of ranks which are of small magnitude compared with the number which are of large magnitude.

(ii) The contribution which the squared difference scores related to these sums of ranks make to the magnitude of $\underline{s}$ expressed as percentages.

Such measures were derived from the stop set and are presented in Table 7. Arbitrary limits were set for the upper and lower ends of the scale of magnitude for sums of ranks, such that at the lower end the sum of ranks should be greater than 32, and at the upper end should be less than 12. In the case of perfect correlation this would limit the cases considered to those where all scores were ranked 1 or 2, or 9 or 1o. Columns 1 and 3 of Table 7 show the number of times in each type that sum of ranks scores falling within the lower and upper limits respectively occurred. Columns 2 and 4 show the percentage contribution made to $\underline{s}$ by the squared difference scores associated with these sums of ranks scores.

There does not appear to be much information to be gleaned from the statement of numbers of sums of ranks scores falling in the two limits. In two cases the number in the lower limit is greater than that in the higher limit, the two contain equal numbers on three occasions, and the higher limit contains more occurrences than the lower in one case. A Mann-Whitney U test was applied to the percentage scores, however. A value of 15 was computed for U, which value has an associated significance of .35. Hence the test fails to find a significant difference between the two sets of percentages.

We may conclude, therefore, that insofaras the test described captures predictions under the "overshoot" hypothesis of Section 1, it does not provide any evidence in favour of it.

5.3. The first criterion of reliability

For the stop sequence, except for the single case of cap vs. cab, the linguistic categories with which the stimulus set was constructed, fail to provide any key to the irregularities revealed in Table 5, presenting results of Spearman Rank Correlation tests on the amplitude data. One is prevented from an automatic adoption of the view that EMG amplitude measures are unreliable by the fact that high and consistent correlations are observed for the case of cool. A saving hypothesis could be constructed by supposing that the signals were more variable in the case of the stop sequence. Now, amplitude measures for the stop sequence are of generally higher amplitude than for the rounded vowel (for example, the range for right centre electrode for cool is 13 - 38.5, while for the signals from the right centre electrode for the stop sequence taken together the range is 2o.5 - 53). Experience suggests further that the durations of the EMG signals will be shorter in the case of stop consonant gestures. It might be suggested that EMG measures for fast, high-amplitude gestures might be more variable than for slow gestures with lower amplitudes, such differences perhaps reflecting a difference in control.

It must be remembered, however, that Mansell (1970b) showed both amplitude and duration measurements for [u] in the frame [hək _ tə] to be highly variable, and to show only a small positive correlation with each other.

In the absence of a satisfactory criterion for explaining the asymmetries of Table 5, and lacking satisfactory evidence for a saving hypothesis, it must be provisionally concluded that the reliability criterion of 3.2. (i) is not met. The reservation must be included, however, that this conclusion applies only to the stop series.

5.4. The 2nd criterion of reliability

Table 6 shows that the second criterion of reliability, given in 3.2. (ii) is met in all except one case. Further, since in this case, the comparison between amplitude scores at left centre electrode for cap and cam, the variance is not bilateral, we must conclude that this is a case of absolute unreliability. More evidence comes from a comparison between the patterns for cap and cam on Tables 3 and 5. In both cases the relationships for the left centre site and the right centre site and the left centre site and the left corner site show the same patterns of significance and lack of it. Although this result is at a relatively low level of significance and is found in the context of 35 satisfactory scores, I take it to demonstrate that there will be cases where unfortunate placement of an electrode may led to distictions between linguistic types on the basis of peak amplitude measures which will not reflect the activity of the lip as a whole. It must be noted that the rounded vowel in cool was not included in the tests reported in this section.

5.6. The cause of variation

Finally, since the reliability of the amplitude measures has not been shown, it must be noted that the possibility of a peripheral cause for signal variation is not ruled out. Further it is not possible on the basis of this experiment to distinguish between variation arising from factors at the electrode-tissue interface and variation in the events taking place beneath the electrode. If the former set of factors were responsible for variation, however, it is rather surprising that this study has shown that, with two exceptions, the range of variations from comparable parts of the lips is the same. In the case of the latter cause of variation, the explanatory problem that was the subject of Mansell (1970b) remains.

6. Conclusions

This study has investigated the dependence of EMG peak amplitude measures on electrode position, and the reliability of such measures as revealed by the degree of correlation among scores from electrode sites in a Common Reference derivation, and the degree of constancy across these sites in distinguishing between peak amplitude activity for different linguistic types. It was further possible to provide a crude measure of the plausibility of an "overshoot" hypothesis to account for EMG variation.

It was concluded that EMG peak amplitude scores are highly dependent on electrode positioning, with the highest amplitude being found closest to the midline of the lips for both stop consonants and spread vowel. The EMG peak amplitude measures for the stop consonants were judged unreliable, since the patterns of correlations among sites were complex, and not explicable in terms of the linguistic oppositions employed in deriving the stimulus array. Further it was possible in one case to make a presumably spurious distinction between linguistic types at one site, whereas the other sites in this case showed no difference in amplitude scores. It must be noted that the rounded vowel was not studied in the latter test, and in the former test of reliability proved to meet the criteria adequately.

The crude test of the overshoot hypothesis failed to find confirmatory evidence.

Finally, it is concluded that the degree to which lip activity as a whole can be studied in detail as a result of its use is sufficient proof of the usefulness of Common Reference deployment of electrodes in speech EMG research.

Acknowledgements:

The experimental work on which this paper is based was carried out in the Speech Research Laboratory of the Language Centre, University of Essex. I should like to thank Dr. M.A.A. Tatham for his help and cooperation in the running of the experiment. The processing of the data was carried out while the author was a visiting researcher at the Institute of Phonetics, University of Copenhagen, supported by a grant from the Rask-Ørsted Foundation. I should like to thank Professor Eli Fischer-Jørgensen and the staff of the Institute, especially Børge Frøkjær-Jensen, for placing their facilities at my disposal.

References

Cooper, R., Ossleton, J.W. and Shaw, J.C. 1969: EEG Technology. Butterworth.

Dedo, H. and Dunker E. 1966: "The Volume Conduction of Motor Unit Potentials", Electroenceph. Clin. Neurol. 2o, p. 6o8-613.

Hamilton, W.J. 1958: Textbook of Human Anatomy. Macmillan.

Hollinshead, W.A. 1968: Anatomy for Surgeons. Vol. I. The Head and Neck. 2nd. ed.

Mansell, P. 197oa: "Linguistic Parameters in Performance Models". University of Essex, Language Centre Occasional Papers, 8.

" 197ob: "On the Nature of EMG Variations". Proc. Essex Symposium on Speech Production Models, UELC Occasional Papers, 9.

Ohman, S., Leanderson, R. and Persson, A. 1965: "Electromyographic Studies of Facial Muscles During Speech", STL-QPSR 1/1966, p. 1-14.

Sicker, H., DuBrul and Lloyd, E. 197o: Oral Anatomy. 5th ed. C.V. Mosby.

Siegel, S 1956: *Nonparametric Statistics for the Behavioural Sciences.* Mcgraw Hill.

Sonesson, B. 1960: "On the Anatomy and Vibratory Pattern of the Human Vocal Folds", *Acta Oto-Laryng.* Suppl. 156.

SYLLABLE AND ACCENT IN JAPANESE[1]

Hideo Mase

1. Introduction

Such place-names as Yokohama, Tookyoo and Nippon have roughly the same length when spoken. Each word is written in the same number of (i.e. four) Japanese letters or syllabic signs (except kyo in Tookyoo, in which one has a device of a two-letter combination in order to show a palatalized, non-syllabic y between k and o). Each word is said to consist of the same number of "mora"s, i.e. of four moras.

According to one theory, the prevailing one in Japan, the mora equals a phonological syllable, because all the moras have almost the same length, because they are units of rhythm and of metre, and because a toneme (which is phonetically manifested in relative height of pitch) is put on any of the moras. Thus, each of the above words has four phonological syllables: Yo-ko-ha-ma, To-o-kyo-o, and Ni-p-po-n (where - shows a syllable boundary). This is one of the two major theories concerning the syllable and the accent, and is called the "pitch-level theory". Some foreign linguists hold a similar, or, rather, virtually the same opinion, saying that the mora is the only syllable which is found in Japanese, but differing as to details, claiming either that it is rather a "phonetic" syllable (Bloch 1950 §6.4); or that the mora - a unit of length -

---

1) I am very much obliged to Niels Ege for valuable comments on the first draft of this paper.

is the only "phonemic" syllable (Hockett 1955 §2214; 1958 §11.7) - based on Bloch's description; or that the acoustic phonetic data show that all the moras have virtually the same length, which fact shows the distinctive function of a mora as an "emic" syllable (Han 1962 p. 63f; Weitzman 197o p. 6-7).

The other opinion advocated in Japan is that the mora is indeed a phonological "unit", but not a phonological "syllable", since the phonological behaviours or functions of moras are different according to their distribution and accentual function. Some moras are "free" moras which can occur in any position in the utterance, for example yo, ko, ha and ma, while others are "bound" moras which can occur only after a free mora in the utterance, for example the second component of the long vowel (the second o's in too and kyoo in the word Tookyoo), p between i and p in the word Nippon, and n in the same word. Further, only a free mora can be an accent-bearer, i.e. a mora which can be (phonetically) higher pitched immediately before a lower pitched mora. Here, the tonemes are not considered to be distinctive accent units. This latter theory is Hattori's and is known as the "accent-kernel" theory. In this theory a phonological syllable is defined as consisting of either a free mora or a free mora followed by a bound mora. Thus, in this theory the above-mentioned words are divided into syllables as follows: Yo-ko-ha-ma, Too-kyoo, and Nip-pon.

Pike (1948 p. 14 fn. 29) and Garde (1968 p. 136-137) hold a similar opinion on the accent to Hattori's. McCawley (1968) sets up a similar theory using the approach of generative phonology. Interestingly enough, the mora is here taken up as a phonological unit from another point of view. Usually one has relied mainly on phonetic phenomena to establish the mora as a functional unit, but McCawley describes the mora essentially as a unit on the morphophonemic level, which makes it possible to apply in an ordered way various accent manifestation rules such as "accent-deletion" and "pre-accentuation" rules.

In the following pages I will consider only the standard dialect (or standard Japanese), since almost all the works on Japanese phonology by foreign scholars are restricted to the description of the standard dialect. It should be mentioned, however, that not all the features in the standard dialect are the same in other dialects. The standard dialect is one of the simplified dialects, as far as the accent system is concerned, so that it is not at all the most representative dialect. Some features found in other dialects will be briefly mentioned at the end of this paper.

2. The pitch-level theory

In this theory a mora is considered to be a phonological syllable,[2] since a mora functions as a unit of rhythm and of metre and functions as the minimal unit on which a toneme is put.[3] There are two distinctive tonemes: /high/ and /low/. Their syntagmatic functions in single-word utterances are shown below. Japanese ac-

---

2) Hereafter, "syllable" in my text refers to a "phonological" syllable, when there can be no confusion, while a "phonetic" syllable is always referred to by me in its full form.

3) The following description is based on the articles of Haruhiko Kindaichi which are collected in his book from 1967 (= Kindaichi 1967): 1947 = 1967 p. 269-289; 1950a = 1967 p. 133-153; 1950b = 1967 p. 299-307; 1957 = 1967 p. 198-230; 1960 = 1967 p. 233-268; 1963 = 1967 p. 113-132; 1965a = 1967 p. 367-391; 1965b = 1967 p. 40-57; and 1967 = 1967 p. 58-77.

But, as I mentioned in the Introduction, the pitch-level theory is the one prevalent in Japan, so that Kindaichi is not the only advocate of the opinion.

cent is essentially a word-accent. Examples are:[4]

1,a) /$\overline{\text{ka}}\underline{\text{ki}}$/ ('oyster')
1,b) /$\underline{\text{ka}}\overline{\text{ki}}$/ ('persimmon')('fence')

2,a) /$\overline{\text{ha}}\underline{\text{na}}$/ (a girl's name)
2,b) /$\underline{\text{ha}}\overline{\text{na}}$/ ('nose') ('flower')

3,a) /$\overline{\text{o}}\underline{\text{n}}$/ ('favour')
3,b) /$\underline{\text{o}}\overline{\text{n}}$/ ('sound')

4,a) /$\overline{\text{ho}}\underline{\text{o}}$/ ('direction')
4,b) /$\underline{\text{ho}}\overline{\text{o}}$/ ('law, ordinance')

5,a) /$\overline{\text{sa}}\underline{\text{too}}$/ ('the opposition party')
5,b) /$\underline{\text{sa}}\overline{\text{too}}$/ ('gradation')
5,c) /$\underline{\text{sa}}\overline{\text{to}}\underline{\text{o}}$/ ('sugar')

Here, in each pair or set, the words have the same segmental constituents. Permutation of tonemes makes paired words different words. Two tonemes are said to be necessary and sufficient in describing the accent system of the standard dialect. Kindaichi claims that this small number of tonemes (i.e. two in number) can describe not only the accent system of this dialect, but also those of other dialects in Japanese, and, further, is valid to a diachronic description.[5] This means that a description applying two tonemes is economical and simple.[6]

The segmental structure of the mora, i.e. of the syllable, is one of the following four types, of which the first is the most general one but the other three, especially the last two, are more or less typical in Japanese.[7]

---

4) $\overline{\quad}$ = the high toneme, ___ = the low toneme. The notation of segments is a phonemic one, but a loose one. As a consequence it is almost the same as the orthography by means of the Latin alphabet. Problems of segmental phonemes are not discussed in this paper.
5) Kindaichi 1967 p. 253.
6) Ibid. p. 252ff.
7) Ibid. p. 116-117.

1) /(C)(S)V/ (C = consonant, S = semivowel, V = vowel)

2) /N/ a nasal consonant of one-mora length

3) /Q/ the first component of a geminate consonant, which has one-mora length

4) /:/[8] a phoneme which prolongs the preceding vowel by one-mora length, or in other words, the second component of a sequence of two identical vowels.

Details are discussed in 4. and the following sections.

3. Syllable in the accent-kernel theory

It is Hattori[9] who first introduced in Japanese phonology a phonological syllable not equal to a mora. First, he explains how moras and phonetic syllables relate to each other:

> "In the phonological interpretation of the pronunciation of the Japanese language, it is necessary to assume a phonological unit 'mora' which corresponds roughly but not exactly to the phonetic unit 'syllable'. While a syllable which ends with a short vowel corresponds to one mora,

---

8) Kindaichi uses the symbol "-" instead of ":".

9) I am not sure when Hattori first published his idea of a phonological syllable. As far as I can trace back, the idea is seen in his article from 1949, "'Bunsetsu' to akusento" ("On the accentual unit"), Hoogen to Minzoku Nos. 3 & 4, 1949 = Gengogaku no hoohoo ("Methods in linguistics"), Tokyo, 196o, p. 428-446. This summary of his ideas is based on his articles from 1949 = 196o p. 428-446, 195o = 196o p. 751-763, 1954 = 196o p. 24o-272, 1958 = 196o p. 36o-364, the article from 1961, and Onseigaku ("Phonetics"), Tokyo, 1951. Articles from 1949, 195o, 1954 and 1958 are collected in his book from 196o, with some but often important corrections and additions, so that I refer to these articles as published in the book. But except for very crucial points, details of page number etc. are omitted.
Examples are in most cases my own.

> a syllable which ends with a long vowel corresponds to two morae. The so-called 'haneru on' ('syllabic nasal') and 'tsumaru on' (choked sound) correspond to one mora." (1960 p.751).

> "/CVN/ [N = "syllabic nasal", H.M.] corresponds usually to one syllable. ... /Q/ ["choked sound", H.M.] stands only after a vocalic phoneme and before /k, t, p, s, c/. The consonantal phonemes which follow /Q/ correspond to geminated consonants, while /Q/ itself corresponds to the laryngeal tension during the first half of the geminated consonants. /CVQ/ corresponds usually to one syllable." (Ibid. p. 753)

A "syllable" in the above citation is of course a "phonetic" syllable. When Hattori says that /CVN/ and /CVQ/ usually correspond to one phonetic syllable, he means that any mora can be pronounced as an independent phonetic syllable, when spoken very slowly (1960 p. 361 and elsewhere).[10] The Japanese word /wan/ (with high-low pitch) ('bay'), for example, can be pronounced as [wa-n·], i.e. in two phonetic syllables, while the English /wʌn/ ('one'), for example, cannot be pronounced in two phonetic syllables. Hattori defines a phonetic syllable by acoustic or/and auditive sonority and articulatory explosion-implosion. A two-mora word like kaki is pronounced in two phonetic syllables [ka] and [ki],[11] while two-mora words like hoo and on are usually pronounced in one phonetic syllable, [ho:] and [kan] , respectively (1951 p. 169).

Now, what is a phonological syllable? There are "free" and "bound" moras.[12] A free mora is one which consists of a

---

10) Hattori's terminology "syllabic" nasal is not suitable. It should be a 'postvocalic nasal of one-mora length'. Or else, the above statement of his does not make sense.
11) All phonetic transcription is very broad.
12) The expressions "free" and "bound moras" are adopted from Martin (1967 p. 246 fn.2).

consonant and a vowel[13] with or without a semivowel between the two phonemes, while a bound mora is the second component of a long vowel, or a "syllabic nasal", or a "choked sound". A free mora can make a phonological syllable by itself, because it can form an independent phonetic syllable, and because it can occur in any position in the utterance without being followed by a bound mora. On the other hand, a bound mora cannot make by itself a phonological syllable, because it does not usually form one phonetic syllable by itself, and because it always occurs after a free mora. Possible combinations of segmental phonemes in the phonological syllable are:

1) /C(S)V/
2) /C(S)VV/ (where /VV/ should be a sequence of two identical phonemes),
3) /C(S)VN/,
4) /C(S)VQ/.

A bound mora, which is a phonological unit, corresponds neither to a phonological nor a phonetic syllable. The role of a bound mora as a phonological (i.e., functional) unit in the standard dialect is not explicitly mentioned, even if we take the accent into account.[14] Although

---

13) The syllable-initial vowel (or, rather, vocoid) is always preceded by a glottal constriction, if the vocoid is not preceded by a consonant (or, rather, contoid). Therefore, this glottal constriction is interpreted as a consonant phoneme, which is the counterpart of the voiceless /h/. (Cf. Hattori 1961.)

14) In some dialects /N/ and the second vowel of /CVV/ can be accent-bearers. So that all the moras except /Q/ are "free" moras in that they can be accent-bearers, but a "bound" mora in the standard dialect seems to be also a "bound" mora in other dialects (perhaps with a few minor exceptions) in that a bound mora does not usually form a phonetic syllable, and that it cannot occur in utterance-initial position. (Cf. Hattori 196o p.247.)
In a later page in the same article Hattori mentions that the mora is not a syllabic unit universal to all languages, so that we had better reconsider whether it is necessary to set up "mora" as a phonological unit in the description of Japanese (196o p.27o; fn.26 p.27o).

the phonological syllable in Hattori's theory cannot be described sufficiently without taking the accent into account, his phonological syllable, in other words, is essentially the unit which best describes segmental combinations, and which corresponds to a phonetic syllable.

4. Validity of tonemes

Let us take the example words mentioned in 2. again, and describe their accents, by applying the technique of so-called autonomous phonemics.

For the sake of this argument, we assume that the mora is a phonological syllable, so that the following words in each pair have the same segmental constituents and all have the same number of syllables. The paired words are differentiated in meaning by a certain permutation of the pitch function. What is common to the four pairs of words is that all of them have the same number of syllables (=moras), and what is common to (a) or (b) is that they have the same pitch contour.[15] (i) and (ii) are alternate forms.

(I)	Pair 1	Pair 2	Pair 3	Pair 4
(a)	kaki	hana	on	hoo
(b,i)	kaki	hana	on	hoo
(b,ii)	-	-	on	hoo

(a) and (b) are in contrast in any environment, therefore the difference of pitch is invariably "emic". (i) and (ii) in Pairs 3 and 4 are not-contrastive, and it is the form (ii)

15) Whether one perceives the relative pitches on the syllables or perceives the pitch change between the syllables is not discussed here.

which is the most frequent.[16] Then, if we take (ii) as representative of 3 and 4, we get

(II)

	1	2	3	4
(a)	kaki	hana	on	hoo
(b)	kaki	hana	on	hoo

In Pairs 3 and 4 the first syllable in (a) does not contrast to that of (b). The low-high pitch sequence in (1,b) and (2,b) is paralleled by the high-high pitch sequence in (3,b) and (4,b). Then, what is common to all cases in (b) is the high pitch on the second syllable. The pitch of the first syllable in (b) is not contrastive to the high pitch in (a), because the former is either low-pitched as in 1 and 2, or high-pitched as in 3 and 4. When the second syllable is high-pitched, the pitch information of the first syllable in (b) must be redundant, or else we would not get two noncontrastive forms in (3,b) and (4,b). Then, the contrast between (a) and (b) must lie in the different positions of the last high pitch.

---

16) Cf. Hattori 1960 p.246. But see the arguments by Han (1962) and Weitzman (1970). Cf. also Mase (1969 p.144).

Though Kindaichi admits the existence of forms like (3,b,ii) and (4,b,ii), he does not show them in the "phonological" notation, since the even-pitched pronunciation does not occur in 'deliberate' speech (1967 p.367-91, esp. p.373-74). Kawakami (1966) is strongly against Kindaichi, saying that the two forms are just two free variants of one accent pattern. The low-high pitch pattern of (b,i) when it occurs, occurs only phrase-initially, therefore, the low-pitch is interpreted as a kind of intonation. Further, there are many standard dialect speakers who have no low-high pitch pattern for words like (3,b) and (4,b), says Kawakami.

If we mark only the distinctive last high pitch, we can show the pairs of words as:

(III)

	1	2	3	4
(a)	kaki	hana	on	hoo
(b)	kaki	hana	on	hoo

The distinctive last pitch in words (a) is of course the high pitch on the first syllable.

This analysis, however, which is basically that of the pitch-level theory, does not tell the truth. As mentioned in the preceding section, a bound mora cannot be an accent-bearer in this dialect. But in (3,b) and (4,b) above, the accent is put on the bound mora, i.e., on /n/ and the second /o/, respectively.

What the real situation is, becomes clear when we add the particle ga after the noun. This particle indicates roughly the subject case, and has no fixed accent.

(IV)

	1	2	3	4
(a)	kakiga	hanaga	onga	hooga
(b,i)	kakiga	hanaga	onga	hooga
(b,ii)	-	-	onga	hooga
(c)	kakiga	hanaga	-	-

((1,b) 'persimmon', (1,c) 'fence', (2,b) 'nose', (2,c) 'flower'.)

In the above list (IV), the third pattern (c) is not found for Pairs 3 and 4. That is, /n/ in (3,c) and the second /o/ in (4,c) do not occur in the position where the distinctive last high pitch is put.

If we put only the distinctive accent mark in the way shown in (III), we get:

(V)

	1	2	3	4
(a)	kakiga	hanaga	onga	hooga
(b,i)	kakiga	hanaga	onga	hooga
(b,ii)	-	-	onga	hooga
(c)	kakiga	hanaga	-	-

Then, we take off the particle, keeping the accent mark on the noun word.

(VI)

	1	2	3	4
(a)	kaki	hana	on	hoo
(b)	kaki	hana	on	hoo
(c)	kaki	hana	-	-

Notice that the nouns (1,b) and (1,c), and (2,b) and (2,c) have the same phonetic pitch contour, i.e., [kaki] and [hana], respectively (cf. lists (I) and (IV)), but in lists (V) and (VI) they are shown differently. This is a crucial point. The high pitch on the second syllable (= the second mora) in words (b) is not followed by a syllable with low pitch on it (here, one-syllable morpheme ga), while in (c) it is followed by a low-pitched syllable.[17] On the other hand, (a) and (c) of 1 and 2 are distinguished by the position of the last high pitch, or in other words, by the position of the high pitch immediately followed by a low-pitched syllable.

Leaving problems of detail to later pages, we now go to Hattori's definition of the accent.

17) Actually, the pitch on ga depends upon the accent of the foregoing noun. Cf. McCawley 1968 p.138 ff., accent rules for nouns.

5. The accent-kernel theory

Hattori says that in order to describe the accent patterns of a dialect (or a language) phonologically, we should only take up their distinctive features. If we define a high pitch immediately followed or that may be immediately followed by a low pitch as the "accent kernel",[18] then the distinctive features of accent patterns in the standard dialect are:

1. whether there is an accent kernel (in the accent pattern), or not, and
2. if there is, on which mora[19] it is.

The accent pattern which has no accent kernel is usually called "even" or "level" pattern, and that which has an accent kernel is called "uneven" or "falling" pattern, or pattern with a kernel.[2o]

When we show the syllable with accent kernel with the accent mark "⌝" on it, list (VI) above appears in the following way.

(VII)

(a)	ka⌝ki(ga)	ha⌝na(ga)	on⌝(ga)	hoo⌝(ga)
(b)	kaki(ga)	hana(ga)	on(ga)	hoo(ga)
(c)	kaki⌝(ga)	hana⌝(ga)	-	-

That is, the distinctive feature (1) is found in the opposition between (a,c) and (b), and the distinctive feature (2) is found in the contrast between (a) and (c). (NB. 3 and 4 are one-syllable words.)

---

18) The accent kernel in any dialect of Japanese is the high pitch immediately before the low pitch, so that we can say that the Japanese accent is, so to say, a "falling" or "descending" accent. (Hattori 1961 p.8). But cf. McCawley 1968 p.19off.

19) "Mora" should be "syllable". But it is not contradictory. In Hattori's theory a bound mora cannot take an accent kernel, then it is only a free mora in the phonological syllable which can take the accent kernel.

2o) Hattori 196o p.249-51; p.268; p.363-64.

The first mora in the syllable with the kernel has phonetically higher pitch, and often greater intensity and greater duration than all the following. The lower pitch on the first mora in the first syllable in (b) and (c) patterns above is a redundant feature in this dialect, since it is conditioned by the place of occurrence: when the first syllable has no kernel, the first mora is lower-pitched. Both distinctive and redundant features should be realized in the accent pattern in actual speech in order to make it different from other patterns and to make it sound natural.[21]

6. Even (=non-falling) and uneven (=falling) patterns

As mentioned in 5., in the kernel theory the accent patterns of hana ('nose' - (2,b)) and hana ('flower' - (2,c)) are functionally distinguished, even though, it is true, they have phonetically one and the same pitch contour, in any circumstance[22] – whether they occur as single-word utterances or as part of a longer phrase consisting of the word plus a particle.[23] On the other hand, in the pitch-level theory both words have the same toneme combination, i.e., /low-high/ (and, as a matter of course, have the same phonetic pitch contour), i.e., they have one and

---

21) Hattori 1960 p.250-52

22) It is not impossible to make the two words phonetically (and phonemically) different. The pitch difference is greater, i.e. the second syllable is a little higher, in hana ('flower' (2,c)) than in hana ('nose', (2,b)), especially if the words are spoken with emphasis (Neustupný 1966 p.84; Han 1962 p.112). This difference may serve to establish a phonological contrast batween bisyllabic words of the two accent types for some speakers who either keep up or try to keep up the difference. But this is probably rather rare (Neustupný 1966).

23) This "phonetic" neutralization occurs when the final syllable of the concerned types is short (i.e., one mora in length). See later in this section.

the same accent pattern. In the pitch- level theory, they do not say it this way. But they should say so, or else the pitch-level theory loses its raison d'être. To give the reason why they should is the purpose of this section.

Recently, Weitzman (197o) set up a new pitch-level theory based on the method of autonomous phonemics. He defines a mora as a structural syllable (p.6-7), and describes accent patterns by different combinations of two "tactic phonemes" (of accent) which are put on the sequence of moras (p.7-8; and Chap. VII). According to him, his theory is "more abstractly conceived" than the other theories, but I cannot see any real difference between his and the pitch-level theories. The reason why I call his theory 'new' is that he is so consistent that he does not admit any accent pattern for one syllable (=one mora) words.[24]

Now, Weitzman argues against Hattori's theory as follows:

> "Hattori...distinguished three accent patterns for two-mora words, or, in general, n+1 contrastive accent patterns for words n mora in length." (p.16)
>
> "In the treatment of the actual accent patterns themselves and of how those patterns differ phonologically, the representation of a word that takes into account morphophonemic considerations is not of relevance. Therefore,as single-word utterances, hana ('nose') and hana ('flower') have the same accent pattern. This fact, however, does not vitiate the basic idea of the accent kernel theory, but only (1) forces a modification of it in regard to the number of accent patterns that need to be described by the theory and (2) places a restriction on the location of an accent kernel to any place except the final mora of an utterance." (p.18)

A problem here is that of so-called biuniqueness in the description. It is true that hana ('nose') and hana ('flower') invariably have the same pitch contour as sin-

24) "...the nature of Japanese [accentual phenomenon, H.M.] is that it normally can be observed only in utterances two or more moras in length. Normally no accentual distinctions are observed in utterances only one mora in length." (Weitzman 197o p.94)

gle occurrences.This fact, however, does not and should not lead to the conclusion that both *words* have one and the same *accent pattern*, since they evidently behave functionally differently when they occur in a phrase consisting of the word plus an enclitic.[25] If we say that both words have the same accent pattern, we must still make a rule that the second (strictly speaking, the last) syllable of *hana* ('flower') is, let us say, accented, therefore the following enclitic *ga* which has no fixed accent,[25] should be, let us say, unaccented and low-pitched, not high-pitched, or/and that the second syllable (i.e., the last syllable) of *hana* ('nose') is unaccented, so that the following enclitic *ga* should keep the same pitch (i.e., high pitch) as the last syllable of the preceding unaccented word. (NB. As is clear here, the "unaccented" syllable is not equal to a syllable with "low toneme" or with "low pitch", or to a syllable with "high toneme" or with "high pitch".)

As long as we stick to the pure pitch-level theory, we cannot find any motivation or reason why the toneme /high/ should be followed by an enclitic *ga* which has the toneme /low/ in the case of *hana* ('flower') and why the toneme /high/ of *hana* ('nose') should be followed by the toneme /high/. If one insists on avoiding a mixture of levels, one cannot describe accent patterns of single words. One should take morphophonemics into consideration in order to represent accent patterns of single words. In the lexicon *hana* ('nose') and *hana* ('flower') should be marked as having different accent patterns.

In the pitch-level theory in Japan, *hana* ('nose', /hana/ in the notation of kernel-theory) and *hana* ('flower', /hana⌝/ in the notation of kernel-theory) are shown as hav-

---

25) Particles are included in the morpheme category "enclitics". Not all the enclitics lack a fixed accent. See, for example, McCawley 1968 p.134 ff.

ing different accent patterns. *Hana* ('nose') is shown as /hana/ and *hana* ('flower') as /hana/.[26] Even one-syllable (one-mora) words are given two different accent patterns, for example /ha/ ('leaf') and /ha/ ('tooth')[26] (/ha/ and /ha/, respectively, in the notation of the kernel-theory). Such descriptions are contradictory to the "toneme" idea. If one admits the difference of accent patterns between /hana/ and /hana/, then how can one describe the difference

/hana/, /hana/ and /hana/,

by different combinations of two tonemes? If one admits the difference between /hana/ and /hana/, it is simply another version, and a poor one, of the accent-kernel theory. Consequently, if one is a consistent advocate of the pitch-level theory, one should have virtually the same viewpoint as Weitzman's.

Another problem is that of possible accent patterns of words and phrases consisting of *n* syllables (not *n* moras!) in length. There is a serious mixture of "syllable" and "mora", and of "single word", "single-word utterance" and "utterance" (="phrase" in my usage) in Weitzman's statement cited above. See list (VIII) on the following page. *Hana* ('nose') and *hana* ('flower') as single-word utterances cannot be distinguished phonetically. This "phonetic" neutralization occurs between two accent patterns where one has no accent and the other has its accent on the final syllable of the word. But this occurs only when the final syllable is one mora in length. When the final syllable has two moras, the two accent patterns are manifested in different phonetic forms, for example, [satoo] ('gradation') and [satoo] ('sugar') (i.e. /satoo/ and /satoo/ in the notation of the kernel theory). (Cf.

26) See Kindaichi (1958).

p. 139, l. 2 from bottom, read: But the first mora of the unaccented first syllable...

(B), and (E) in list (VIII).)

Two facts should be mentioned concerning possible accent patterns. First, the number of possible underlying (or functional) *word*-accent patterns is *n+1* for words of *n* syllables (not *n* moras) in length.[27] Second, when a phrase consists of more than two morphemes,[27a] the number of possible underlying and surface accent patterns is *n* for phrases of *n* syllables in length. For this second reason, it is often said that Japanese accents function essentially at the phrase level.[28] As for the single-word phrases, it depends upon the number of moras in the last syllable whether the word of *n* syllables are "phonetically" distinguished in *n+1* ways.

(VIII)

Possible accent patterns of n-syllable nouns[28a]

	single word	phrase	meaning
A) 1-syllable, 1-mora nouns:			
1)	$\overline{\text{ha}}$	$\overline{\text{ha}}$-ga	'tooth'
2)	ha	ha-ga	'leaf'
B) 1-syllable, 2-mora nouns:			
1)	$\overline{\text{on}}$	$\overline{\text{on}}$-ga	'favour'
2)	on	on-ga	'sound'

---

27) Actually the noun is the only morpheme class where all possible accent patterns are found.

27a) A compound word is taken here as one morpheme.

28) Cf. Shibata (1955).

28a) The notation of examples is phonemic, the accent marks being placed according to the kernel theory. "-" shows the syllable boundary. Phonetically, pitches are kept high up to and including the first mora of the accented syllable. But the first mora of the first syllable is usually low pitched. Other syllables are low pitched.

single word	phrase	meaning
C) 2-syllable, 2-mora nouns:		
1) ha-na	ha-na-ga	'girl's name'
2) ha-na	ha-na-ga	'nose'
3) ha-na	ha-na-ga	'flower'
D) 2-syllable, 3-mora nouns:		
i) 2-mora syll. + 1-mora syll.:		
1) kan-zi; hen-zi	kan-zi-ga; hen-zi-ga	'secretary'; 'accident'
2) kan-zi	kan-zi-ga	'Chinese letter'
3) hen-zi	hen-zi-ga	'reply'
ii) 1-mora syll. + 2-mora syll.:		
1) sa-too	sa-too-ga	'the opposition party'
2) sa-too	sa-too-ga	'gradation'
3) sa-too	sa-too-ga	'sugar'
E) 2-syllable, 4-mora nouns:		
1) sen-koo	sen-koo-ga	'incense rod'
2) sen-koo; sen-see	sen-koo-ga; sen-see-ga	'flash';'oath'
3) sen-see	sen-see-ga	'teacher'

Since a bound mora cannot be an accent-bearer in this dialect, the number of possible accent patterns (of single-words and of phrases) will not be consistent, if we say that the mora is the structural syllable. Examples (C,D,E) are all two-syllable nouns, so that they have 2+1 possible word accent patterns. Examples in (D) have 3 moras, and those in (E) have 4 moras. If the number of possible accent patterns is calculated on the basis of mora (D) should have 3+1 patterns, and (E) should have 4+1 patterns, but in reality they have only 3 possible patterns.

Weitzman requires a modification in regard to the number of accent patterns which are set up in the kernel theory. I do not think a modification is necessary. One should only keep in mind the distinction between the word and phrase accent patterns, and between the single-word phrase and the phrase consisting of two or more morphemes. The second modification required by him is that the accent kernel can be put to any place except the final mora (should be "syllable") of an utterance. Here again, it depends upon the character of "an utterance", whether the modification is possible or not. This matter of modification is not a trifling matter. On the contrary, this matter is the most essential one in the description of the accent system of this dialect.

7. Paradigmatic and/or syntagmatic function of accent

I take it for granted that such difference of accent as in /hana/ ('nose') and /hana˥/ ('flower'), i.e., difference between words with and without accent kernel, is necessary in the description of functional accent.

According to the pitch-level theory, accent patterns of words are contrasted with each other by permutation of tonemes which are put on a sequence of syllables.[29] This "accent pattern" corresponds to Hattori's "prosodeme" (in his own teminology). Just the same as the accent pattern, a prosodeme is a functional unit which integrates a sequence of syllables into one accent-phrase (Cf. Hattori 1960 p.268). But prosodemes are different from accent patterns in that they are distinguished from each other by

29) A sequence of syllables should be understood to include the case where a sequence consists of one syllable.

virtue of the two heterogeneous distinctive features of accents, mentioned earlier.[3o] First, a prosodeme with kernel is opposed to a prosodeme without kernel, i.e.,

a prosodeme with kernel: /o⌉oo/ /oo⌉o/ /ooo⌉/

↑ ↑ ↑

(by commutation of kernel)

↓ ↓ ↓

a prosodeme without kernel: /ooo/ /ooo/ /ooo/

(o = syllable)

In this case we are dealing with "opposition" or "commutation".

Secondly, prosodemes can be distinguished by the location of the kernel, and in this case we are dealing with "contrast" or "permutation", i.e.

3o) In this connection, I would like to say that Weitzman (197o) misunderstands some of Hattori's ideas.
Weitzman says that the prosodeme "is similar to the Akusento no kata [= accent patterns, H.M.] in that it is something superimposed upon a sequence of "moras."
He is right in this point.

He continues, "The Akusento-so, however, is more encompassing in that it is manifested not by pitch alone, as the Akusento no kata is said to be, but by other phonetic features as well." (p.15)

It seems to me that Weitzman considers that the difference between the Akusento no kata and Akusento-so lies only in the phonetic substance. But the difference lies, as I am discussing above, in their different formal approach.

one type: /o˥oo/

↑

(by permutation of kernel)

↓

another type: /oo˥o/

As is clear here, the kernel theory is based on a completely different concept from that of the toneme. In the kernel theory Japanese accents are considered to have both commutation and permutation function. This theory tells us the nature of accent in Japanese, at least that of the standard dialect.

Characteristic features in the accent system of this dialect are that the system has a formal (but not necessarily phonetic) characteristic of the so-called tone language, such as Chinese, since accents are commutable, and that it has a feature which is similar to that of a stress-accent language, since accents are permutable. However, I hesitate to state that an accent feature of this dialect is just the same as that in Chinese, since in Chinese an accent is put on the syllable, but not on the sequence of syllables as is done in Japanese.

8. Tone languages and stress-accent languages

In this dialect of Japanese maximally one accent is put to a phrase, i.e., to a sequence of syllables, and separate information about individual syllables is not necessary.

Concerning the accent phenomenon, McCawley (1968) distinguishes three types of languages, (1) a stress-accent language such as English, (2) a pitch-accent language such as Japanese, and (3) a true tone language such as Chinese. Types (2) and (3) resemble each other, for they use the same "substance" i.e., pitch movements, to realize functional accents.[31] Types (1) and (2) resemble each other, since accents are essentially of syntagmatically contrastive nature. Types (1) and (3) have a common formal similarity in that, admitting more than three stress degrees in the type (1), not the postion of one stress in (1), nor one tone in (3), is predictable from the position of another stress, nor another tone. Thus, he suggests "either grouping of types 1 and 3 or the position that type 1 is intermediate between types 2 and 3." (p.136-37)

As a conclusion, he says,

> "A language with a "pitch accent" system like Japanese and a language with a "stress accent" system like Russian have the formal similarity that the accentual information which must be recorded in dictionary entries is at most the location of some accentual phenomenon, rather than separate information about individual syllables, as is needed for pitch in the dictionary entries of a true tone language like Chinese." (p.182)

> "I feel that the above considerations give a rather convincing case that "stress accent" and "pitch accent" are merely two manifestations of the same linguistic phenomenon and that it consequently is a mistake to label as " tone language" those languages which have a "pitch accent" system such as Japanese." (p.183)

That is because, in addition to the above reason, cyclic accent reduction rules are applied to both "stress-" and "pitch-" accent languages, but not to a tone language. For

> "Pitches in a tone language, on the other hand, is not subject to "reduction" rules but merely to the same kinds of rules as affect segmental features (assimilation, dissimilations, etc.). (p.183)

---

31) But,it is known today that there are many "stress-accent" languages where the acoustic fundamental frequency is a distinctive cue for stress perception.

As has been discussed, nouns have "n+1" accent patterns. Using McCawley's expression, the accent of nouns is put on the first, second, ..., or no syllable. (But it is true that the noun is the only morpheme class in which the location of accent is completely free. In verbs and adjectives, which are the only morpheme classes with inflections, there are only two possible accent patterns: with or without accent.)

Instead of saying that an unaccented morpheme has accent on no syllable, McCawley defines accent as occurring between the syllables, so that an unaccented morpheme is defined to have "pre-accent".[32] Thus, the three hana's are specified in the underlying notation in the forms:

ha'na	(= Hattori's	/$\overline{\text{ha}}\rceil$na/)	'girl's name'
'hana	(= "	/hana/)	'nose'
hana'	(= "	/$\overline{\text{hana}}\rceil$/)	'flower'

Then, the possible places of accentuation in n-syllable morphemes are equal to n+1 syllable boundaries. (p.169-7o) The argument here is that Japanese accent is described most generally (or economically) as to its function, when one specifies accent and describes manifestation rules in the way of generative phonology.

A characteristic feature of accent placement in Japanese should be mentioned. In one morpheme class (i.e., nouns) restriction to two possible places of accent applies to one-syllable morpheme only, while there are two and only two possible places of accent for morphemes in another class (i.e., verbs and adjectives), independent of the syllable number of the morphemes.

---

32) There are "pre-accented" nouns and "pre-accented" affixes. Affixes are governed by the "pre-accentuation" rule, but pre-accented nouns are not.

When compared with a "stress-accent" language, Japanese morphemes with one syllable (i.e., nouns) have, then, two possible accent placements, while in a stress-accent language one-syllable morphemes in a class have no difference of accent placement, i.e., they are simply "accented" (or, possibly, only "unaccented"). This fact shows a basic difference between Japanese and stress-accent languages.

All of the above discussion has been about the standard dialect. When we take other dialects into consideration, we are not so sure about the extent of the formal similarity between Japanese as a whole and a stress-accent language as a whole. This matter will be mentioned in Section 1o.

9. The role of mora

It seems to be true that the mora plays some functional role in Japanese, but no Japanese linguist gives a clear argument for its functional role.

Hattori says that the mora is a phonological unit but not a phonological syllable. (Cf. Section 3.) He does not say that the kernel occurs on the syllable, but says that it occurs on the mora. In the Kyoto dialect the phonemes /N/ and the second vowel of a sequence of two identical short vowels can be accent-bearers. But they are still "bound" moras in that they always occur in syllable-final position.[33] (see Section 1o). If we take Hattori's phonological syllable to be the unit which can best describe segmental combinations, then the phonological syllable will be valid to almost all, if not all, the dialects in Japanese. If the phonological syllable is to be a unit which has to do with both "accent" and "segmental combinations", then it is not applicable to all the dialects. He might be taking an "over-all" description into account. But, as far

33) Actually, in the Kyoto dialect a sequence of two identical vowels belong always to two different syllables. (Cf. fn.36.)

p. 147, l. 16-17, read: "Japanese has phonological rules which depend on the number of moras, but none to my knowledge which depend on the number of syllables. For example..."

as the accent phenomenon and the phonotactic combinations of the standard dialect are concerned, we do not find any strong motivation in his theory that the mora is a phonological unit.

McCawley (1968) says, "There are many reasons why the notion of mora must be used in describing standard Japanese." (p.131) The reasons are:

(1) "the mora functions as the unit of length in the language; not only is the length of a phrase roughly proportional to the number of moras it contains ..., but the meters of Japanese are based on the number of moras per line". (p.131)

(2) "the acoustic realization of Japanese accent can only be described by stating which moras are high or low pitched." (Ibid.)

(3) "Japanese has phonological rules which depend on the number of syllables. For example, in a certain class of foreign loan words the accent is put three moras from the end of the word. The mora is thus the "unit of phonological distance" in Japanese." (p.133)

But McCawley emphasizes that the "prosodic unit of Japanese is the syllable and not the mora", since "there is no contrast between "accented first mora" and "accented second mora" in a long syllable" (p.134). Thus, Japanese is a "mora-counting syllable language" (Ibid).

The role of mora in metrics is definite and important. But, sometimes "bound" moras do not seem to be counted as an independent unit (here, syllable) in poems, especially in modern ones. At the same time, we also find one (but rarely more) extra free mora (here, syllable) in a line. It is difficult to say whether the surplus bound moras are meant to be extra moras or violence to the rigid metrics.

McCawley's point (2) is an evident fact, at least auditorily. It does not, of course, necessarily mean that the pitch contour on one mora acoustically keeps the same pitch level all through the mora.[34]

In regard to point (3), McCawley says,

> "rules such as the one putting accent three moras from the end of foreign words will put the accent on the fourth mora from the end if the third mora from the end happens to be the second mora of a long syllable; the correct form of the rule is thus "place accent on the <u>syllable containing the third from last mora</u>."" (p.134)

Another example from McCawley is a rule which "operates at the distance of one mora". An accent attraction rule for the verbal infix morpheme of the "provisional" mode operates only when the morpheme is one mora in length (i.e. /CV/), but when the morpheme is more than one mora in length (i.e. /CVC-/) the attraction rule does not operate. This rule is quite important in his rules. The morpheme with /CVC/ structure is governed by a (consonant) deletion rule, but this rule should be applied after the accent attraction rule.[35]

Taking all three points together, I think that the mora is a functional unit.

10. Some remarks on accent in the Kyoto dialect compared to that in the standard dialect

As I have mentioned in a couple of places, the Kyoto dialect has some features that differ from those of the standard dialect. The so-called syllabic <u>n</u> (/N/) in the

---

34) Cf. Weitzman (1970), Chap. IV; Mase (1969)

35) In this connection, I would like to say that in Japanese the syllable and morpheme boundaries do not always coincide. Therefore, the following description by Garde (1968 p.25) is incorrect: in French "les limites de syllabes et de morphèmes ne coïncident pas ... Mais en japonais toute limite de morphème coïncide avec une limite de syllabe, donc la syllabe est incluse dans le morphème."

Kyoto dialect can be accent-bearer: "oNna" ('woman'), "koNbaN" ('tonight'). This holds true also of the second vowel in the sequence of two identical short vowels: "keeba" ('horse race'), "keesi" ('disregard'), "keezi" ('policeman'), "keezi" ('notice').[36]

In two-syllable nouns we find four accent patterns:

	K.-a.	K.-b.	Hat.	McC.[37]	
1)	hana, hana ga	hana	hana	hana	('nose')
2)	yama, yama ga	ya⌉ma	⌈ya⌉ma	ya'ma	('mountain')
3)	sora, sora ga	⌉sora	⌊sora	'sora	('sky')
4)	saru, saru ga	⌉saru⌉	⌊saru⌉	'saru'	('monkey')

In this dialect, the pitch on the first mora is phonologically distinctive (cf. fn.16). According to Hattori, the accent system of the Kyoto dialect is described by three distinctive features:

1) with or without kernel,
2) where is kernel, and
3) low- or high-initial.

---

36) Examples are from Hirayama (1960)
As for /Q/, I found under the head-word "ke-" and "ko-" in Hirayama (1960) such types as "geQka" and "keQka", but not "keQka", or "keQka" types. It seems to me that there is not contrast between /VQ/ and /VQ/, or between /VQ/ and /VQ/.
/Q/ in Japanese as a whole occurs only phrase-medially, except some peripheral cases (i.e., in exclamatory words and onomatopoeia). The phrase-medial /Q/ is obtained from the underlying oral obstruent segment (except /Q/, which occurs in some onomatopoeia (cf. McCawley (1968)). Thus, we perhaps need not set up /Q/.

37) Examples are phonemic notation. But K.-a. system shows actual pitch patterns.

K.-a. = Kindaichi 1960 p.255; K.-b. = Kindaichi 1960 p.305; Hat. = Hattori 1961 p.4; McC. = McCawley 1968 p.192.

Even in the pitch-level theory, unidimensional tonemes cannot describe the accent patterns of the Kyoto dialect. In order to distinguish (3) and (4), one must put a single /high/ toneme on the second syllable of (3) and a two-toneme combination /high-low/ on the second syllable of (4) (—K.-a.). Or else, one must apply "pre-accentuation" (—K.-b.).[38]

In any event, one must assign two accents to the accent-phrase of the dialect. It should be well noticed that all the combinations of the higher and lower registers are found in underlying two-syllable nouns.

$\overline{\text{hana}}$

$\overline{\text{ya}}\underline{\text{ma}}$

$\underline{\text{sora}}$

$\underline{\text{sa}}\overline{\text{ru}}$

The accent system of this dialect does not particularly resemble that of a stress accent language. But it is true that the same kind of reduction rules as for the standard dialect and stress languages will apply to this dialect, too.

Japanese accent is described as functioning syntagmatically in the pitch-level theory, paradigmatically and syntagmatically in the accent-kernel theory, and syntagmatically in generative phonology. The difference between the three approaches is that of levels of description. The method of the pitch-level theory is that at the first stage one should identify accent function mainly at the phonemic level. The method of the kernel theory is the next stage: identifying the accent function at the morphophonemic level. The method of generative phonology is applied to describe the accent function based on the results of the first two ana-

38) Shibata (1955) criticizes Kindaichi, saying that Kindaichi's toneme-conception is incompatible with his idea of "pre-accentuation". (See Shibata 1955 p.53-55)

lyses. One should not take it for granted that the method of generative phonology is the best one, since different methods could be applied at different stages of linguistic analyses.

References

Bloch, Bernard 195o: "Studies in Colloquial Japanese IV, Phonemics", Language 26, p. 86-125.

Garde, Paul 1968: L'accent (Paris).

Han, Mieko Shimizu 1962: Japanese Phonology, An Analysis Based upon Sound Spectrograms (Tokyo).

Hattori, Shiro 1951: Onseigaku ('Phonetics') (Tokyo).

" " 196o: Gengogaku no hoohoo ('Methods in Linguistics') (Tokyo).

" " 1961: "Akusento-so, onsetsu-koozoo, koo'on-onso" ("Prosodemes, Syllable Structure, and Laryngeal Phonemes"), Study of Sounds IX, p. 1-21.

Hirayama, Teruo 196o: Zenkoku akusento jiten ('Accent Dictionary of Japanese Dialects') (Tokyo).

Hockett, Charles F. 1955: A Manual of Phonology (Baltimore).

" " 1958: A Course in Modern Linguistics (New York).

Kawakami, Shin 1966: "Teineina hatsuon to akusento no kata" ("The Deliberate Pronunciation and the Word-Pitch Pattern"), Kokugogaku 64, p. 1-1o.

Kindaichi, Haruhiko (ed.) 1958: Meikai nihongo akusento jiten ('Japanese Accent Dictionary') (Tokyo).

" " 1967: Nihongo on'in no kenkyuu ('Studies in Japanese Phonology") (Tokyo).

Martin, Samuel E. 1967: "On the Accent of Japanese Adjectives", Language 43, p. 246-277.

Mase, Hideo 1969: "Acoustic Cues for the Perception of Accent in Japanese when the Accented Vowel is Devoiced", ARIPUC 3/1968, p. 143-176.

McCawley, James D. 1968: The Phonological Component of a Grammar of Japanese (The Hague-Paris).

Neustupný, J.V. 1966: "Hana ('flower') to hana ('nose') no hatsuon wa hatashite chigau ka" ('On the Difference of Pronunciation between hana ('flower') and hana ('nose')", Gengo seikatsu 172, p. 80-87.

Pike, Kenneth L. 1948: Tone Languages (Ann Arbor).

Shibata, Takeshi 1955: "Nihongo no akusento-taikei" ("Accent Systems of Japanese Dialects"), Kokugogaku 21, p. 44-69.

Weitzman, Raymond S. 1970: Studies in the Phonology of Asian Languages IX. Word Accent in Japanese (Acoustic Phonetic Laboratory, University of Southern California).

ENERGY MEASUREMENTS OF SPEECH SOUNDS

Carl Ludvigsen

1. Introduction

Recording of so-called intensity curves frequently forms a part of the routine procedure for processing speech material collected in a phonetic investigation. Devices used for recording of intensity curves are described by Fant (1958 p. 57).

The interpretation of such curves is generally not an easy task as is shown by Fant's characterization of the intensity meter as "... a device that produces an electrical voltage which represents but is not necessarily proportional to the intensity of the speech wave"(1958 p. 57).

The aim of this paper is twofold: to discuss very briefly some aspects of the registration of acoustic energy, and to describe a device constructed at the Institute of Phonetics for measurements of the energy of sound segments. Segments measured with this device may be very short, i.e. of a duration of a few milliseconds.

2. Acoustic measurements

Peterson and McKinney (1961 p. 81) argue that "the measurement of spech power level is relevant to the study of linguistic stress." However, very little attention has been given to the problem of which among several possible measures of speech power level is of major relevance. The following section is devoted to some loose considerations of this question.

If we want to study the acoustics of speech sounds with reference to their properties as stimuli to the ear the measurement of speech power introduced by Fletcher (1953 p. 68) does not seem to be ideal: Fletcher's speech power measures are measures of the total speech energy radiated while a person is talking, whereas here we are only interested in that part of the energy which is active in stimulating the listener. Due to the shadowing effect of the talker's head the two measures are not proportional. Furthermore, it is very difficult to perform the practical measurements.

Another frequently employed measure is obtained by placing a microphone at a distance of e.g. one meter from the talker and feeding the electrical signal from the microphone into an "intensity meter" or a sound level meter. Several problems are involved in this registration procedure: Firstly, does the electrical signal from the microphone represent the sound pressure as it would have been if no microphone were present? This is only the case if the sound field is simple, e.g. a frontal free field, and the microphone is modified to have a corresponding free field response. Secondly, is the sound uniquely described by the electrical signal from the microphone if the sound is regarded as a stimulus to a listener placed in the sound field? This is true only with regard to one type of sound field, cf. ISO Recommendation 454.

The conclusion seems to be that the sound stimulus should be registered at the listener's ear drum or, more conveniently, from a pressure response microphone placed at the end of the ear channel of a dummy head. A recording of the sound pressure at this place will presumably be the best physical description of the stimulus.

3. Processing of the speech signal

The material for further processing is supposed to be an electrical signal proportional to the instantaneous sound pressure recorded as described above. In order to obtain a measure of the power of the signal some kind of rectification and a smoothing process have to be performed.

3.1. The rectifier

If we want a DC-voltage proportional to the instantaneous power of the microphone signal we must use a square law rectifier (SLR). It is known from the theorem of Parseval that for a periodic[1] signal the DC-value of the squared signal will be proportional to the sum of the mean powers from each of the frequency components of the original signal. Furthermore, we know that this DC-component of the squared signal is independent of the phase properties of the input signal.

The full wave linear rectifier (FWLR), on the other hand, gives an output signal whose DC-value is proportional to the square root of the mean power[2] of the input signal only for very simple input signals. When the signal is not very simple we cannot generally predict the deviation from the correct value. Furthermore, the output from a linear rectifier is sensitive to phase changes of the input signal.

To demonstrate differences and similarities between the two rectifiers a few examples will be examined. The DC-component emitting from a true RMS-detector consisting of a square law rectifier, a low-pass filter, and a square root device is compared with the DC-component emitting from a mean value detector, i.e. a full wave rectifer followed by a low-pass filter.

---

1) Similar statements can be made for aperiodic signals.
2) i.e. the RMS value of the signal.

(a) Pure tone with amplitude A:

RMS-detector: $A/\sqrt{2}$

Mean value detector: $2A/\pi$

(b) Narrow band noise with mean power P_m: [3]

RMS-detector: $\sqrt{P_m}$

Mean value detector: $\sqrt{2P_m/\pi}$

(c) Amplitude modulated pure tone:

Mean amplitude: A, Index of modulation: m

RMS-detector: $\sqrt{A^2/2 + m^2A^2/4}$

Mean value detector: $2A/\pi$

Now, suppose that we tested two intensity meters: One with a full wave linear rectifier (FWLR) followed by a low-pass filter and the other with a square law rectifier (SLR) followed by a low-pass filter and a square root circuit, i.e. a true RMS-detector. If we adjusted the two instruments to give the same reading for a pure tone input signal, then we would not obtain identical readings on the two instruments for any of the other two signals mentioned above. For the amplitude modulated tone we could, by varying the degree of modulation (m), obtain a varying reading on the SLR, but a constant reading on the FWLR.

None of the examples mentioned were speech sounds, and experiments (cf. Peterson and McKinney 1961 and Fant 1958 p. 57) have shown that the differences occur-

---

3) Cf. Rice (1944).

ring when speech sounds are considered are moderate. However, with the linear rectifier it is somewhat dubious what we actually measure, and conclusions drawn from intensity curves should perhaps be of a qualitative rather than a quantitative nature. Furthermore, the use of an FWLR will complicate a statistical treatment of the data from intensity meters since differences in the means of two variables may to a certain degree be due to unsatisfactory rectification.

Nevertheless, many rectifiers used in intensity meters are of the linear type. The reason obviously is that an SLR is more difficult to construct if it is to have a sufficiently large dynamic range. However, as technology has advanced in the last years we are now able to construct SLR's which satisfy our demands and which are not excessively expensive.

3.2. Integration

If we use an intensity meter with an SLR followed by a smoothing low-pass filter it is no simple task to measure the energy of a single sound segment. Even if the segment is isolated the interpretation of the output is not a straightforward matter. If the duration of the segment is long compared to the transient time[4] of the filter the mean of the output signal equals the average power of the segment provided that the AC-part of the output signal is small compared to the mean, i.e. to the DC-component of the output signal. If, however, the duration of the segment approximates the transient time of the filter the output signal will be highly dependent on the shape of the impulse response of the filter. In this case the set-up does not

---

4) Often called "the integration time" of the filter, cf. Fant (1959 p. 9).

give useful measurements. If the duration of the segment is short compared to the transient time then the peak of the output signal will correspond to the energy of the input signal. When the set-up is calibrated for this type of measurement and the duration of the signal is known we can calculate the average power.

From these considerations we may conclude that if the energy of transients such as stop bursts is measured by the output of an ordinary intensity meter the results are subject to several uncertainties.

To overcome the difficulties involved in energy measurements of very short segments of speech one may replace the low-pass filter by a true integrator. The integration should start at the beginning of the sound segment in question and stop at the end of the segment. The final output voltage from the integrator would then be proportional to the average power of the segment, irrespective of its duration.

4. Construction of a square law rectifier and a true integrator

Fig. 1 shows a simplified diagram of the device.

4.1. Principle of the square law rectifier

The characteristic of the rectifier is a piecewise linear approximation to a parabola. The input signal will be fed into a voltage divider via an FWLR. The voltage levels on the voltage divider will trigger the four transistors on and off. The emitter resistors are selected so that the output current is approximately proportional to the square of the instantaneous input voltage to the squaring circuit.

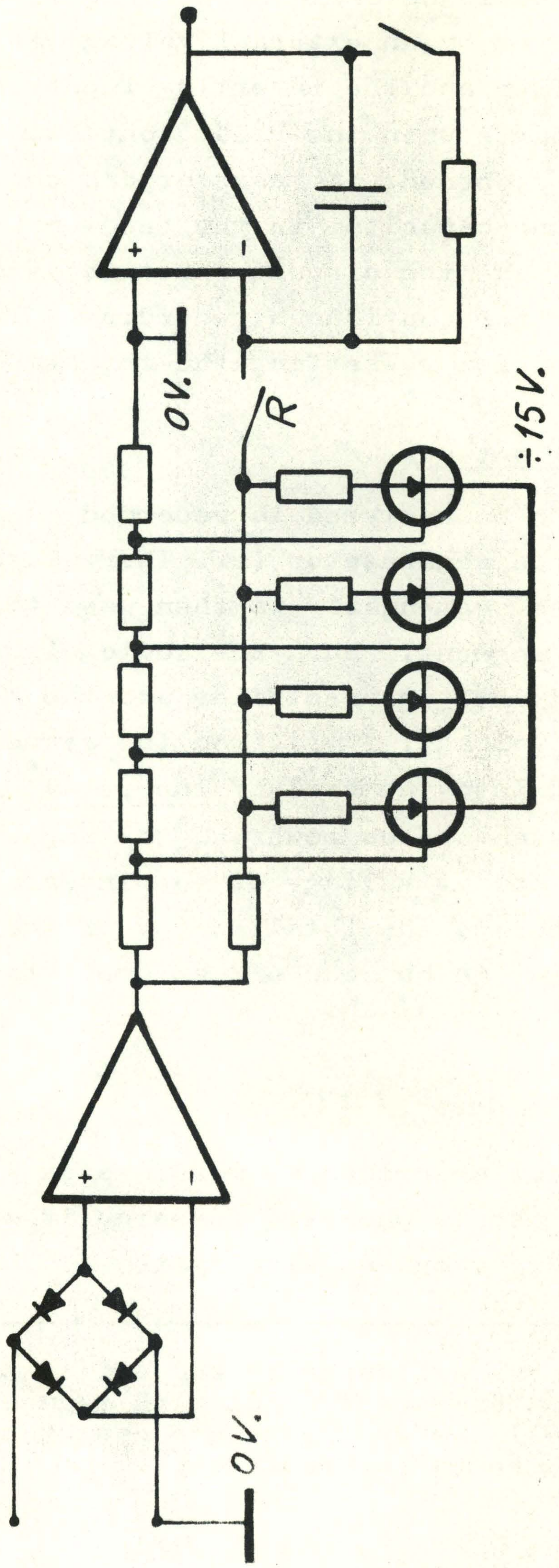


fig.1

4.2. The integrator

The integrator is an ordinary summing integrator. A relay, R, triggered by an external voltage is inserted between the rectifier and the inverting input terminal. The integration starts when the relay contacts close and stops when they are opened. A resistor can be coupled in parallel with the capacitor in the feed-back loop of the integrator to make the circuit function as a simple RC low-pass filter for continuous recording. This resistor is also used for re-setting the integrator.

4.3. Experimental set-up

The signal to be analyzed is recorded on the loop tape-recorder of the segmentator (cf. Thorvaldsen 197o). The "window" of the segmentator is then used to isolate the desired sound segment. When the audio signal from the segmentator is fed into the rectifier and the "gating signal" (i.e. a step voltage: _⎍_) from the segmentator is simultaneously fed into the coil of the relay[5] an integration of the instantaneous power of the segment takes place. The final output voltage of the integrator will then be proportional to the total energy of the sound segment. This voltage can be read off an amplifier DC-voltmeter.

5. Final remarks

The instrument described is meant as an experimental design only. For use in phonetic research it must be incorporated in a more comprehensive system. The addition

---

5) This synchronous switching on and off of the input to the integrator is essential with segments of very low intensity in order to exclude contributions from background noise on the tape loop.

of a compressor as well as a more sophisticated low-pass filter for continuous recordings will be an obvious improvement.

References

Fant, Gunnar 1958: "Modern Instruments and Methods for Acoustic Studies of Speech", Acta Polytechnica Scandinavica.

" 1959: "Acoustic Analysis and Synthesis of Speech with Application to Swedish", Ericsson Technics No. 1.

Fletcher, Harvey 1953: Speech and Hearing in Communication. D. Van Nostrand Company, Inc.

ISO Recommendation 454: Relation between the loudness of narrow bands of noise in diffuse-field and in a frontally incident free-field.

Peterson, Gordon E. and McKinney, Norris P. 1961: "The Measurement of Speech Power", Phonetica Vol. 7.

Rice, S. O. 1945: "Mathematical Analysis of Random Noise", Bell System Technical Journal (January 1945).

Thorvaldsen, Poul 1969: "Construction of a Segmentator", ARIPUC 4, p. 8-14.

COMPARISON OF SOUND PRESSURE LEVEL AND LOUDNESS(GF) MEASUREMENTS ON SPEECH SOUNDS

Carl Ludvigsen and Nina G. Thorsen

1. Introduction

Lately, an instrument called a loudness analyzer (HP 8051A Loudness Analyzer), which operates according to Zwicker's method of estimating loudness, has been made commercially available.

Zwicker's method is based on a division of the signal in critical bands. In each critical band the signal is weighted according to the sensitivity of the ear. The contribution of the sub-signals are added up, allowance being made for masking phenomena. - The output of the loudness analyzer is called loudness(GF)* and is measured in sones(GF). "G" refers to critical bands (Frequenzgruppe) and "F" refers to a frontal sound field.

The aim of this investigation was to compare sound pressure level (SPL) measurements with loudness(GF) measurements on some consonants and vowels.

The investigation must be regarded as a pilot experiment since only two subjects were involved and the material was recorded once only, by each subject.

2. The material

The material consisted of isolated bisyllabic words of the CV:Cə type with trochaic stress pattern, which is a common type in Danish. C represents s, f, n, m, and V represents i e ɛ a y ø œ u o ɔ ɑ ɒ. The consonants and

*) It is also possible to get an output called loudness(GD), where "D" refers to a diffuse sound field.

vowels were combined in all possible combinations, forming a list of 192 different words. - The material was originally designed for a different purpose, namely an examination of the relations between fricatives, nasals, and vowels. We found it suitable for the present purpose too, since the unvoiced fricatives, the nasals, and the vowels are three acoustically distinct classes of speech sounds with very different spectra.

The words were recorded once by two subjects, one male (CL) and one female (NT), (the authors), who both speak a form of advanced Standard Copenhagen Danish. - We attempted to produce the words at a comfortable constant level, but we cannot exclude that our knowledge of the purpose of the experiment has influenced the reading of the words. This question will be treated again briefly in section 5.1.1.

3. Instrumental set-up

The recordings were made in the sound studio at the Institute of Phonetics. The subjects were placed comfortably at a distance of approximately 1 meter from a free field response microphone (B&K, type 4131).

The response of a free field microphone is adjusted so as to deliver a signal which corresponds to the SPL value obtained if no microphone were present. In the B&K, type 4131 microphone this is accomplished by a damping of the frequencies above 2000Hz of approximately 6dB/octave. This characteristic of the microphone has been calculated for use in free, plane sound fields, 0^o incident, and this situation is approached by placing the microphone at a distance of 1 meter from the subject.

The signal from the microphone was picked up on a tape recorder via a microphone amplifier (B&K, type 2603).

A pure tone of 1000Hz preceded and followed the recordings for the sake of calibration.

3.1. Processing of the speech signals

3.1.1. Tracing of words

The signal from the tape recorder was fed into different analyzing devices, the outputs of which were traced on an ink writer[1]. The analyzing devices were

(1) fundamental frequency detector

(2) intensity meter

(3) loudness analyzer.

3.1.2. Tracing of isolated speech sounds

The peak mode operation of the loudness analyzer implies that a segment with a small loudness(GF) value immediately preceded by a segment with a larger loudness(GF) value will not cause any further deflection of the loudness curve. - To render possible loudness(GF) measurements of such weaker segments all segments, except the schwa vowels, were isolated on a segmentator and recorded on tape. The segmentator is described by Thorvaldsen (1970). The rise and fall times of the segments were set at 5ms. - The segments were then analyzed in the same way as the words (cf. 3.1.1.).

4. Calculations

The data were subjected to statistical treatment. Four variables were included in the process

(1) the duration of each segment

(2) the maximum SPL of each segment

(3) the maximum loudness(GF) of each segment

---

1) Instruments which are not further specified are listed in ARIPUC vol. 4, 1969, p. II-V.

(4) the $\log_2$ of the maximum loudness(GF) of each segment.

The total population of sound segments analyzed was treated in various sub-populations, as well as pooled. This was done in order to watch the variation in the SPL and loudness(GF) measurements in vowels in different surroundings before pooling these measurements.

The statistical treatment involved calculations of mean, standard deviation, range, skewness, mean error, and 95% and 99% confidence limits. Each subject was treated separately.

5. Results

5.1. The vowels

(1) A cursory glance at the tracings of the whole words tells us that the course of the loudness(GF) and SPL curves are not always correlated. - In most words the loudness(GF) and SPL maxima were distributed in the same way over the two syllables, i.e. either the full vowel had the greater loudness(GF) and SPL values, or both maximum values were reached in the unstressed vowel. - However, there were instances where the maximum value of loudness(GF) in a word was reached during the unstressed schwa, but the maximum SPL value was measured in the stressed vowel, and vice versa.

(2) In Table I are given the means of the duration, SPL, and loudness(GF) values for the (isolated) vowels.

Table I

Means of duration, SPL, and loudness (GF) for 12 vowels, all surroundings pooled, 16 examples of each vowel. Two subjects, NT and CL.

	duration [ms]		SPL [$dB/2x1o^{-5}N/m^2$]		loudness(GF) [sones(GF)]	
	NT	CL	NT	CL	NT	CL
i	2o9	2o6	74,6	73,1	1o,7	1o,8
e	251	225	7o,2	74,5	1o,9	14,2
ε	264	251	68,5	72,3	15,6	15,3
a	268	266	68,4	72,3	17,9	16,9
y	241	218	75,3	74,4	12,o	11,2
ø	26o	239	7o,3	74,1	12,1	14,3
œ	273	26o	69,4	71,9	17,2	15,2
u	256	22o	74,1	73,8	1o,3	1o,o
o	264	251	7o,4	72,1	1o,1	1o,8
ɔ	268	269	7o,o	7o,8	15,1	12,3
ɑ	288	281	71,6	69,9	23,1	16,7
ɒ	279	262	7o,4	7o,9	16,4	12,4

The SPL and loudness(GF) means were subjected to a ranking correlation test. The ranks (1, 2, ...) were assigned to each value according to its magnitude.

(a) The ranks of the two subjects' SPL means were not significantly correlated. The Spearman rank correlation coefficient, r_s, was calculated to be o,2o. In order for the rank correlation to be significant at the 95% level it should exceed o,5o.

(b) However, the ranks of the two subjects' loudness(GF) means were significantly correlated. The r_s value was calculated to be o,86. In order for the correlation to be significant at the 99% level it must exceed o,7o, only. - This must be due to the fact that the SPL measures depend on the total power of the sound, whereas the loudness(GF) measures are dependent also on the shape of the spectrum, which dependency has been decisive for the output of the loudness analyzer.

(c) The ranks of the loudness(GF)/SPL means for the subject NT were negatively correlated, r_s = -o,52, which exceeds the demand for significance at the 95% level. The same ranks for the subject CL were likewise negatively correlated, r_s = -o,3o, which value is not significant at the 95% level.

5.1.1. Vowel SPL and loudness(GF) vs. F1-frequency

The SPL means in Table I are displayed graphically in Fig. 1. It will be seen that in most vowels small SPL corresponds to high F1. Correlation calculations yielded: The rank correlation between SPL means and F1-frequency for both subjects was found to be negative at a level of significance greater than 95%[2]. This is surprising in the light of earlier investigations, cf. for instance House, Fairbanks and Stevens (195o) and Lehiste and Peterson (1959). The reason may very well be found in the reading of the words. For example, in trying to produce all the words with the same strength we may subconsciously have compensated for the smaller SPL of the close vowels. Furthermore, the results we obtained gave rise to a renewed inspection of the word lists and the tape. This revealed that the originally intended randomization had failed in certain respects. - We cannot exclude that a certain rhythmic effect may have been active, but a phonetically trained listener did not detect

---

2) Vowels of approximately the same F1 were tied in four groups (iyu)(eøo)(ɛœɔ)(aɑɒ). These ties were incorporated in the ranking tests.

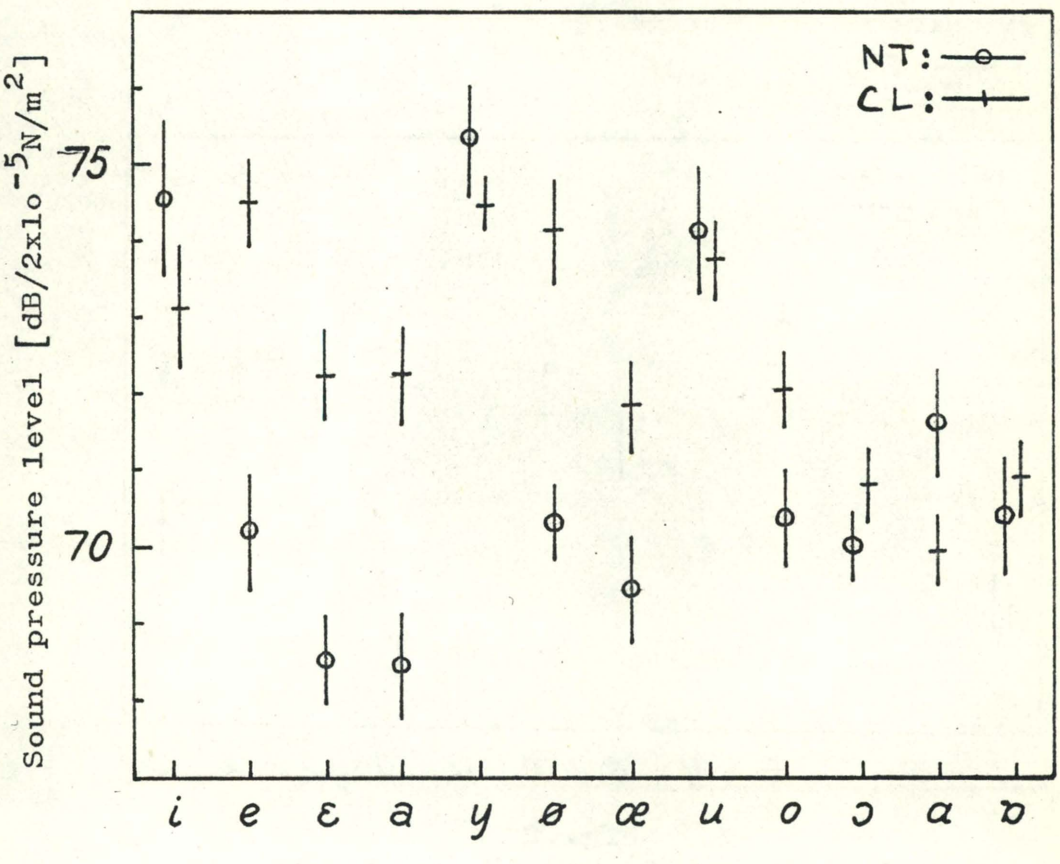


Fig. 1
Means of SPL and 95% confidence limits for 12 vowels, all surroundings pooled, 16 examples of each vowel for each of the two subjects, NT and CL.

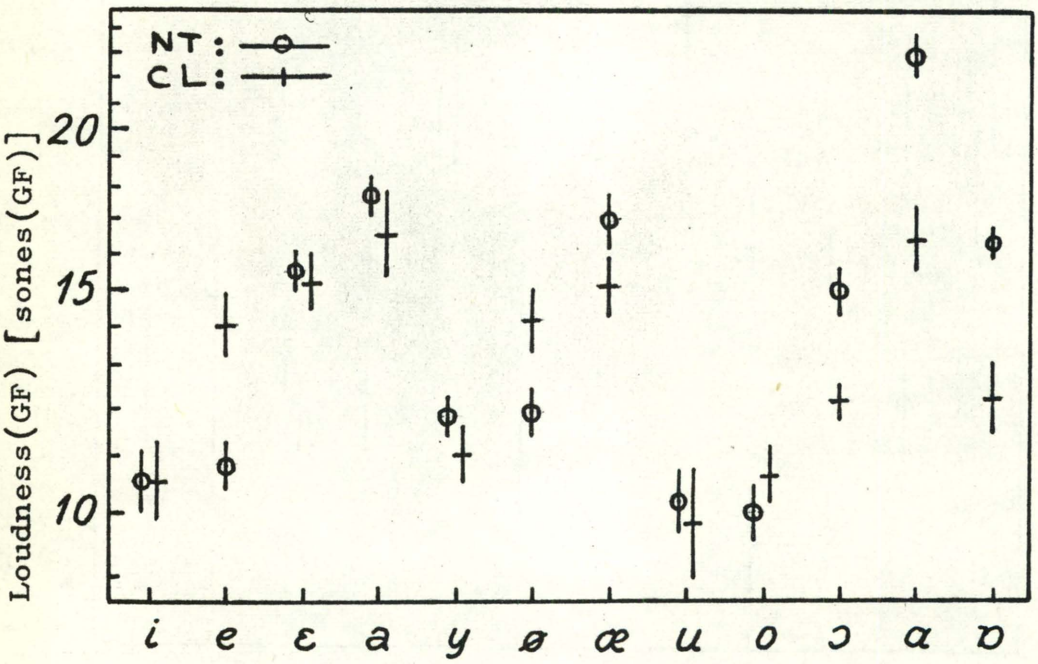


Fig. 2
Means of loudness(GF) and 95% confidence limits for 12 vowels, logarithmic scale, all surroundings pooled, 16 examples of each vowel for each of the two subjects, NT and CL.

any particular rhythm during a casual listening to the tape. Thus, our results cannot invalidate what is a very common experience, namely that vowels with low F1 have smaller SPL than vowels with high F1.

On the other hand, the means of the loudness(GF) obviously rise with rising F1, cf. Fig. 2. We found a significant positive correlation between loudness(GF) and F1. For both subjects the rank correlation calculations yielded values that exceed the demand for significance at the 99% level. - This is in accordance with the frequency response of the loudness analyzer.

5.1.2. Vowel duration

In Table I can be seen that vowel duration increases with increasing mouth opening, which corresponds well with earlier findings, cf. for instance House and Fairbanks (1953).

5.2. The consonants

In Table II are given the means of the duration, SPL, and loudness(GF) values for the (isolated) consonants.

TABLE II

Means of duration, SPL, and loudness(GF) for 4 initial and 4 intervocalic consonants, all surroundings pooled, 48 examples of each consonant. Two subjects, NT and CL.

	duration [ms]		SPL [dB/2x1o^{-5}N/m^2]		loudness(GF) [sones(GF)]	
	NT	CL	NT	CL	NT	CL
s-	161	19o	59,1	57,4	7,8	7,8
f-	133	161	52,3	52,1	4,3	5,o
n-	122	115	73,1	71,2	8,2	6,8
m-	1o7	113	72,8	71,3	7,9	6,9
-s-	1o8	1o7	59,4	57,4	7,4	6,7
-f-	11o	111	51,6	52,8	4,7	4,3
-n-	79	66	77,5	7o,1	9,4	5,1
-m-	88	91	78,4	7o,2	1o,4	5,4

The SPL value of s is only 5-7dB higher than that of f. This difference is smaller than that found in many earlier investigations (15-2odB). This may be due to the operation of the free field response microphone, cf. section 3. - The loudness(GF) difference, however, is evident, since the noise of s lies in a frequency domain where the ear, and therefore the loudness analyzer, is very sensitive.

6. Final remarks

An evaluation of the loudness analyzer as a tool in speech research must await further experiments, which must obviously involve listening tests.

References

Fairbanks, G., A.S. House and E.L. Stevens 195o: "An Experimental Study of Vowel Intensities", JASA 22, 4, p. 457-459.

House, A.S. and G. Fairbanks 1953: "The Influence of Consonant Environment upon the secondary Acoustical Characteristics of Vowels", JASA 25, 1, p. 1o5-113.

Lehiste, I. and G.E. Peterson 1959: "Vowel Amplitude and Phonemic Stress in American English", JASA 31, 4, p. 428-435.

Thorvaldsen, P. 197o: "Construction of a Segmentator", ARIPUC 4/1969, p. 8-14.

A STUDY OF CONSONANT QUANTITY IN WEST GREENLANDIC

Hideo Mase and Jørgen Rischel

1. Introduction

The phonetic structure of Eskimo words is characterized by very heavy constraints on the possible clusterings of consonantal segments. These constraints become heavier the further one proceeds toward the East. In Greenland, which is the easternmost part of the contiguous region in which Eskimo is spoken, the surface phonetic word structure is characterized by a very monotonous distribution of consonant segments (see section 2. below; also cf. Bergsland 1955 p. 3, where it is shown that this is an innovation of modern Greenlandic).

On the other hand, Greenlandic is characterized by a high degree of utilization of the contrast between short and long, or single and geminate, segments. This is true of vowels as well as consonants. The importance of quantity was emphasized already in the earliest grammar of West Greenlandic (Egede 1760 p. 6-7), where it is stated that there are two kinds of "accent", viz. long and short. The data given are somewhat confusing, but there are at least some examples which - also in modern West Greenlandic - are perfect minimal pairs, e.g. /manna/[1] 'this' (Egede: Mánna) versus /maanna/'now' (Egede: Mânna); /agi(j)utaa/ 'his file' (Egede: Aggiutá) versus /aggi(j)utaa/ 'his day of arrival'

1) For practical reasons we frequently refer to a slightly modified version of the autonomous phonemicization found in such works as Swadesh (1946), Bergsland (1955). Slant lines are used only with autonomous phonemic notations; abstract representations of a generative phonology are underlined when occurring in running text.

(Egede: Aggiúta). - In contemporary language the functional load of the quantity contrast has obviously increased compared to the eighteenth century, because far-reaching assimilation processes have been at work after that time. Several of these processes survive as rules of grammar in modern West Greenlandic: there are numerous clusters of unlike consonant segments or of unlike vowel segments in the morphophonemic representation, but most of these do not appear at the surface phonetic level, because they are subject to assimilation (or other processes, e.g. metathesis).

The long or geminate consonant segments that occur in West Greenlandic sometimes have a clearly heteromorphemic origin. There is a suffix meaning 'new' which occurs with initial /t/ in such forms as /irnirtaaq/ 'new son' (cf. /irniq/ 'son', which shows that the uvular element "/r/" is a variant of the stem final segment). When the suffix is added to stems ending in a non-uvular consonant, e.g. the stem in /kamik/ 'boot', we get /tt/ (ɔ: [t:]) by complete assimilation: /kamittaaq/ 'new boot'. This is an extremely common type of process with most types of suffix initial consonant.

There is, however, another major source of Greenlandic long/geminate consonant segments, namely the process generally referred to as gemination (see Bergsland 1955 p. 9-13). This process is apparently rather independent of morpheme boundaries. In many cases it affects a (synchronically) morpheme internal consonant which is geminated in connection with the addition of certain suffixes (that also often condition a modification of the final part of the stem), e.g. /nukaq/ 'younger sibling (of same sex)', plural /nukkat/; /sanawuq/ 'he works (with woodden material)', /sannawik/ 'a workshop'. (An attempt at a description of this process will be given in section 2.4. below.)

There is much disagreement among scholars on the cause of gemination as a historical process, and it is also a crux in synchronic phonology. It is obviously of importance for the appraisal of different explanations of this process to study various aspects of the phonetic relationship between forms with long and short consonant segments in Eskimo.

In the current phonemicization long/geminate consonants are denoted by double consonants (irrespective, of course, of their origin). In addition it posits two types of consonant clusters: (1) uvular fricative plus consonant, and (2) a unique cluster /ts/. It has, however, been noted by various scholars that the quantity relationships in these clusters are very peculiar.

In clusters such as /rn/[2] in /irniq/ 'son' the second segment of the alleged consonant cluster (i.e. /n/) clearly sounds long, and native speakers asked to utter the word slowly and chopped into syllables definitely tend to say "irn-niq" rather than "ir-niq". However, it is open to question whether different types of consonants behave differently in this respect. Moreover, it is open to question whether the post-uvular segment should be classified in systematic phonetic terms as a truly long segment (essentially similar to /nn/ in /inniq/ 'fire'), or whether it is only a somewhat prolonged variety of the short type of segment (as in /ini(q)/ 'room').

As for the cluster that is generally rendered as /ts/ (in e.g. /tatsit/ 'lakes', /natsat/ 'caps', /tuqutsiwuq/ 'he performs killing') the phonetically trained listener hears the first, i.e. occlusive, portion of it as long or geminate, whereas the fricative portion seems to be short. It thus makes sense to transcribe the auditory impression of this cluster as a long affricated [t:^{s}] or as a combination of stop plus affricate [t¢].[3] However, its rela-

---

2) The symbol r is generally used to denote a uvular fricative in Eskimo.

3) The symbol ¢ denotes a rather advanced affricate.

tionship to other types of short (single) and long (geminate) coronal stops is a problem.

The question of *short (single) or long (geminate)* in these clusters is of interest from a general theoretical point of view, because the phonological processes that apply to them are controversial. Before entering into this it is necessary to repeat that the clusters in question are not the only *underlying* clusters in the language. Long consonant segments often reflect a succession of two consonants separated by morpheme boundary, since in most types of clusters (except uvular plus consonant and *t+s*) we find a complete, regressive assimilation, e.g. *k+t* → *tt* ([t:]), *k+s* → *ss* ([s:]). Hence it makes sense to assume a priori that clusters like *r*(+)*n*, *t*(+)*s* should be comparable in duration to the ordinary long consonant segments occurring in the phonetic surface representation of West Greenlandic, since the underlying representations are of similar complexity.

As for the clusters of uvular plus consonant the general problem may be put quite briefly like this: are these clusters (1) subject to a peculiar phonological rule that lengthens the second segment, or are they (2) subject to a very ordinary assimilation rule according to which a greater or lesser part of the first segment is assimilated to the second (with regard to some or all features) and hence produces the effect of a prolongation of the latter? - As for /ts/ one might likewise suggest two phonological explanations of the longer closure hold: (1) there is a rule that lengthens stop before homorganic sibilant, (2) the sibilant is partly assimilated to the preceding homorganic stop.

It has been amply demonstrated in phonetic literature that the phonological categories "long" and "short", or "single" and "geminate", cannot be generally defined in terms of absolute values of physical duration. The actual duration of a segment is conditioned by a complex of factors having to do both with the context in which the segment occurs and with inherent features of the segment itself. Nevertheless, it can hardly be denied that a phonological grammar that assigns the values "long" and "short" to particular phonetic segments is more successful in doing so if there is a reasonable correlation between these terms and measurable quantities. If, now, we assume that the extra segment length in the clusters above is due to segment lengthening rules, then this account would seem to be corroborated if the total duration of these clusters clearly exceeds the duration of other entities of the same phonological size (i.e. if the total duration of, say, q+t -> rt considerably exceeds that of k+t -> tt, [t:], etc.). Conversely, the assumption that we have to do with a purely assimilatory phenomenon would seem to be corroborated if the total duration of the clusters in question is essentially similar to that of ordinary long (geminate) segments, such as [t:] from underlying k+t, etc.

As regards /ts/ the phonological situation is quite complex. There exists in the language an affricate [t^s] or, in an alternative transcription, [¢], which in all respects counts phonologically as a single, short segment. It is the obligatory reflex of underlying t before i (e.g. in such forms as /tii(q)/ 'tea', /atiq/ 'name', /aallartippaa/ 'he lets him travel away'), and this affrication is found also with the long stop segment arising from non-uvular consonant plus t before i (e.g. in /allatti/ 'secretary', with tt from underlying k+t). Obviously, t and s are neutralized in the position between t and i. Phonetically there is only one long segment or cluster beginning with a coronal stop in the position before i, which may be transcribed [$t{:}^s$] or [t¢]. It sometimes derives from underlying t(+)s and sometimes from underlying {t/k}(+)t, but there is no distinction between "/ts/" and "/tt/" before /i/ on an autonomous phonemic level (see section 2.2. below concerning the inadequacy of the current autonomous phonemicization). This double origin of the long coronal affricate must be taken into consideration in formulating the rules deriving it from t(+)s. To put it in a somewhat oversimplified way, one may ask whether tt becomes ts before i (and hence is affected by some, probably late, rules which modify the durational relationship between t and s in this cluster), or whether ts and tt become a long (geminate) affricated stop segment by entirely different rules (these very rules producing the neutralization).[4] It must clearly be of interest to check the phonetic adequacy of such different analyses by measuring whether the affrication of the long affricate is so similar to the affrication of the short affricate (derived from single t before i) that they can be considered the same articulatory gesture. If this appears to be true, it seems beyond dispute that we should derive the affrication of underlying tt and t before i by the same rule(s) rather than assume that /tt/ takes a different path (via ts) through the phonological rules. If, on the other hand, we are faced with two clearly different kinds or degrees of affrication, the other solution may not be a priori out of question.

The phonological problems mentioned above are clearly not just problems of the synchronic description but just as much problems of diachrony. What is the nature of the historical processes that have produced the present consonant pattern of Greenlandic Eskimo? Measurements of physical segment duration do not in themselves provide answers to such questions, but they may help to show whether the phonetic framework on which the phonological explanations are based, is a plausible model of what actually goes on.

---

4) The apparent third possibility: ts → tt/_i is out of question, since ts behaves essentially alike before all vowels.

The remainder of this paper contains first a more detailed presentation of the role of consonant length (gemination) and affrication in the phonology of modern West Greenlandic, arguments being adduced in favour of a particular conception. - After this follows the presentation of a rather extensive set of measurements of consonant duration in West Greenlandic, and finally a brief discussion of how the results agree or disagree with the phonetic model implied by the phonological rules.

2. Elements of West Greenlandic phonology

2.1. The traditional description: autonomous phonemic pattern

The phonemicization that is generally used in transcribing West Greenlandic and related dialects in linguistic papers and manuals, comprises the following inventory (disregarding some quite peripheral items):

vowels:	/a i u/
stops:	/p t k q/
nasals:	/m n ŋ/
sibilants:	/s ṣ/
other continuants:	/w l j g r/

Some of these symbols are sufficiently self-explanatory (in terms of the IPA standard) for the purpose of the present paper. Note, however, the following:

(stops:) /k/ is midpalatal/prevelar; /q/ is uvular.

(sibilants:) /ṣ/ (generally transcribed / ʃ /) is nonanterior.

(others:) /l/ is sometimes lateral, sometimes flapped or articulated as a slightly fricative tongue-tip [ɹ] ; /j/ (generally transcribed

/y/) is a palatal glide; /g/ (often transcribed /ɣ/ is a midpalatal/prevelar fricative; /r/ (often transcribed /ɣ̣/)[5] is a uvular fricative.

The most important[6] rules of allophonic variation according to this type of analysis are:

(aperture:) all vowels are lowered and retracted before uvulars; on the other hand, /a/ is raised before non-uvular consonants in the same syllable (i.e. before consonant clusters or word final consonants); in other positions /i u/ are (essentially) high, /a/ low.

(length:) in the cluster /ts/ the first consonant is long; in other clusters of two different consonants the second consonant is long.

(voice:) stops and sibilants are unvoiced; nasals are voiced; other consonants are voiced in those positions where they are manifested as short segments, but unvoiced elsewhere.

<u>The rules of combinability</u> of autonomous phonemes depend heavily on the interpretation of such sequences as [...i^{j}a...], [...u^{w}i...] (see section 2.2. below). If we assume (somewhat arbitrarily) that there is a consonant segment in this position, the rules for syllabification become very simple, and the other rules of combination become simple as well:

(syllabification:) Two vowel phonemes in succession form one syllable. The vowels of adjacent syllables (within

---

5) The presentation in Hill (1958 p. 420-421) is erroneous due to a confusion of the symbols used by Swadesh (1946).

6) Phenomena not directly relevant to the present paper are disregarded here.

one "word")[7] are separated by one or two consonants. The initial syllable of the word may begin with zero or one consonant, the final syllable may end with zero or one consonant.

(distribution of vowels:) All vowels occur single and geminate. There is one diphthong /ai/, which however occurs only in absolutely word final position.

(distribution of consonants:) All consonants (excluding /j/, which was defined above as a glide) occur single as well as geminate. Word initially only noncontinuants, i.e. stops and nasals (except /ŋ/), are found. Word finally only stops are found. In addition to the phonemic geminates there is a series of clusters consisting of /r/ plus consonants which occur intervocalically. The second constituent of these clusters may be any consonant, except that /k g/ as well as the glide /j/ do not combine with /r/. Finally there is a unique cluster /ts/ occurring in intervocalic position.

2.2. Some major inadequacies of the autonomous phonemic notation

In contemporary phonological work the interest has shifted very much from autonomous phonemics to models that can be integrated into a transformational generative grammar. It has been argued that autonomous phonemic descrip-

---

7) For practical reasons we use the straightforward term "word" (i.e. phonological word) instead of some other term such as "phrase" or "stretch" (Bergsland's term) in referring to such complex entities as /qiturnartaaŋŋu(w)aq/ 'a little newborn child'. From a phonological point of view, at least, they are no less words than such forms as loneliness, regretfully, or man-eater in English.

tions do not constitute a significant level of linguistic structure. Nevertheless, people still find it expedient to refer to some kind of broad, i.e. more or less genuinely autonomous phonemic, notation in discussing matters of phonology. In our opinion the expediency of such a level of reference is most conspicuous in the case of Eskimo, and we have chosen to use it in the later, experimentally oriented sections of this paper. It may, therefore, be well-motivated to consider why this paradoxical situation arises for languages in general and for Eskimo in particular.

Probably the most general argument against autonomous phonemes in a generative grammar is that phonological rules operate on single phonological features (such as [±voice]) rather than on phoneme-sized segments. This means that phonemes must at any rate be interpreted as abbreviations (ad hoc symbols) for feature matrices. In itself, however, this is no real break with earlier theories.

A more specific objection against autonomous phonemes is based on the fact that phonological rules seem to be at least partly ordered. This means that there is no particular level between beginning and end of a rule complex. It should be noted, however, that this does not a priori _exclude_ the possibility of devising a broad or autonomous phonemic notation whose symbols can be interpreted as abbreviations for feature matrices at some specific point in the sequence of phonological rules. If we consider the autonomous phonemic description of Greenlandic outlined above, it is rather evident (on closer inspection) that its allophonic (manifestation) rules would somehow reappear as late generative rules with no effect on deeper processes of the phonological component. This does not directly mean that the phoneme symbols reflect matrices as they appear at a particular break between "early" and "late" rules, but rather that they can be easily translated into such matrices. Therefore, the autonomous notation constitutes an expedient frame of reference in illustrations of early as well as late rules: in the former case it is possible to refer to the output of the rules without cumbersome and irrelevant indications of phonetic phenomena introduced by late rules (such as the aperture of Greenlandic /a i u/ or the voicedness of nonsibilant continuants), in the latter case it is possible to mention forms that satisfy the structural description of the late rules without indicating irrelevant high-level information. It may be theoretically preferable to have specific, letter-sized transcriptions for the input and output of every rule (as is widely done in Chomsky and Halle 1968), but this gives many technical problems and sometimes makes it rather difficult for readers not familiar with the subject-matter to keep track of the material that is used for documentation.

Now, in order for a broad or autonomous transcription to be at all meaningful as a frame of reference it must at least be a notation in which neutralized segments are indicated by one common symbol (Kiparsky 1968 takes the autonomous phonemic specification to mean essentially this). However, much of the work that was done previously in autonomous phonemics was based on a conception of the phoneme which did not allow for neutralization. Hence in cases of factual neutralization one or the other phoneme is indicated, often arbitrarily. This is also true of the phonemicization of Eskimo. A closer consideration of some of these cases neatly reveals the fallacy of the belief that such problems can be resolved by appeal to "symmetry of distribution".

As regards /j/ and /w/, the former is not distinct from zero at the surface level in the sequences /i(j)a/, /i(j)u/, and the latter is not distinct from zero at the surface level in the sequences /u(w)a/, /u(w)i/. These segments have a twofold origin: from underlying obstruents and as secondary segments due to a rule of glide insertion.

p. 186, l. 3-4, read: We follow Bergsland's (1955) practice of indicating these consonants (though without parentheses);

Since the glides are weak (often virtually inaudible) it is a problem in autonomous phonemics whether to represent them as segments or not. We follow Bergsland's (1955) practice of giving them in parentheses; from a generative point of view it can be shown that they are segments at a near-surface level, but it would be outside the scope of the present paper to go further into this.[8]

The affricate cluster /ts/ or /tt/ before /i/ poses another problem, which was outlined in section 1. above. Comparison with the behaviour of /t/ before /i/ suggests that we have /tt/ in forms like "/tatsit/" 'lakes', whereas comparison with the long affricate that occurs before /a/ in forms like "/natsat/" 'caps' suggests /ts/. From a generative point of view the long affricate has a double origin, but there is no obvious reason to claim that either ts → tt, or tt → ts, before i. Rather, if we define an affricate ¢, it is reasonable to claim that

ts → t¢ (in all environments)
tt → t¢/ __ i

by two different processes. If it is to be meaningful, the autonomous phonemic notation must represent the outputs of these two rules, and we therefore suggest that the cluster should be rendered as /t¢/. Since the latter part of the cluster auditorily resembles the manifestation of single /t/ before /i/, it would seem more consistent also to restate the analysis of the latter by splitting the autonomous phoneme "/t/" into two, viz. /t/ and /¢/ (a form like "/atiq/" 'name' being restated as /a¢iq/). Adherents of autonomous phonemics should note that /t/ and /¢/ are

---

8) Positing these segments allows us to refer to syllable number on a well-defined basis.

distinct in the environments /t__a/, /t__u/, cf. /attat/ 'button' versus /nat¢at/ 'caps', or /tuttut/ 'reindeers' versus /put¢ut/ 'clouds (of smoke or the like)'.

From a generative point of view the restatement of the single coronal stop as an affricate before the high front vowel implies that the second rule above is just a special case of the more general rule

$$t \rightarrow ¢/\_\_ \underline{i}$$

In the absence of evidence to the contrary it seems strongly motivated to claim that this is the case, i.e. that affrication of all occurrences of underlying t before i takes place before the point in the rules that is (roughly) reflected by the autonomous notation.

We therefore tentatively change the notation as suggested above, i.e. we contend that there is a segment ¢ with the same feature specification in both /a¢iq/ and /tat¢it/, or in the pair /na¢iq/ 'floor' - /nat¢iq/ 'a kind of seal'. The validity of this contention from an acoustic point of view will be tested below.

There remains one major inconsistency in the autonomous notation, viz. the uvular entity in such forms as /aqqa/ 'his name'. The non-uvular stops occur, according to the autonomous notation, both as geminates (/tt/ etc.) and in clusters with preceding /r/ (/rt/ etc.). This leaves us with a distributional problem: should the form above be transcribed /aqqa/ or /arqa/? The existence of this problem is, in our opinion, a symptom of the basic inadequacy of transcriptions like /rt/. A satisfactory solution implies a complete restatement of clusters in general, and it would thus remove the notation very much from current practice. We have preferred to leave the autonomous notation used as reference as it is on this point (i.e. we write /qq/ but /rt/), but the problem will be approached within a generative framework in the section below.

2.3. The regressive assimilation rule

As stated in section 1. above there is an assimilation rule which assimilates a non-uvular consonant to a following consonant within the same word. If ordered after the (progressive assimilation) rule that changes t+s into tś, the regressive assimilation rule applies without exception, as long as the first consonant is non-uvular, i.e. does not have the feature configuration

$$\begin{bmatrix} +\text{back} \\ -\text{high} \end{bmatrix}$$

(according to the feature theory of Chomsky and Halle (1968)).

Now it lies close at hand to assume that the (subjective impression of) length in the second segment of clusters like "/rp/", "/rt/", "/rn/" is a consequence of the same rule rather than due to a spurious lengthening rule. This would mean that the regressive assimilation applies also if the first segment is uvular, but only partially in this case. The difficulty with the formulation of such a partial assimilation rule is that it is not easy to determine on a purely auditory basis what is going on. The uvularity is signalled most clearly by the influence on the preceding vowel.

It holds true generally that vowels assume a highly different quality before uvulars than before non-uvulars, cf.

qimmiq	[qim:ə-q]	'dog'
qimmit	[qim:iṭ]	'do., pl.'
tuukkaq	[tu:k:ɑq]	'harpoon'
tuukkat	[tu:k:et]	'do., pl.'

and the (late) rules that introduce this modification ("uvularization") of vowel before uvular applies also when a vowel is followed by one of the controversial clusters exemplified above. It is not clear, on the other hand, how much is left of the uvular consonant segment.

In terms of ordered rules nothing prevents us from assuming that the "uvularization" of vowels takes place *before* regressive consonant assimilation. If this is true, the assimilation of a uvular to a following consonant may be complete without any loss of distinctiveness, the distinction between forms like "/irniq/" and "/inniq/" being carried over to the preceding vowel. However, nothing else seems to suggest that "uvularization" of vowels is an early rule. On the contrary, it must at any rate be later than the affrication rule (see section 2.2.). Affrication of *t* (*rt*, *tt*) takes place also before *i* modified by a following uvular, but not before the high variety of *a*, cf.

atiq [aøə-q] vs. attat [et:et]

This does not seem phonetically reasonable unless the affrication takes place *before the vowel modifications.*

The alternative, and immediately more attractive solution is that there is partial assimilation of uvular to a following consonant, but still a uvular component left to trigger the "uvularization" of the preceding vowel by a later rule. This may be conceived either in such a way that the initial portion of the uvular keeps intact (cf. the partial progressive assimilation of *s* to preceding *t*), or so that the segment is completely assimilated to the following except for the features defining its "uvularity" (in clusters like "/rt/" or "/rs/" this implies that we get a coronal stop or continuant with "uvularity" as an extra articulation - it must then be left to later rules to specify how such segments are realized in the phonetic output, be it with "pre-uvularization" or in some other way).

Though apparently less natural, the latter interpretation may seem to get support from another rule according to which *t* is assibilated to *s* after an underlying *i*. This latter rule (whose applicability has to do with very abstract properties of the underlying pattern) applies also if a consonant intervenes, cf. in conventional autonomous notation /suraarsippaa/ (from *sura+ir+t...*) 'he lets him discontinue his work' versus /aallartippaa/ (in our revised notation: /aallarȼippaa/) 'he lets him travel away'. *If* this is a rule conditioned by phonological features rather than morphological diacritics (which is disputable since is depends on the analysis of the underlying vowel system as a three-vowel or four-vowel system) it seems immediately most favourable for the formulation of this rule to have regressive assimilation precede it. The two forms above then have the coronal sequence ^{r}*tt* as input to the rule, and we get straightforward assibilation to ^{r}*ss* after *i* in /suraarsippaa/ (but later affrication of ^{r}*tt* to ^{r}*tȼ* before *i* in /aallarȼippaa/). However, there are good reasons for claiming that the alternation between stop and sibilant is *not* a phonological process.

Leaving open the question whether phonological simultaneity (in the output of the assimilation rule) of "uvularity" and other articulations implies ultimate phonetic simultaneity or successivity, we may tentatively formulate the assimilation rule without regard to the temporal relationships within the first segment of clusters. The formulation becomes rather involved since it must state that assimilation with regard to the features [±high] and [±back] fails to take place if the first segment has "minus" for the first and at the same time "plus" for the second feature:

$$\begin{bmatrix} -\text{syllabic} \\ \left\langle \left\{ \begin{matrix} [+\text{high}] \\ [-\text{back}] \end{matrix} \right\} \right\rangle \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\,\text{anterior} \\ \beta\,\text{coronal} \\ \gamma\,\text{continuant} \\ \delta\,\text{nasal} \\ \left\langle \begin{matrix} \varepsilon\,\text{high} \\ \theta\,\text{back} \end{matrix} \right\rangle \end{bmatrix} \Big/ \_\_\_ \begin{bmatrix} -\text{syllabic} \\ \alpha\,\text{anterior} \\ \beta\,\text{coronal} \\ \gamma\,\text{continuant} \\ \delta\,\text{nasal} \\ \varepsilon\,\text{high} \\ \theta\,\text{back} \end{bmatrix}$$

(It should be noted that the complexity of this rule may be due to inadequate specification of the features of underlying forms. It is likely that either the so-called velars or midpalatals should be specified as nonback, or the uvulars (Swadesh's "velars") are characterized by a special feature of extreme backness. Such a feature is, however, not available in the currently used feature theory. It would lead too far to discuss the evidence for either of the two proposals in the present paper.)

If we denote residual uvularity by an index letter $\underline{r}$, the derivation of forms like "/qajartaaq/" 'new kayak' can be represented like this:

qajaq+taaq $\longrightarrow$ qaja^{r}ttaaq

Note that the assimilation thus stated does <u>not</u> imply that the first segment has an intervening stage where it is a fricative /r/ of full segment length.

2.4. Syllable structure and consonant gemination

If we adopt the glide notation discussed in section 2.2., the semi-surface structure of Greenlandic wordforms can be given like this:

(C)V(V)(C)CV(V)(C)CV(V)(C)....

i.e. except for the first syllable every syllable contains the CV sequence, which may be followed by none or one vowel segment plus none or one consonant segment. This kind of structure may be generated by the syllable formula

$$[CV(V)(C)]_0$$

if we add the rule that word initial C may be specified as zero,[9] whereas initial C in other syllables is obligatory (see the discussion of glides in section 2.2.).

It is postulated by the formula that (i) an intervocalic consonant goes to the following syllable, and (ii) syllable initial or syllable final clusters do not occur: in case of intervocalic consonant clusters the first constituent goes with the preceding syllable and the second with the next.

The adequacy of this analysis is shown by the fact that vowels are influenced strongly by the first member of a following consonant cluster but not by a following intervocalic consonant, unless it is uvular. Word final consonants have the same effect as the first member of clusters, i.e. in both cases do we have a closed syllable. The syllable division is easily heard in slow speech, where there is loose contact before one intervocalic consonant but close contact before a cluster. The syllable division in clusters was noted already by Kleinschmidt, who for example indicated the pronunciation of /aggirpuq/ 'he comes' as <u>ag-gerp-poq</u> (Holtved 1964 p. 29). Note that this analysis applies also to the long consonants generated from uvular plus consonant.

As for "uvularization" of a vowel by a following uvular this phenomenon is found also when there is a syllable border according to the formula above, e.g. in forms like

---

9) Bergsland (1955 p.4) gives the taxonomically more adequate formula (C)V(V)[(C)CV(V)](C). However, it seems to us that the unit within square brackets has no relationship to any linguistically significant unit (such as the syllable). We therefore prefer the formula given above in spite of the need of accompanying restrictional rules.

/niqi/ 'meat'. However, the domain of the rule involved does not extend further than the syllable just before the uvular, cf. that /ii/, /aa/ are modified in their entirety in such forms as /¢iiq/ 'tea', /taaq/ 'shadow', but only /a/ is modified in /agijaq/ 'fiddle', and only /i/ in /tuwiq/ 'shoulder' (note that the controversial segments /j/ and /w/ define syllable boundary in these forms).

The impression of word rhythm in Greenlandic seems to depend on the specific "weight" of each syllable. According to Kleinschmidt (1851, also see Holtved 1964 p. 30) a vowel segment counts as 2, and a syllable final consonant as 1. Hence syllables have increasing weight as follows:

V	VC	VV	VVC
2	3	4	5

There is no lexically distinctive category of stress. A certain degree of stress is generally heard on the ultimate and/or the antepenultimate syllable (in longer words sometimes on earlier syllables as well), but the impression of stress placement seems rather dependent both on the specific weight of the syllables and on the intonation. As for the latter, the tonal contour of the word is closely bound to the vowel segments. For example, the typical contour before a non-final pause is characterized by nonlow tone on the antepenultimate and ultimate vowels, and low tone on the penultimate, no matter how they are distributed on syllables;[10] schematically:

(—) ‾ _ ‾ aawaa 'he fetches it'   (—) ‾ _ ‾ akiwara 'I answer him'

The tonal rules are as yet quite imperfectly understood, but there is evidence that the vowel segment (the "syllabic") plays at least as important a role as the syllable in prosodic rules.

---

10) A somewhat different presentation is given in Petersen (1970).

If we turn now to the segments that contribute to syllable "weight" in Kleinschmidt's sense, it would be expedient to have a common term for these. Although we have not demonstrated any prosodic relevance of these segments, we suggest mora as a classificatory term for them. We may then speak of a vocalic mora, i.e. (C)V(V)(C) (first vocalic mora) or (C)VV(C) (second vocalic mora) and of a consonantal mora, i.e. (C)V(V)C. (Opinions may differ as to the status of word final C with respect to the category "mora", see below.)

The validity of a category such as "mora" in Greenlandic Eskimo appears from the phenomena associated with gemination as a process in Greenlandic morphology. The explanation of this process is one of the most intricate problems in the grammar of the language.[11] Anyway, gemination often has the apparent effect of "compensating" for the deletion of one or several final segments before an affix, cf. the difference between the two plural forms below:

nuna 'country', pl. nunat
qiŋaq 'nose', pl. qiŋŋat

or the difference between the two derivations below:

urnippuq 'comes' - urniguppaa 'comes with it'
ȼiŋiwuq 'flies away' - ȼiŋŋuppaa 'flies away with it'

(stems: urnik versus tiŋi, of which the latter drops the final vowel and gets gemination of the preceding consonant instead).

Gemination also often has the overt effect of compensating for deletion of part of the affix, cf. the two

---

11) See Ulving (1953), Bergsland (1955), Petersen (1970) for some attempts at diachronic and synchronic explanations.
Also see footnote 13 below.

variant forms below:

sanawuq 'works (with wood)' -

$\left\{\begin{matrix}\text{sanaappaa}\\ \text{sannappaa}\end{matrix}\right\}$ 'makes something (of wood) for him'

(the underlying affix begins with u, which is assimilated to a in the first of the two alternative formations but deleted in connection with gemination of stem internal n in the latter).

Gemination is frequently accompanied by changes in the feature composition of the geminated segment (particularly alternation between fricative and stop, cf. /miiraq/ 'child', pl. /miiqqat/). There are even underlying segments which are entirely deleted unless they are geminated (such as underlying intervocalic g in /puuq/ 'bag', pl. /puggut/). The pattern of these alternations is not directly relevant to the quantity problem and will be disregarded here.

If we look closer at the "compensatory" effect of gemination it is interesting that it adds a mora in the sense defined above, and that this "compensation" happens together with the deletion of one (or several) mora. In the immediately transparent inflectional and derivational forms gemination does not normally occur if a consonant that is not mora-forming is dropped. The plurals of /umik/ 'hair of beard' and /inuk/ 'human being' are /umiit/, /inuwit/ with loss of k before the syllabic ending in both cases (and later glide insertion in the second example). Here the loss of a consonant entails no loss of mora, since the underlying full forms would be divided as follows into syllables:

u-mi-kit   i-nu-kit

with syllable-initial and hence not mora-forming k. In plural forms like /qiŋŋat/, on the other hand, a mora is deleted no matter whether we assume underlying it or t as the plural affix; if we construct an underlying form by adding the former affix shape we get a hypothetical syllable division

*qi-ŋa-qit

and if we construct it with the latter affix shape we get a hypothetical

*qi-ŋaqt

Deletion of qi in the former case means loss of a vocalic mora, viz. i, and deletion of q in the latter case means loss of a consonantal mora. Gemination of the stem internal consonant in both cases restores the mora number of the full plural forms:

qiŋ-ŋat

As regards the number of consonantal mora in the wordform, the effect of gemination with this type of plural formation is the same as the effect of adding an overtly syllable-forming ending in another type of plural formation, cf. (with hyphens still indicating the mechanical division into syllables):

tu-piq 'tent' - pl. (tup-qit →) tuq-qit (added: one mora)

In this word the second vowel of the stem - which may be defined as an underlying segment whose feature specification differs somewhat from that of a "normal i" - is syncopated with the effect that the preceding consonant p comes to stand in a position where it forms a mora.[12]

Thus, if we count mora, we find a high degree of regularity in noun inflection (though the pattern is variegated enough in other respects). There are two reservations that must be made to this. Firstly, the gemination rule has no effect if there is already a consonant cluster at the place where gemination should occur, i.e. in such forms as

---

12) Cf. the vowel deletion process in Japanese discussed by McCawley (1968), p. 115f.

qim-miq 'dog' - pl. qim-mit

and secondly it must be stated that the types of plural formation exemplified above are not distributed according to quite rigid principles in modern usage (one hears such forms as /tupit/ 'tents'; cf. below). Nonetheless, the addition of one mora in plural versus singular forms is found in a vast number of words.

The tendency that inflection or derivation with a particular affix morpheme entails a fixed increment of the mora number in spite of morphological variation, can be observed in several types of formation, cf. the (semantically equivalent) morphological variants (explained above):

sa-naap-paa }
san-nap-paa } identical number of mora

or the (semantically differentiated) variants /sannawik/ 'workshop' and /sanawwik/ 'place where to work':

san-na-wik }
sa-naw-wik } identical number of mora

In trying to explain the mechanism behind this surface regularity one must realise there is a skewness in the relationship between stems ending in a vowel and stems ending in a consonant. This is clearly seen if we compare /nuna/ and /qiŋaq/ inflected with a nonsyllabic and a syllabic ending:

singular	plural	stem + a̲ 'his'
nu-na	nu-nat	nu-naa
qi-ŋaq	qiŋ-ŋat	qi-ŋaa

If a word final consonant counts as a mora then we have the same increment of mora number in the two plural forms. If, on the other hand, the final consonant is considered irrelevant to the "weight" of the word, we may state that we have the same increment of mora number with a syllabic affix (note that a stem final consonant is retained or dropped before such an affix, depending on the vowel preceding it; this does not affect the mora number at all, since the consonant will become syllable initial anyway). It seems that deletion of an underlying consonant segment only counts as a mora loss if the non-deletion of the segment would result in a consonant cluster. (The deletion of q in the plural of qiŋaq is thus a loss of mora since its retention before plural t would result in the structurally impossible form *qiŋaqt.)

The tendency in contemporary Greenlandic goes in the direction of morphological simplification, including the gradual abolishment of certain forms with "irregularities" such as gemination. It may be foreseen that it will become increasingly more common to say /qiŋat/ 'noses' instead of /qiŋŋat/, etcetera. This can be explained as partly due to a formal ambiguity of singular forms in q. In essential respects this segment behaves as part of the stem, but it is clearly felt as a singular marker. If it is taken as part of the stem it is clear that the plural formation entails a mora loss which must be compensated for by gemination. If, on the other hand, the final consonant is interpreted as a member of a paradigm

$$\left\{\begin{matrix} \text{q} & \text{'singular'} \\ \text{(i)t} & \text{'plural'} \end{matrix}\right\}$$

it is obvious that the substitution of one ending for the other should not lead to any compensatory effects; it would be entirely parallel to the inflection

nuna - nunat

where t replaces another singular marker, viz. the absence of any affix.

As for the compensatory effect of consonant gemination it is not simple to determine with certainty whether it compensates for the deletion of an underlying mora-forming consonant or vowel in the forms cited above (Bergsland 1955 claims that it always accompanies syllable syncopation). We shall not go further into this here, nor attempt to formulate a synchronic syncopation rule as a phonological process (Underhill (1970) argues that gemination should be considered "morphological" in contemporary West Greenlandic).

3. Acoustic-phonetic investigation

As shown by the preceding sections there are three main questions concerning consonant duration in Greenlandic: (i) the relationship of "long" or "geminate" to "short" or "single" consonants, (ii) the relationship of rC-clusters to either of these types, and (iii) the relationship of the long coronal affricate to other segments or clusters. The remaining sections present some phonetic data bearing upon these questions.

3.1. Recording of material

The material for this investigation (see section 3.2. below) was recorded at the Institute of Phonetics (on Lyrec tape recorders) at several occasions in 1970-71. Informants were three male adults, Mr Robert Petersen, mag. art., amanuensis at the Institute of Eskimology ("RP"), Mr Carl Christian Olsen, teaching assistant at the Institute of Eskimology ("CCO"), and Mr Isak Heilmann, teacher ("IH"). The informants speak slightly different varieties of standard West Greenlandic as spoken in the central part of West Greenland. Besides being native speakers of Greenlandic all three informants have a scholarly insight into the language which has been of much help to us. The research presented in this paper was in fact inspired by discussions with them and with other participants at Mr Petersen's seminars on problems of Greenlandic grammar. We are highly indebted to them for this co-operation.

It is a well-known fact that durational measurements require a carefully selected material in order to be meaningful. First of all, items to be compared must of course be recorded under comparable conditions, uttered with comparable speed, etc. Even if this is so, there is a wide range of possible types of material, ranging from a piece of normal conversation as one extreme to a set of nonsense words as the other extreme. The former of these extremes is preferable if it is crucial to find the absolute average durations of segments in the particular language in running speech; the other extreme is preferable if the study is to throw light upon basic, general phonetic questions concerning articulatory gestures and their timing.

The present study does not satisfy any of these desiderata. It is in fact rather a prerequisite to such studies, the emphasis being exclusively on the basic questions presented above, i.e.: what is the difference or ratio between segments that are generally considered long and segments that are generally considered short? and: what is the status of certain controversial segments such as consonant after uvular or the coronal affricate - are these long or short in systematic phonetic terms as regards the oral closure phase?

We have not used nonsense words as material for this study though it would have simplified the first phase of the research. One reason for not doing so is that it would probably be difficult to read such a text aloud with natural tempo relationships among the sound segments. Greenlandic words are highly analysable, and in reading a strange wordform one is likely to make attempts to reinterpret it as something more familiar or reasonable, particularly with the current orthography, which the reader may not expect to be correctly used in a handwritten or typewritten text. Even with grammatical wordforms this problem may arise. In our text /tutuppuŋa/ 'I am covered with ingrained dirt' was occasionally misread as /tuttuppuŋa/ 'I have caught a reindeer'.

Another problem is that since Greenlandic words are often extremely long (the average syllable number seems to exceed 4 in normal, running prose) the use of short test words (i.e., ideally of the type VCV or VCCV, where the consonants to be studied intervocalically are put in a minimum frame) would not necessarily show the typical manifestation of the quantity contrast. It would at least be desirable to see also what happens in words of average length, rather than dealing only with the extreme case.

We have preferred to use natural words for the present study, and in order that the raw data should also throw some light upon the range of manifestation of the durational features we have measured a variety of wordforms of varying length. A logical follow-up would be to investigate the influence of word length and other factors on the values measured. We have devoted a few remarks to this problem here and there below, but a closer analysis of the data from this point of view has not yet been undertaken. In the present paper, on the contrary, we have set out to find out whether there is a certain constancy in the durational relationships <u>in spite of</u> variations in word structure.

The type of constancy one would look for in durational measurements is of course not to be found in the absolute values but in the difference or ratio between the durations of segment types. Hence the material was so arranged that there were minimal or sub-minimal pairs for the segment types to be compared. - The subjective impression of stress and tone contours suggests that in order for segments in analogous environments to be durationally comparable the two words should exhibit the same number of syllables, and each syllable pair should have the same number of vocalic and consonantal mora. We tried to avoid examples that did not exhibit syllable or mora isomorphy, but is was not always possible. We also attempted to find word pairs with identical vowel environments for the consonants to be compared; among the pairs of words that were subminimal in this respect some were left out because the degree of openness of a preceding or following vowel obviously influenced the consonant duration considerably, but others had to be included because we had not been sufficiently successful in finding good minimal pairs for the contrast in question. - Some of the word pairs in the material differ as to the presence or absence of a word initial consonant. The measurements showed that the durations of all segments of the first and second syllable are influenced by the presence or absence of this consonant, i.e. there is an apparent tendency toward similar overall duration of such sequences as /#CVCV.../ and /#VCV.../. A subminimal pair like /anija/ 'his pain': /mannija/ 'its egg' thus gives a somewhat erroneous figure because the segments of the latter word are somewhat shortened (in this case the difference /n/:/nn/ becomes less than it is in truly minimal pairs). Some of the recorded words had to be left out on these grounds. (For a recent study of tendencies toward constant rate of syllable production see Slis 1968.)

3.2. Word lists

The material consists of the word pairs listed in Table I.[13] The words are given in autonomous phonemic notation according to section 2. above. In the case of /tȼ/ (see section 2.2.) the morphophonemic and orthographic status is indicated in parentheses: "(tt)" means that the long affricate is derived from underlying t(+)t (in the Kleinschmidt orthography ′t, vt, or gt), "(ts)" means that it is derived from t+s or is the geminate counterpart to surface /s/ in forms like /tasiq/ - /tatȼit/ (orthographically ts). - The consonants measured are underlined.

---

13) Several of the word pairs were suggested by Isak Heilmann or Robert Petersen, the rest by one of the authors (JR). Although some of the latter examples may be slightly old-fashioned or unusual in some other respect, none of the wordforms were deemed phonologically abberant at all.

The original, randomized lists were composed at different times. Due to an oversight three word pairs (i.e. 6 words in all) occurred twice in the word lists recorded, and separate averages of the measures were made. We have left them in, but they are marked as repetitions in Table I. -(Cf. section 3.4. below.) - Note also that there are several words (occurring once only in the randomized lists) which enter into two, sometimes even three different contrastive pairs, e.g. /t/ in /niqiturpuq/ is contrasted to /tt/ in /nipitturpuq/, /rt/ in /amirturpuq/, and /tȼ/ in /akitȼurpuq/. Thus the recorded material is not as huge as it looks in Table I.

In a couple of instances the words in CCO are numbered differently than in RP and IH, because some few words were not spoken by CCO. These differences are indicated in Table I.

3.3. Phonetic analysis

All the material was registered on mingograms displaying a "duplex oscillogram", a fundamental frequency curve, and two, three, or mostly four intensity curves (normally with an integration time of 5 ms) for the purpose of temporal segmentation.[14] As regards the signal to the intensity meters different kinds of frequency filtering (highpass, lowpass, or bandpass filtering with different cutoff frequencies) were employed in order to facilitate the segmentation, which was found to be difficult particularly with /n/. - Examples of mingograms are shown in Figs. 1-7.

As for the strategy of segmentation we did our best to define the starting-point of the stop or nasal in such a way that no possible element of fricative or vocalic "r" in the sequences /rp, rt/ etc. was included. The oral clo-

---

14) The devices employed (mingograph, fundamental frequency meter, intensity meter, and filters) are surveyed in ARIPUC 4/1969 (197o), p. IIff.

TABLE I

List of word pairs

(Numbers in parentheses refer to CCO's word list only.)

/pp/ - /p/

1. /ippak/ - /ipak/
2. /napparput/ - /napapput/
3. /¢illuppallaqaa/ - /¢illupallaqaa/
4. /nappaqqut/ - /napaqqut/
5. /uppippuq/ - /upippuq/
6. /¢ippapput/ - /¢ipapput/
7. /aluppaat/ - /alupaat/
8. /uqaluppaluppuq/ - /uqalupiluppuq/
9. (8) /anippallappuq/ - /anipallappuq/
1o. (9) /nappaqqut/ - /napaqqut/ (repetition of pair No.4)
11.(1o) /tappipput/ - /tapipput/
12.(11) /apputaa/ - /aputaa/
13.(12) /qamippaa/ - /qalipaa/
14.(13) /ippirtuwuq/ - /ipirtuwuq/
15.(14) /kamippak/ - /qalipak/
16.(15) /ulappupput/ - /ulapipput/
17.(16) /aappaluppuq/ - /taapalippuq/
18.(17) /aapparput/ - /aapapput/

/rp/ - /p/

1. /nirpiqarpuq/ - /nipiqarpuq/
2. /a¢irpallarijarlugit/ - /a¢ipallarijarlugit/
3. /alurpaat/ - /alupaat/
4. /anirpallappuq/ - /anipallappuq/
5. /illarpalaarpuq/ - /illapalaarpuq/
6. /sirpiqarpuq/ - /¢ipiqarpuq/
7. /tarparpuq/ - /tapappuq/
8. /uqarpugut/ - /aqupiwuq/

TABLE I (continued)

/pp/ - /rp/

1. /nipaappuq/ - /niparpuq/
2. /uppik/ - /urpik/
3. /nipipput/ - /nipirput/
4. /siippuq/ - /siirpuq/
5. /u¢ipput/ - /u¢irput/
6. /tuttupput/ - /tutturput/
7. /miluppaa/ - /milurpaa/
8. /uppiŋŋuwaq/ - /urpiŋŋuwaq/
9. /tukkupput/ - /tukkurput/
1o. /tuuppaa/ - /tuurpaa/
11. /miluppara/ - /milurpara/
12. /aluppaat/ - /alurpaat/
13. /uqaluppaluppuq/ - /sijanirpaluppuq/
14. /anippallappuq/ - /anirpallappuq/
15. /asaluppuq/ - /asalurpuq/
16. /qaappuq/ - /qaarpuq/

/tt/ - /t/

1. /nunatta/ - /nunataq/
2. /tuttuppuŋa/ - /tutuppuŋa/
3. /¢ikittarput/ - /¢ikitarput/
4. /puuguttaminik/ - /puugutaminik/
5. /naalattuwarpara/ - /asatuwarpara/
6. /mattuwuq/ - /matuwuq/
7. (5) /¢ikittarpuq/ - /puwisitarput/
8. (6) /nipitturpuq/ - /niqiturpuq/
9. (7) /kasuttakkat/ - /kutappuq/
1o. (8) /imaattuwinnarpuq/ - /kaatuwinnarpuq/
11. (9) /pilattuut/ - /silatuut/
12.(1o) /attappaa/ - /katappaa/
13.(11) /mattuwuq/ - /matuwuq/
14.(12) /atturpaa/ - /aturpaa/

<u>TABLE I</u> (continued)

/rṯ/ - /ṯ/

1. /tuqqurṯaq/ - /uqquṯaq/
2. /imirṯarput/ - /puwiṣiṯarput/
3. /amirṯurpuq/ - /niqiṯurpuq/
4. /aturṯarpuq/ - /kuṯappuq/
5. /imaarṯuwinnarpuq/ - /kaaṯuwinnarpuq/
6. /tarṯunaq/ - /maṯuwuq/
7. /arṯurpaa/ - /aṯurpaa/
8. /purṯuwuq/ - /puṯuwuq/

/ṯ¢̲/ (ts) - /ṯ/ (comparison of stop intervals)

1. /aaṯ¢̲aat/ - /aaṯaat/
2. /iriniṯ¢̲apput/ - /puwiṣiṯarput/
3. /akiṯ¢̲urpuq/ - /niqiṯurpuq/
4. /kuṯ¢̲appuq/ - /kuṯappuq/
5. /qaṯ¢̲urpuq/ - /aṯurpuq/
6. /puṯ¢̲uwalippuq/ - /puṯuwijarpuq/
7. /aaṯ¢̲arput/ - /aaṯarput/
8. /akiṯ¢̲urpuq/ - /niqiṯurpuq/ (repetition of pair No. 3)
9. /siwiṯ¢̲uut/ - /silaṯuut/

/ṯ¢̲/ - /¢̲/ (comparison of stop and fricative intervals)

(a) /t¢/ = "ts"

1. /naaṯ¢̲irpaa/ - /naa¢̲irpai/
2. (1) /¢ikiṯ¢̲irpara/ - /kuul¢i¢̲irpara/
3. (2) /naṯ¢̲iq/ - /na¢̲iq/
4. (3) /puṯ¢̲irpuq/ - /u¢̲irpuq/
5. (4) /inaṯ¢̲it/ - /iga¢̲it/
6. (5) /naaṯ¢̲irpaa/ - /naa¢̲irpai/ (repetition of pair No. 1)

(b) /t¢/ = "tt"

7. /masaṯ¢̲irpara/ - /kuul¢i¢̲irpara/ (second word same as in 2.)
8. (6) /kiṯ¢̲irut/ - /qi¢̲irut/

TABLE I (continued)

/tt/ - /rt/

1. /inuttut/ - /inurtut/
2. /uqquttarpaat/ - /tuqqurtarpaat/
3. /¢ikittarpuq/ - /imirtarput/
4. /nipitturpuq/ - /amirturpuq/
5. /kasuttakkat/ - /aturtarpuq/
6. /imaattuwinnarpuq/ - /imaartuwinnarpuq/
7. /mattuwuq/ - /tartunaq/
8. /atturpaa/ - /arturpaa/
9. /paatturluni/ - /paarturluni/
10. /taatturippuq/ - /naartulirpuq/

/tt/ - /t¢/(ts) (comparison of stop intervals)

1. /¢ikittarpuq/ - /irinit¢apput/
2. /nipitturpuq/ - /akit¢urpuq/
3. /kasuttakkat/ - /kut¢appuq/
4. /pilattuut/ - /siwit¢uut/
5. /paatturluni/ - /aat¢urluni/
6. /misigittarpuq/ - /irinit¢appuq/

/rt/ - /t¢/(ts) (comparison of stop intervals)

1. /imirtarput/ - /irinit¢apput/
2. /amirturpuq/ - /akit¢urpuq/
3. /aturtarpuq/ - /kut¢appuq/
4. /paarturluni/ - /aat¢urluni/

/qq/ - /q/

1. /quuqqut/ - /quuqut/
2. /aniqquwaa/ - /aniquwaa/
3. /tuqqurtaq/ - /tuqutaq/
4. /aqqarpaa/ - /aqarpaa/

TABLE I (continued)

5. /nuwiqqarpuq/ - /nuwiqarpuq/
6. /aallaqqugumma/ - /aallaqugumma/
7. /nunaqqa¢iqarpuq/ - /nunaqa¢iqarpuq/
8. /qulissirasuwaqqullugu/ - /qulissirasuwaqullugu/
9. /uqquttarpaat/ - /tuquttarpaat/
1o. /piqqarpuq/ - /piqarpuq/
11. /aqquppaa/ - /aquppaa/
12. /iqqaawuq/ - /iqaarpuq/

13. /siniqqarpuŋa/ - /iniqarpuŋa/
14. /nijaqqugut/ - /nijaqurut/
15. /iqqurpuq/ - /iqurpuq/
16. /tuqqugaq/ - /tuqutaq/
17. /uqqippuq/ - /uqippuq/
18. /siniqqarpuq/ - /iniqarpuq/

/nn/ - /n/

1. /maanna/ - /maana/
2. /sinniq/ - /piniq/
3. /alijannaarpuq/ - /alijanaappuq/
4. /akunnappuq/ - /takunarpuq/
5. /anniq/ - /maniq/
6. /inniqarpuq/ - /iniqarpuq/
7. /unnirluni/ - /unilluni/
8. /maanniq/ - /taaniq/
9. /annippaa/ - /anippaa/
1o. /uunnappuq/ - /uunarpuq/

/rn/ - /n/ (comparison of nasal intervals)

1. /maarniq/ - /maana/
2. /pirniq/ - /piniq/
3. /agijarnarpuq/ - /alijanaappuq/
4. /takurnarpuq/ - /takunarpuq/

TABLE I (continued)

5.	/marniq/ - /maniq/
6.	/irniqarpuq/ - /iniqarpuq/
7.	/urnilluni/ - /unilluni/
8.	/qaarniq/ - /taaniq/

/nn/ - /rn/ (comparison of nasal intervals)

1.	/maanna/ - /maarniq/
2.	/sinniq/ - /pirniq/
3.	/alijannaarpuq/ - /agijarnarpuq/
4.	/akunnappuq/ - /takurnarpuq/
5.	/anniq/ - /marniq/
6.	/inniqarpuq/ - /irniqarpuq/
7.	/unnirluni/ - /urnilluni/
8.	/maanniq/ - /qaarniq/

sure hold, and the explosion or affrication were measured separately for stops before the total duration was calculated as the sum of the two. In the case of single /q/ the point of explosion was often impossible to determine since the whole consonant was more or less fricative. It seems to be a general difference between /q/ and /qq/ that the single consonant is much more loosely articulated. The total duration could, however, be measured with a fair degree of accuracy, in spite of the unstable character of the oral articulation. There were no similar problems with the other types.

The measurements of consonant duration was made by one of the authors (HM) according to "conventions" for segmentation agreed upon. Random checks on examples measured by both authors showed generally a good agreement; in a minority of cases (less than ten per cent of the single word tokens) there was a disagreement amounting to one cs or (occasionally) more. This means that there is a margin

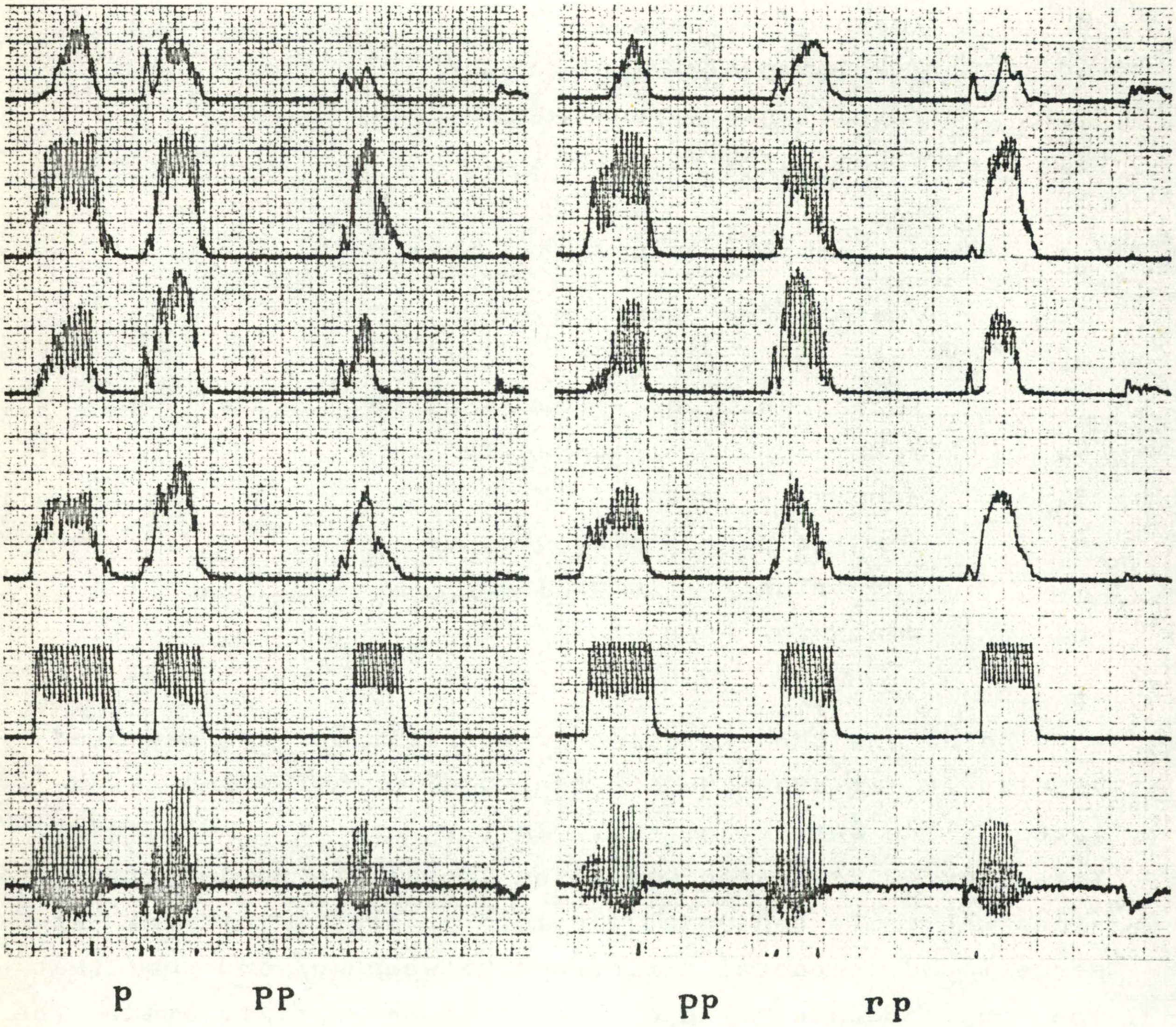


Fig. 1.

Words: /napapput/ and /napparput/. Speaker: RP.

Traces from top to bottom: four intensity curves with different filtration (integration time: 5 ms), fundamental frequency curve, and duplex oscillogram.

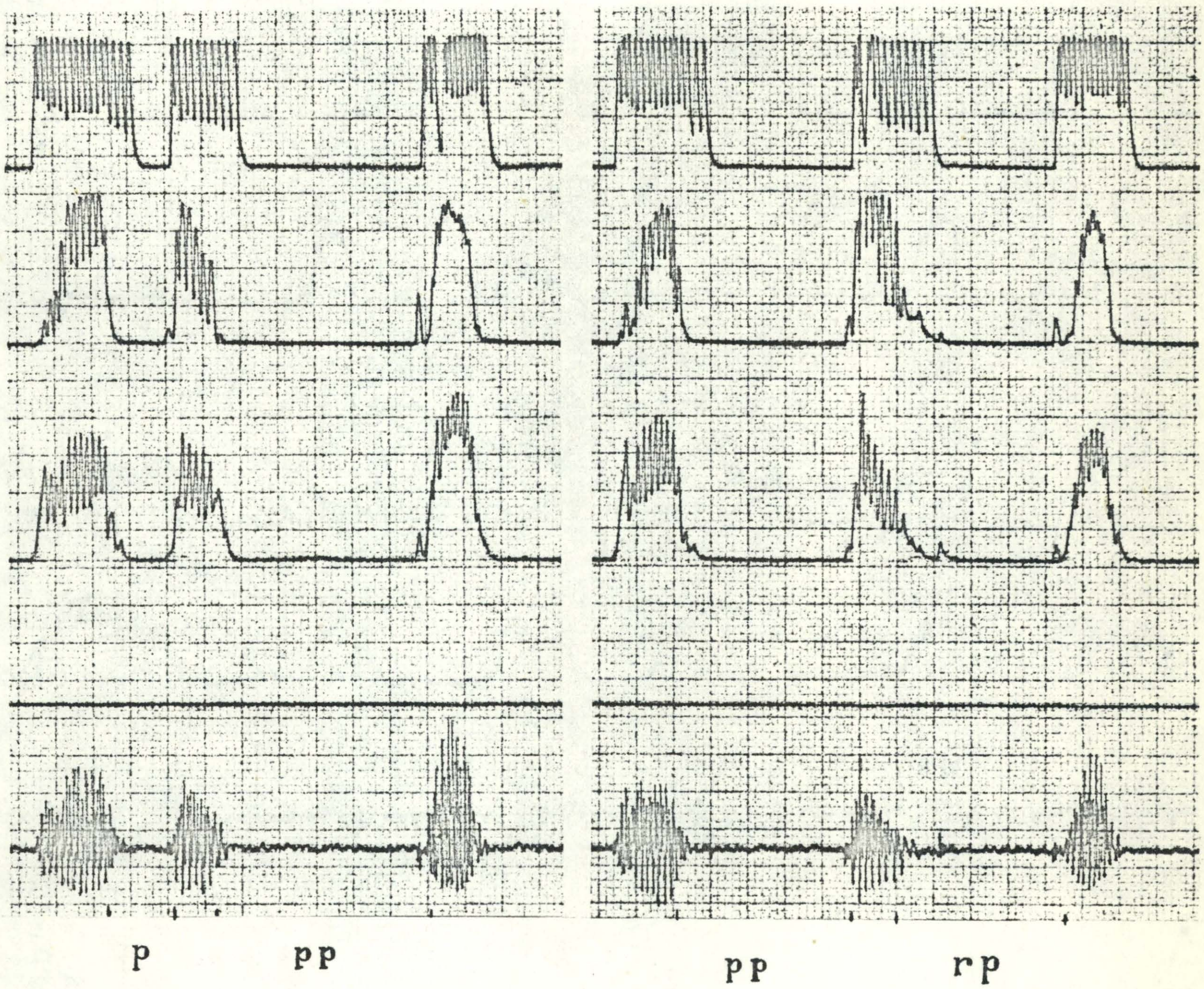
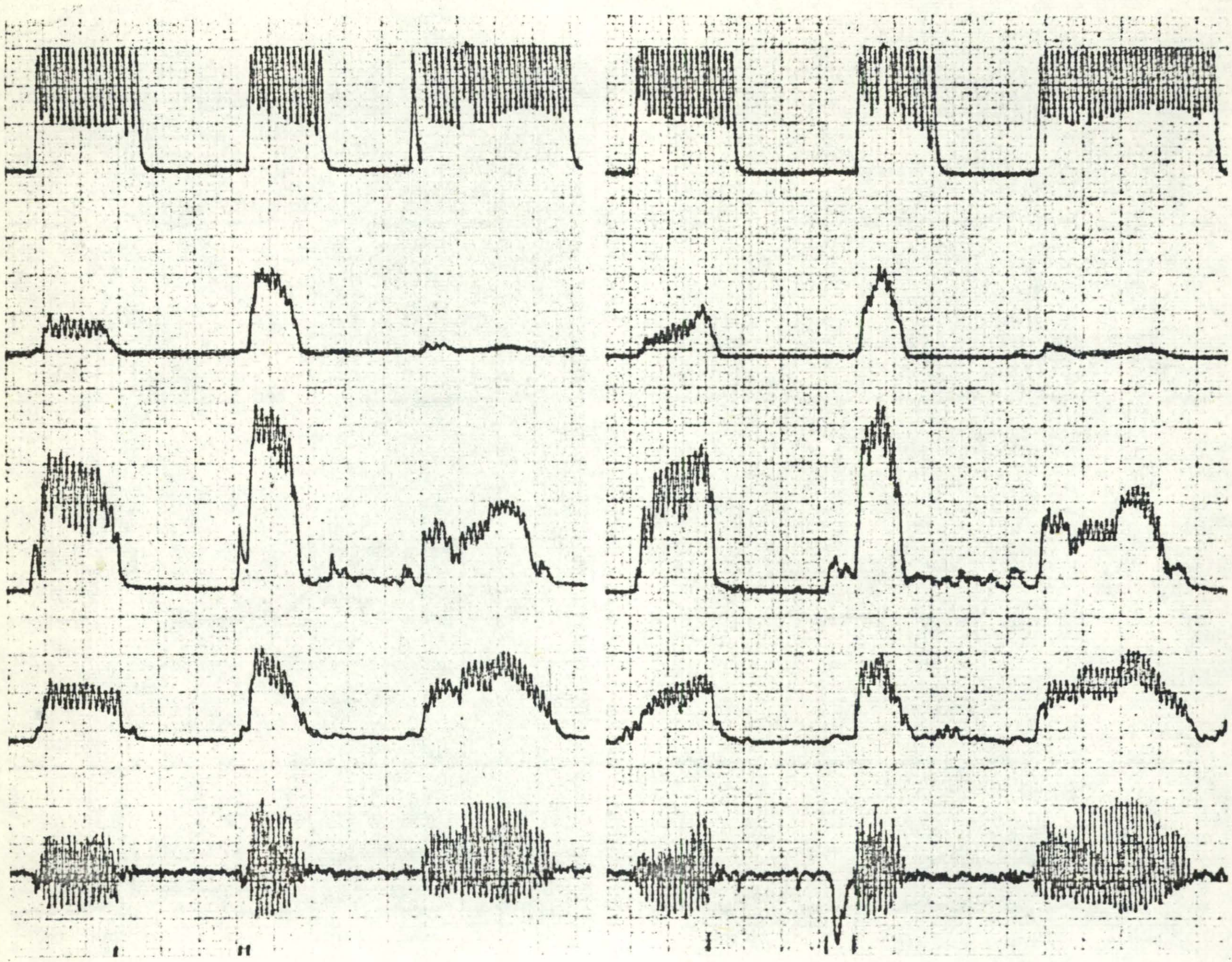


Fig. 2.

Words: /napapput/ and /napparput/. Speaker: IH.
Traces from top to bottom: fundamental frequency curve, two intensity curves with different filtration (integration time: 5 ms), and duplex oscillogram.



tt	t¢

Fig. 3.

Words: /paatturluni/ and /aat¢urluni/. Speaker: RP.
Traces from top to bottom: fundamental frequency curve, three intensity curves with different filtration (integration time: 5 ms), and duplex oscillogram.

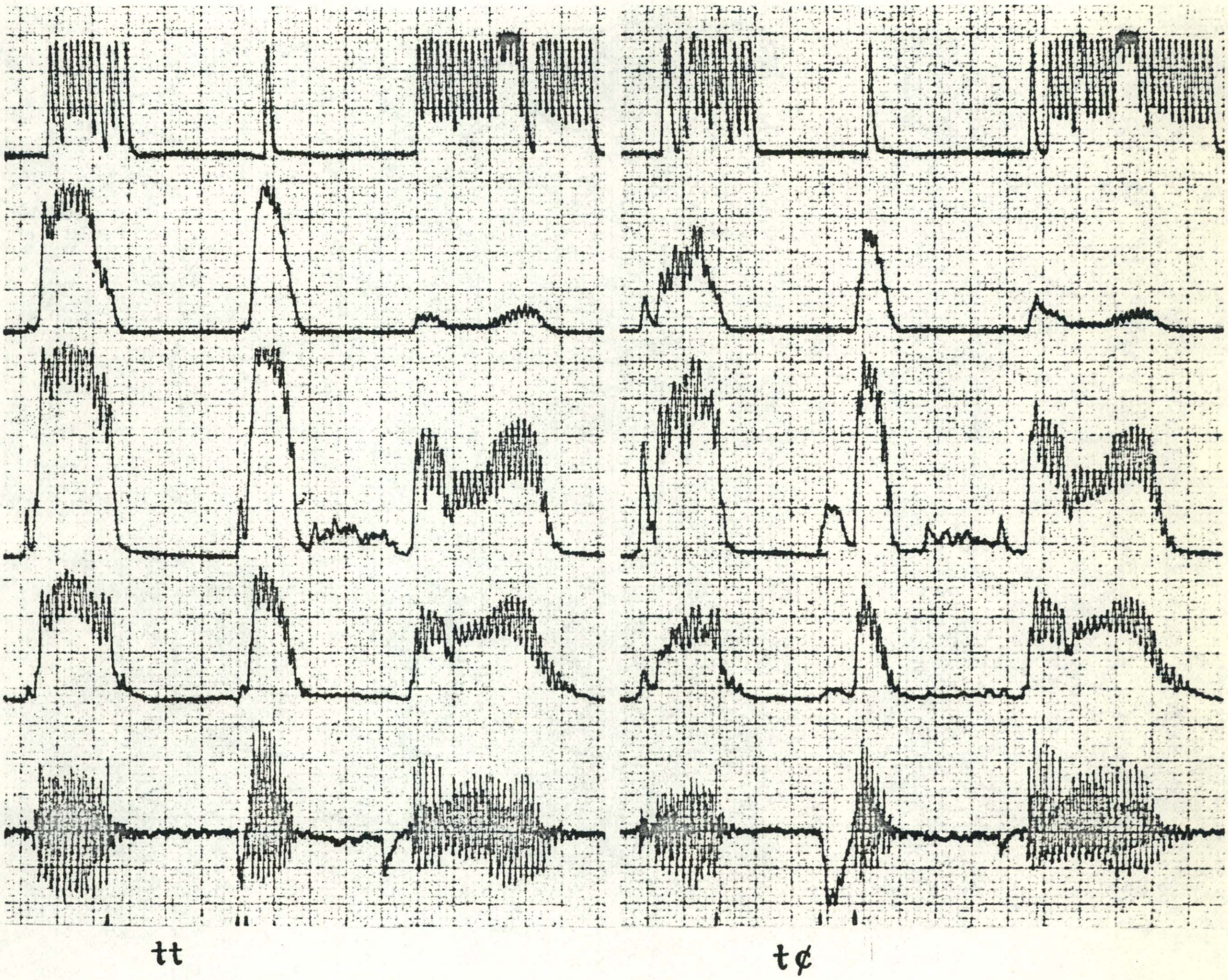


Fig. 4.

Words: /paatturluni/ and /aat¢urluni/. Speaker: CCO.

Traces from top to bottom: fundamental frequency curve, three intensity curves with different filtration (integration time: 5 ms), and duplex oscillogram.

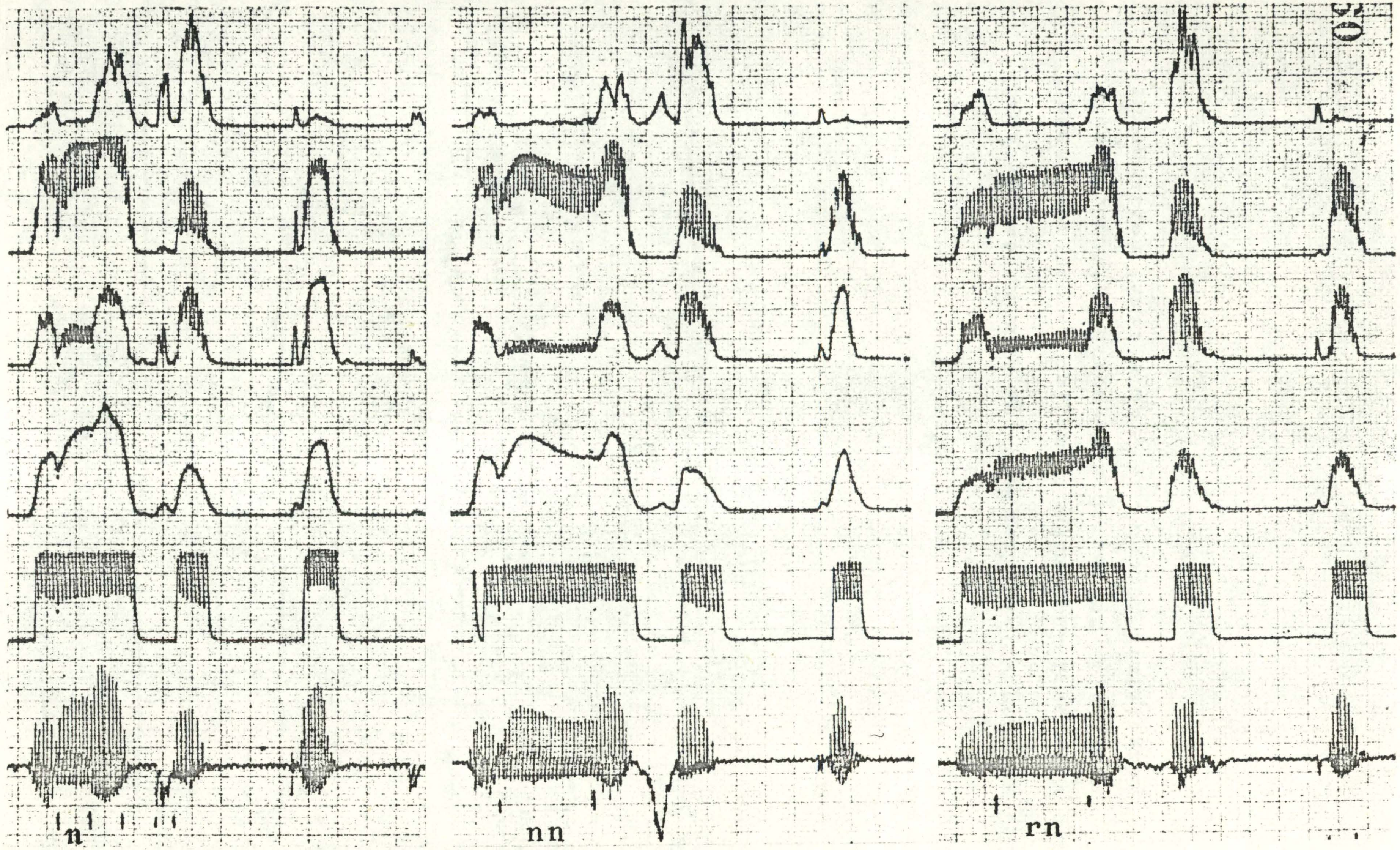


Fig. 5. Words: /iniqarpuq/, /inniqarpuq/, and /irniqarpuq/. Speaker: RP.
As for traces, see Fig. 1.

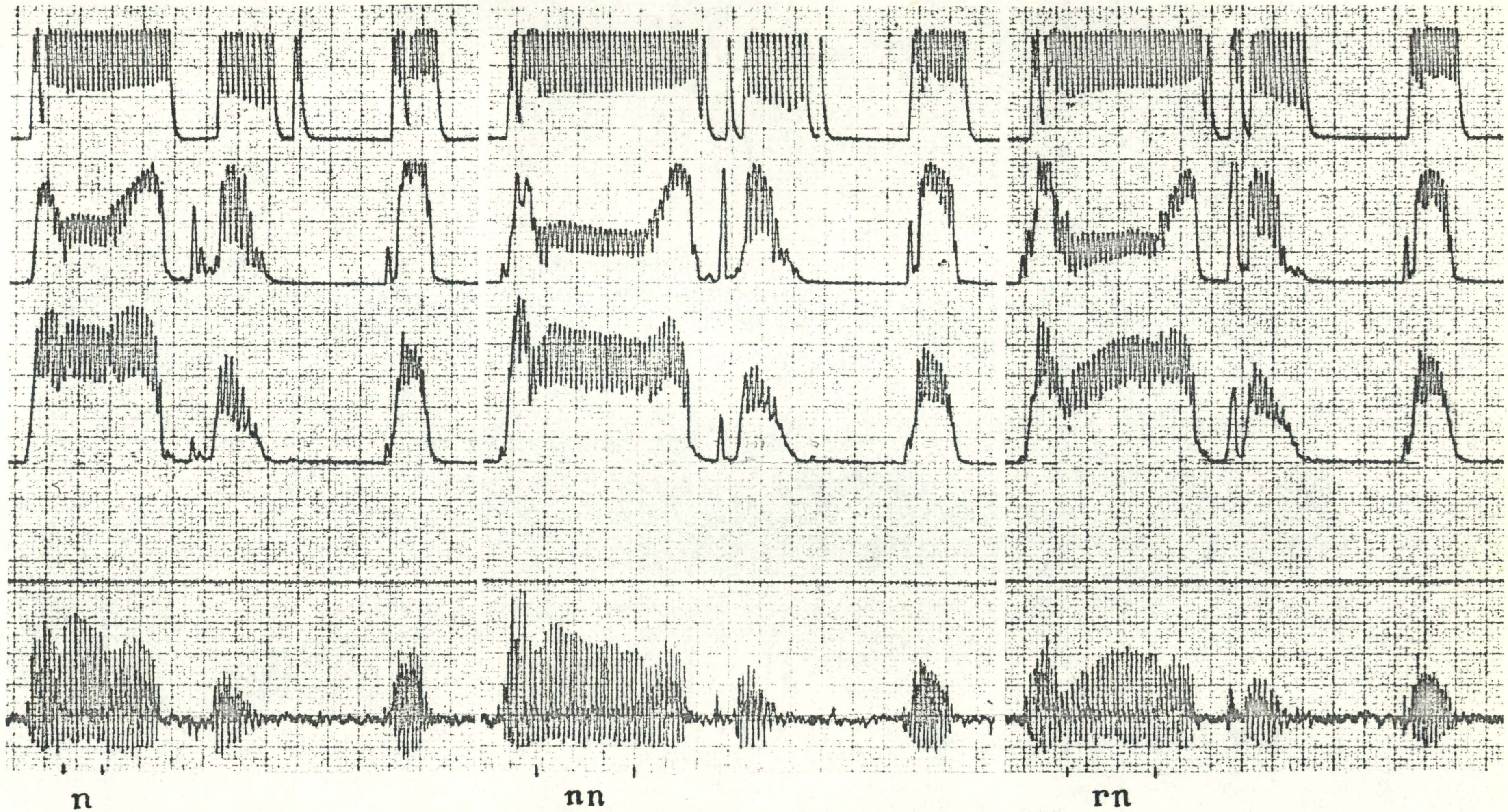


Fig. 6. Words: /iniqarpuq/, /inniqarpuq/, and /irniqarpuq/. Speaker: IH.
As for traces, see Fig. 2.

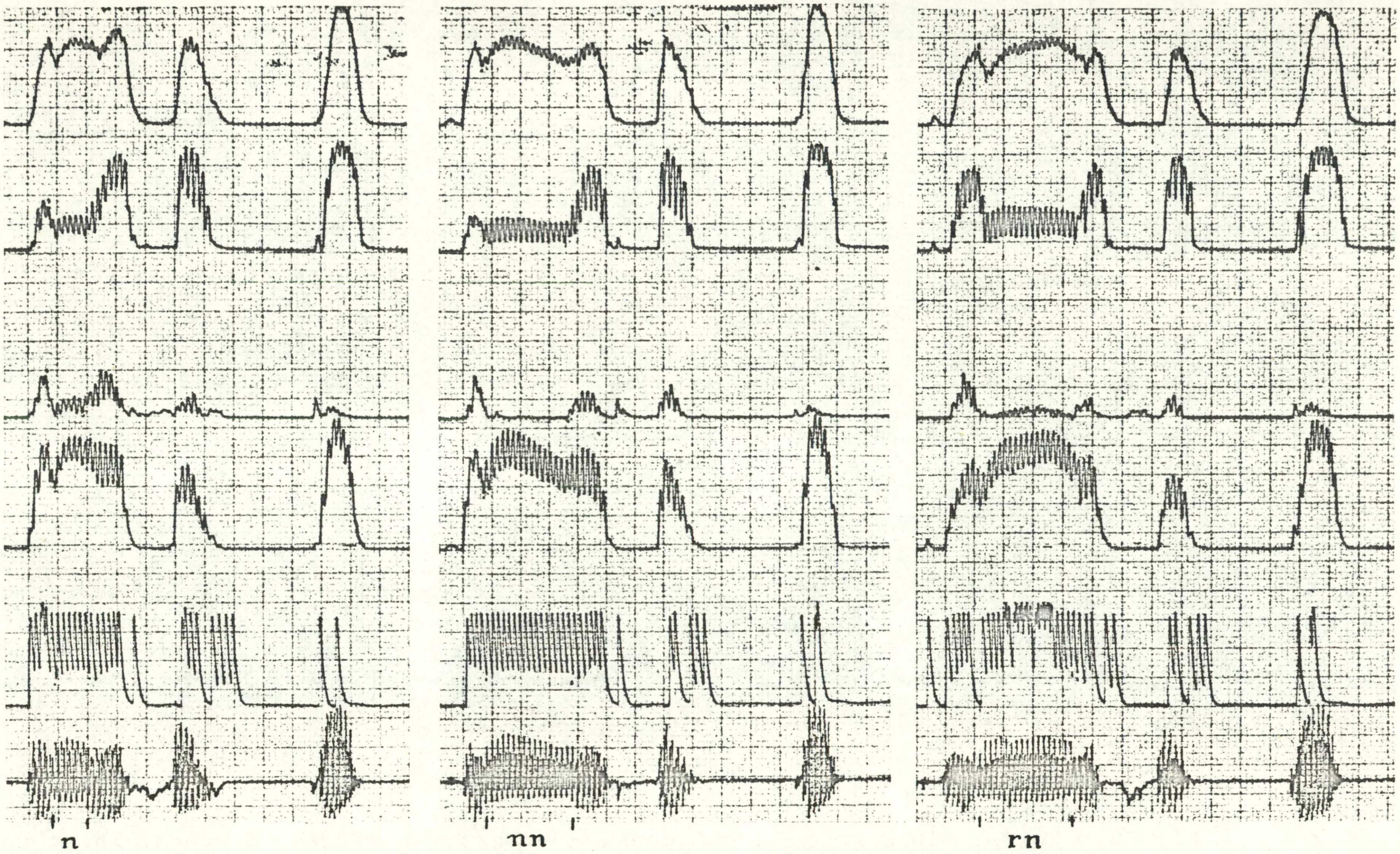


Fig. 7. Words: /iniqarpuq/, /inniqarpuq/, and /irniqarpuq/. Speaker: CCO.
As for traces, see Fig. 1.

of error (which can hardly be avoided since several instances are genuinely problematic due to imperfect mouth closure or other factors which are more or less impossible to recognize on the basis of acoustic curves). However, the error is diminished due to the fact that every bit of data entered in the tables represents an average of typically five individual measures.[15]

3.4. Results

The average measures of the duration of stops and nasals are given in Tables II, III, and IV. The printed symbols are to be understood as follows:

RN VERSUS N etc.: measurements on word pairs differing as stated (in casu /rn/ versus /n/ intervocalically), cf. Table I

RP, QQ, TS, T, C, etc.: duration of (what appears to be) oral closure plus explosion/affrication in wordforms with these phonemic representations (TS = /t¢/, C = /¢/)

RN, NN, N: duration of nasal consonant in wordforms with these phonemic representations

RT/T, etc.: ratio between two measures

RT-T, etc.: difference in centiseconds between two measures

3.5. Discussion

The main part of this section will be devoted to the durations of the stop and nasal segments with which the paper is specifically concerned. The durational characteristics of their environments will be more briefly treated later.

---

15) Word tokens that were obviously misread were of course excluded. The total number of such omissions is very small.

TABLE II

Average measures for each word pair: absolute durations in cs, differences, and ratios. Subject: RP.

SUBJECT RP — CONSONANTS PP VERSUS P

PAIR	PP	P	PP-P	PP/P
1	26.5	10.9	15.6	2.43
2	19.3	8.2	11.1	2.35
3	17.6	8.3	9.3	2.12
4	20.3	9.6	10.7	2.11
5	21.8	8.3	13.5	2.63
6	20.6	8.0	12.6	2.57
7	26.0	12.3	13.7	2.11
8	16.9	9.4	7.5	1.80
9	18.3	8.6	9.7	2.13
10	19.8	7.8	12.0	2.54
11	23.0	10.2	12.8	2.25
12	19.4	10.2	9.2	1.90
13	25.0	11.7	13.3	2.14
14	20.3	9.0	11.3	2.26
15	20.6	10.3	10.3	2.00
16	17.4	8.2	9.2	2.12
17	16.3	8.1	8.2	2.01
18	20.0	8.6	11.4	2.33

SUBJECT RP — CONSONANTS RP VERSUS P

PAIR	RP	P	RP-P	RP/P
1	18.2	9.1	9.1	2.00
2	17.7	8.8	8.9	2.01
3	16.7	9.0	7.7	1.86
4	17.3	8.8	8.5	1.97
5	19.1	7.5	11.6	2.55
6	20.8	10.0	10.8	2.08
7	22.9	12.3	10.6	1.86
8	18.7	8.6	10.1	2.17

TABLE II - continued

SUBJECT RP CONSONANTS PP VERSUS RP

PAIR	PP	RP	PP-RP	PP/RP
1	20.3	20.0	0.3	1.01
2	26.5	25.5	1.0	1.04
3	22.6	22.8	-0.2	0.99
4	21.4	21.5	-0.1	1.00
5	21.5	22.4	-0.9	0.96
6	21.1	22.8	-1.7	0.93
7	24.6	25.5	-0.9	0.96
8	22.1	23.3	-1.2	0.95
9	21.8	22.5	-0.7	0.97
10	24.8	23.5	1.3	1.06
11	19.4	21.2	-1.8	0.92
12	26.0	22.9	3.1	1.14
13	16.9	16.2	0.7	1.04
14	18.3	18.7	-0.4	0.98
15	18.8	19.6	-0.8	0.96
16	20.6	20.4	0.2	1.01

SUBJECT RP CONSONANTS TT VERSUS T

PAIR	TT	T	TT-T	TT/T
1	22.4	8.9	13.5	2.52
2	19.9	9.0	10.9	2.21
3	19.0	7.5	11.5	2.53
4	18.6	8.6	10.0	2.16
5	19.5	10.8	8.7	1.81
6	21.4	10.3	11.1	2.08
7	19.4	7.4	12.0	2.62
8	19.6	7.8	11.8	2.51
9	21.5	7.6	13.9	2.83
10	17.3	11.1	6.2	1.56
11	23.5	11.5	12.0	2.04
12	19.2	7.5	11.7	2.56
13	20.2	10.4	9.8	1.94
14	20.5	8.7	11.8	2.36

TABLE II - continued

SUBJECT RP CONSONANTS RT VERSUS T

PAIR	RT	T	RT-T	RT/T
1	20.1	7.3	12.8	2.75
2	19.0	7.4	11.6	2.57
3	20.3	7.8	12.5	2.60
4	19.4	7.6	11.8	2.55
5	19.3	11.1	8.2	1.74
6	20.6	10.4	10.2	1.98
7	21.2	8.7	12.5	2.44
8	20.1	10.4	9.7	1.93

SUBJECT RP CONSONANTS TS VERSUS T

PAIR	TS	T	TS-T	TS/T
1	19.6	7.4	12.2	2.65
2	21.0	7.8	13.2	2.69
3	19.8	7.6	12.2	2.61
4	25.9	13.1	12.8	1.98
5	24.0	11.5	12.5	2.09
6	19.7	8.4	11.3	2.35
7	20.3	10.3	10.0	1.97
8	18.9	9.0	9.9	2.10
9	19.8	7.2	12.6	2.75

SUBJECT RP CONSONANTS TS VERSUS C

PAIR	TS	C	TS-C	TS/C
1	21.3	11.5	9.8	1.85
2	20.5	8.4	12.1	2.44
3	24.2	11.6	12.6	2.09
4	21.6	9.5	12.1	2.27
5	21.8	12.2	9.6	1.79
6	19.6	10.8	8.8	1.81
7	19.0	8.4	10.6	2.26
8	21.5	13.2	8.3	1.63

TABLE II - continued

SUBJECT RP — CONSONANTS — TT VERSUS RT

PAIR	TT	RT	TT-RT	TT/RT
1	22.0	22.4	-0.4	0.98
2	18.3	17.8	0.5	1.03
3	19.4	19.0	0.4	1.02
4	19.6	20.3	-0.7	0.97
5	21.5	19.4	2.1	1.11
6	17.3	19.3	-2.0	0.90
7	20.2	20.6	-0.4	0.98
8	20.5	21.2	-0.7	0.97
9	17.5	18.6	-1.1	0.94
10	16.2	16.3	-0.1	0.99

SUBJECT RP — CONSONANTS — TT VERSUS TS

PAIR	TT	TS	TT-TS	TT/TS
1	19.4	19.6	-0.2	0.99
2	19.6	21.0	-1.4	0.93
3	21.5	19.8	1.7	1.09
4	23.5	24.0	-0.5	0.98
5	17.5	17.6	-0.1	0.99
6	17.2	18.7	-1.5	0.92

SUBJECT RP — CONSONANTS — RT VERSUS TS

PAIR	RT	TS	RT-TS	RT/TS
1	19.0	19.6	-0.6	0.97
2	20.3	21.0	-0.7	0.97
3	19.4	19.8	-0.4	0.98
4	18.6	17.6	1.0	1.06

TABLE II - continued

SUBJECT RP			CONSONANTS	QQ VERSUS Q
PAIR	QQ	Q	QQ-Q	QQ/Q
1	22.5	12.0	10.5	1.88
2	23.0	12.0	11.0	1.92
3	22.1	11.1	11.0	1.99
4	22.5	10.6	11.9	2.12
5	20.3	10.0	10.3	2.03
6	14.5	9.1	5.4	1.59
7	15.3	8.6	6.7	1.78
8	18.7	9.8	8.9	1.91
9	23.5	10.4	13.1	2.26
10	21.9	9.5	12.4	2.31
11	23.4	10.5	12.9	2.23
12	25.0	11.9	13.1	2.10
13	17.7	10.3	7.4	1.72
14	21.4	10.8	10.6	1.98
15	22.0	10.1	11.9	2.18
16	18.9	11.0	7.9	1.72
17	22.9	10.6	12.3	2.16
18	19.0	9.1	9.9	2.09

SUBJECT RP			CONSONANTS	NN VERSUS N
PAIR	NN	N	NN-N	NN/N
1	16.6	4.0	12.6	4.15
2	19.3	5.3	14.0	3.64
3	16.4	4.8	11.6	3.42
4	16.4	3.8	12.6	4.32
5	22.0	5.8	16.2	3.79
6	16.3	5.1	11.2	3.20
7	17.4	5.5	11.9	3.16
8	12.4	4.8	7.6	2.58
9	16.3	5.8	10.5	2.81
10	16.0	4.5	11.5	3.56

TABLE II - continued

SUBJECT RP — CONSONANTS RN VERSUS N

PAIR	RN	N	RN-N	RN/N
1	17.0	4.0	13.0	4.25
2	18.9	5.3	13.6	3.57
3	15.6	4.8	10.8	3.25
4	14.9	3.8	11.1	3.92
5	19.6	5.8	13.8	3.38
6	15.5	5.1	10.4	3.04
7	15.1	5.8	9.3	2.60
8	16.2	4.5	11.7	3.60

SUBJECT RP — CONSONANTS NN VERSUS RN

PAIR	NN	RN	NN-RN	NN/RN
1	16.6	17.0	-0.4	0.98
2	19.3	18.9	0.4	1.02
3	16.4	15.6	0.8	1.05
4	16.4	14.9	1.5	1.10
5	22.0	19.6	2.4	1.12
6	16.3	15.5	0.8	1.05
7	16.3	15.1	1.2	1.08
8	16.0	16.2	-0.2	0.99

TABLE III

Average measures for each word pair: absolute durations in cs, differences, and ratios. Subject: CCO.

SUBJECT CCO CONSONANTS PP VERSUS P

PAIR	PP	P	PP-P	PP/P
1	25.4	12.2	13.2	2.08
2	22.0	8.2	13.8	2.68
3	17.2	8.0	9.2	2.15
4	21.0	8.9	12.1	2.36
5	21.8	8.7	13.1	2.51
6	20.3	7.8	12.5	2.60
7	26.1	13.2	12.9	1.98
8	20.7	10.0	10.7	2.07
9	22.4	9.2	13.2	2.43
10	24.6	11.5	13.1	2.14
11	21.9	11.7	10.2	1.87
12	25.3	14.0	11.3	1.81
13	22.6	8.8	13.8	2.57
14	20.8	11.4	9.4	1.82
15	21.1	9.1	12.0	2.32
16	18.1	9.9	8.2	1.83
17	20.4	8.9	11.5	2.29

SUBJECT CCO CONSONANTS RP VERSUS P

PAIR	RP	P	RP-P	RP/P
1	18.6	10.8	7.8	1.72
2	18.8	8.9	9.9	2.11
3	26.1	13.2	12.9	1.98
4	19.9	10.0	9.9	1.99
5	16.8	8.4	8.4	2.00
6	19.8	9.6	10.2	2.06
7	22.3	8.5	13.8	2.62
8	22.7	12.9	9.8	1.76

TABLE III - continued

SUBJECT CCO CONSONANTS PP VERSUS RP

PAIR	PP	RP	PP-RP	PP/RP
1	20.4	20.0	0.4	1.02
2	24.7	24.7	0.0	1.00
3	23.2	21.8	1.4	1.06
4	22.5	22.2	0.3	1.01
5	23.0	22.4	0.6	1.03
6	21.6	22.1	-0.5	0.98
7	25.2	27.1	-1.9	0.93
8	21.6	21.5	0.1	1.00
9	21.1	21.2	-0.1	1.00
10	25.6	24.6	1.0	1.04
11	22.0	22.9	-0.9	0.96
12	26.1	26.1	0.0	1.00
13	17.7	16.7	1.0	1.06
14	20.7	19.9	0.8	1.04
15	20.0	20.4	-0.4	0.98
16	22.0	23.3	-1.3	0.94

SUBJECT CCO CONSONANTS TT VERSUS T

PAIR	TT	T	TT-T	TT/T
1	23.0	10.7	12.3	2.15
2	19.0	7.8	11.2	2.44
3	20.9	9.4	11.5	2.22
4	17.4	8.5	8.9	2.05
5	19.6	8.6	11.0	2.28
6	20.8	10.1	10.7	2.06
7	18.6	8.3	10.3	2.24
8	17.5	10.8	6.7	1.62
9	22.6	13.1	9.5	1.73
10	24.0	8.2	15.8	2.93
11	24.8	11.5	13.3	2.16
12	23.3	8.7	14.6	2.68

TABLE III - continued

SUBJECT CCO CONSONANTS RT VERSUS T

PAIR	RT	T	RT-T	RT/T
1	21.7	8.5	13.2	2.55
2	18.8	8.6	10.2	2.19
3	19.4	10.1	9.3	1.92
4	19.9	8.3	11.6	2.40
5	18.0	10.8	7.2	1.67
6	23.5	11.5	12.0	2.04
7	22.6	8.7	13.9	2.60
8	23.4	10.6	12.8	2.21

SUBJECT CCO CONSONANTS TS VERSUS T

PAIR	TS	T	TS-T	TS/T
1	23.4	14.5	8.9	1.61
2	17.8	8.6	9.2	2.07
3	20.5	10.1	10.4	2.03
4	21.3	8.3	13.0	2.57
5	24.4	8.9	15.5	2.74
6	22.2	9.5	12.7	2.34
7	21.2	10.6	10.6	2.00
8	22.7	10.5	12.2	2.16
9	23.6	13.1	10.5	1.80

SUBJECT CCO CONSONANTS TS VERSUS C

PAIR	TS	C	TS-C	TS/C
1	19.4	8.4	11.0	2.31
2	26.7	11.8	14.9	2.26
3	23.4	9.9	13.5	2.36
4	24.1	11.6	12.5	2.08
5	19.8	11.4	8.4	1.74
6	24.1	13.8	10.3	1.75
7	19.3	8.4	10.9	2.30

TABLE III - continued

SUBJECT CCO — CONSONANTS TT VERSUS RT

PAIR	TT	RT	TT-RT	TT/RT
1	21.8	22.6	-0.8	0.96
2	18.0	18.5	-0.5	0.97
3	19.6	18.8	0.8	1.04
4	20.8	19.4	1.4	1.07
5	18.6	19.9	-1.3	0.93
6	17.5	18.0	-0.5	0.97
7	24.8	23.5	1.3	1.06
8	23.3	22.6	0.7	1.03
9	19.0	19.7	-0.7	0.96
10	20.5	18.4	2.1	1.11

SUBJECT CCO — CONSONANTS TT VERSUS TS

PAIR	TT	TS	TT-TS	TT/TS
1	19.6	17.8	1.8	1.10
2	20.8	20.5	0.3	1.01
3	18.6	21.3	-2.7	0.87
4	22.6	23.6	-1.0	0.96
5	19.0	19.8	-0.8	0.96
6	20.6	20.5	0.1	1.00

SUBJECT CCO — CONSONANTS RT VERSUS TS

PAIR	RT	TS	RT-TS	RT/TS
1	18.8	17.8	1.0	1.06
2	19.4	20.5	-1.1	0.95
3	19.9	21.3	-1.4	0.93
4	19.7	19.8	-0.1	0.99

TABLE III - continued

SUBJECT CCO CONSONANTS QQ VERSUS Q

PAIR	QQ	Q	QQ-Q	QQ/Q
1	22.1	12.6	9.5	1.75
2	22.4	13.6	8.8	1.65
3	20.9	10.7	10.2	1.95
4	20.4	7.9	12.5	2.58
5	19.5	9.0	10.5	2.17
6	16.2	9.9	6.3	1.64
7	16.4	8.9	7.5	1.84
8	18.3	7.4	10.9	2.47
9	19.4	9.0	10.4	2.16
10	19.5	8.7	10.8	2.24
11	20.9	9.1	11.8	2.30
12	23.1	10.0	13.1	2.31
13	16.7	8.7	8.0	1.92
14	20.0	12.3	7.7	1.63
15	20.1	8.8	11.3	2.28
16	21.9	10.7	11.2	2.05
17	19.2	8.1	11.1	2.37
18	18.9	10.2	8.7	1.85

SUBJECT CCO CONSONANTS NN VERSUS N

PAIR	NN	N	NN-N	NN/N
1	15.4	7.3	8.1	2.11
2	19.6	6.5	13.1	3.02
3	14.7	7.3	7.4	2.01
4	14.2	5.2	9.0	2.73
5	20.2	6.4	13.8	3.16
6	15.4	6.8	8.6	2.26
7	16.5	4.8	11.7	3.44
8	16.4	6.7	9.7	2.45
9	16.5	4.8	11.7	3.44
10	16.2	6.4	9.8	2.53

TABLE III - continued

SUBJECT CCO CONSONANTS RN VERSUS N

PAIR	RN	N	RN-N	RN/N
1	16.0	7.3	8.7	2.19
2	18.6	6.5	12.1	2.86
3	15.2	7.3	7.9	2.08
4	14.8	5.2	9.6	2.85
5	17.8	6.4	11.4	2.78
6	15.6	6.8	8.8	2.29
7	16.6	4.8	11.8	3.46
8	15.6	6.7	8.9	2.33

SUBJECT CCO CONSONANTS NN VERSUS RN

PAIR	NN	RN	NN-RN	NN/RN
1	15.4	16.0	-0.6	0.96
2	19.6	18.6	1.0	1.05
3	14.7	15.2	-0.5	0.97
4	14.2	14.8	-0.6	0.96
5	20.2	17.8	2.4	1.13
6	15.4	15.6	-0.2	0.99
7	16.5	16.6	-0.1	0.99
8	16.4	15.6	0.8	1.05

TABLE IV

Average measures for each word pair: absolute durations in cs, differences, and ratios. Subject: IH.

SUBJECT IH — CONSONANTS PP VERSUS P

PAIR	PP	P	PP-P	PP/P
1	32.0	12.5	19.5	2.56
2	23.7	9.4	14.3	2.52
3	18.4	8.4	10.0	2.19
4	23.8	9.8	14.0	2.43
5	25.3	9.5	15.8	2.66
6	24.2	10.2	14.0	2.37

SUBJECT IH — CONSONANTS RP VERSUS P

PAIR	RP	P	RP-P	RP/P
1	19.9	9.8	10.1	2.03
2	19.2	8.6	10.6	2.23

SUBJECT IH — CONSONANTS PP VERSUS RP

PAIR	PP	RP	PP-RP	PP/RP
1	24.5	23.8	0.7	1.03
2	32.2	32.0	0.2	1.01
3	28.5	28.0	0.5	1.02
4	26.5	24.7	1.8	1.07
5	27.7	28.2	-0.5	0.98
6	22.9	25.1	-2.2	0.91
7	29.3	28.4	0.9	1.03
8	25.8	24.7	1.1	1.04
9	21.9	22.9	-1.0	0.96
10	28.6	26.6	2.0	1.08
11	23.4	23.3	0.1	1.00

TABLE IV - continued

SUBJECT IH CONSONANTS TT VERSUS T

PAIR	TT	T	TT-T	TT/T
1	30.7	9.5	21.2	3.23
2	22.3	8.7	13.6	2.56
3	24.9	8.9	16.0	2.80
4	21.4	9.4	12.0	2.28
5	19.2	11.5	7.7	1.67
6	24.9	11.9	13.0	2.09
7	23.7	7.7	16.0	3.08
8	23.8	8.3	15.5	2.87
9	23.9	7.3	16.6	3.27
10	17.7	9.0	8.7	1.97
11	24.9	11.5	13.4	2.17

SUBJECT IH CONSONANTS RT VERSUS T

PAIR	RT	T	RT-T	RT/T
1	23.9	9.3	14.6	2.57

SUBJECT IH CONSONANTS TS VERSUS T

PAIR	TS	T	TS-T	TS/T
1	22.1	7.7	14.4	2.87
2	23.0	8.3	14.7	2.77
3	24.7	7.3	17.4	3.38
4	25.4	11.5	13.9	2.21
5	26.0	13.7	12.3	1.90

SUBJECT IH CONSONANTS TS VERSUS C

PAIR	TS	C	TS-C	TS/C
1	21.1	10.2	10.9	2.07
2	23.4	7.3	16.1	3.21

TABLE IV - continued

SUBJECT IH — CONSONANTS — TT VERSUS RT

PAIR	TT	RT	TT-RT	TT/RT
1	28.6	29.9	-1.3	0.96
2	20.0	19.0	1.0	1.05
3	23.7	22.1	1.6	1.07
4	23.8	24.0	-0.2	0.99
5	23.9	22.2	1.7	1.08
6	17.7	18.5	-0.8	0.96

SUBJECT IH — CONSONANTS — TT VERSUS TS

PAIR	TT	TS	TT-TS	TT/TS
1	23.7	22.1	1.6	1.07
2	23.8	23.0	0.8	1.03
3	23.9	24.7	-0.8	0.97

SUBJECT IH — CONSONANTS — RT VERSUS TS

PAIR	RT	TS	RT-TS	RT/TS
1	22.1	22.1	0.0	1.00
2	24.0	23.0	1.0	1.04
3	22.2	24.7	-2.5	0.90

TABLE IV - continued

SUBJECT IH CONSONANTS QQ VERSUS Q

PAIR	QQ	Q	QQ-Q	QQ/Q
1	25.4	10.6	14.8	2.40
2	23.8	12.9	10.9	1.84
3	24.1	10.7	13.4	2.25
4	23.2	9.9	13.3	2.34
5	23.5	8.9	14.6	2.64
6	18.4	9.5	8.9	1.94
7	17.0	8.7	8.3	1.95
8	18.6	9.0	9.6	2.07
9	21.8	7.9	13.9	2.76
10	25.5	9.6	15.9	2.66
11	22.3	9.4	12.9	2.37
12	26.4	10.6	15.8	2.49
13	21.4	8.3	13.1	2.58
14	21.3	11.4	9.9	1.87
15	23.6	9.4	14.2	2.51
16	24.2	10.8	13.4	2.24
17	24.9	8.8	16.1	2.83
18	22.5	8.3	14.2	2.71

SUBJECT IH CONSONANTS NN VERSUS N

PAIR	NN	N	NN-N	NN/N
1	17.3	6.3	11.0	2.75
2	23.0	7.9	15.1	2.91
3	18.4	4.5	13.9	4.09
4	16.3	4.1	12.2	3.98
5	24.5	5.9	18.6	4.15
6	16.4	6.1	10.3	2.69

TABLE IV - continued

SUBJECT IH			CONSONANTS	RN VERSUS N
PAIR	RN	N	RN-N	RN/N
1	17.0	6.3	10.7	2.70
2	21.2	7.9	13.3	2.68
3	17.0	4.5	12.5	3.78
4	14.2	4.1	10.1	3.46
5	19.8	5.9	13.9	3.36
6	15.0	6.1	8.9	2.46

SUBJECT IH			CONSONANTS	NN VERSUS RN
PAIR	NN	RN	NN-RN	NN/RN
1	17.3	17.0	0.3	1.02
2	23.0	21.2	1.8	1.08
3	18.4	17.0	1.4	1.08
4	16.3	14.2	2.1	1.15
5	24.5	19.8	4.7	1.24
6	16.4	15.0	1.4	1.09

3.5.1. Geminate versus single consonant

As stated earlier, every measure of duration in the tables represents an average of generally five individual measurements. A statistical treatment[16] of the original raw data shows that there is a good deal of dispersion of the measures for each item (word). However, the durational difference between phonologically geminate (long) and phonologically single (short) consonant is found to be significant at a 99 per cent confidence level within each (sub-) minimal word pair.

There is absolutely no overlapping between long and short consonant within a word pair, and it was true for all three subjects that all occurrences of the long item were at least 1.7 times as long as any occurrence of the short (in the same word pair), with the following exceptions: nine words had a ratio between 1.6 and 1.7, and two words had a ratio of 1.56 and 1.59 respectively.

The quantity contrast is thus very clearly manifested. This is true both with respect to the absolute difference and with respect to the ratio long:short. The former exceeds 9 cs, and the latter generally exceeds 2:1. Generally, the lowest figures are found with /qq/ versus /q/, for which some word pairs differ by only some 6 cs and show a ratio of little more than 1.6:1. The opposite extreme occurs with /nn/ versus /n/ whose ratio sometimes exceeds 4:1.

The relatively low figures for the quantity contrast of /qq/: /q/ may have to do with the fact that /qq/ : /q/ to some extent differ also by weaker articulation of the latter. They may thus not be quite comparable. The other consonants studied are supposed to exhibit a more purely

---

16) We are indebted to Mr Jørgen Elgaard Knudsen, mag. scient., who has performed the data processing.

durational difference.

As for the nasal the high ratio is not due to excessive duration of /nn/ but to a tendency toward very short duration of /n/. It may be inherently shorter than the stops, but the difference may also in part be due to our principles of segmentation (with exclusion of the entire transitions in order to ensure a meaningful comparison of /nn/ with /rn/). Anyway, the difference between /nn/ and /n/ (which, of course, is not nearly as vulnerable to differences in the duration of /n/ as the ratio is) is similar to that of /tt/ to /t/, etc., viz. typically of the order of lo cs.

A comparison of the two measures of quantity relationship: absolute difference and ratio, shows that within the word list belonging to each phonological contrast there is more variation in absolute difference than in ratio. If, however, we compare the averages for each word list with the averages for other word lists (cf. examples in Table V) the relative variation of the difference is clearly smaller than the relative variation of the ratio. Since each list represents a variety of different word types (differing in syllable and mora number, in the position of the measured consonants within the word, etc.), it seems that environments exert less influence on the ratio than on the absolute difference in duration, whereas the phonetic properties of the segments themselves exert less influence on the absolute difference of duration than on the ratio.

We suggested in section 2. above that geminates differ from single consonants not only by being "double" but also by belonging always to two syllables, as against single, intervocalic consonants which are syllable initial. This applies equally to double consonants that are related to single consonants by the gemination rule (see section 2.4.) and to double consonants that are derived from consonant clusters with complete assimilation (see section 2.3.). Hence, what we really compare is not, systematically speaking, a long and a short consonant in the same position (note that word

TABLE V Specimens of statistical treatment of ratios and differences.

SUBJECT RP	CONSONANTS PP VERSUS P			NUMBER=18
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
PP/P	2.21	0.23	0.05	0.11
PP-P	11.19	2.14	0.51	1.07

SUBJECT CCO	CONSONANTS PP VERSUS P			NUMBER=17
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
PP/P	2.21	0.29	0.07	0.15
PP-P	11.78	1.71	0.41	0.88

SUBJECT RP	CONSONANTS PP VERSUS RP			NUMBER=16
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
PP/RP	0.99	0.06	0.01	0.03
PP-RP	-0.13	1.24	0.31	0.66

SUBJECT CCO	CONSONANTS PP VERSUS RP			NUMBER=16
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
PP/RP	1.00	0.04	0.01	0.02
PP-RP	0.03	0.88	0.22	0.47

SUBJECT IH	CONSONANTS PP VERSUS RP			NUMBER=11
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
PP/RP	1.01	0.05	0.01	0.03
PP-RP	0.33	1.22	0.37	0.82

TABLE V - continued

SUBJECT RP	CONSONANTS TT VERSUS T			NUMBER=14
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
TT/T	2.27	0.36	0.10	0.21
TT-T	11.06	1.95	0.52	1.13

SUBJECT CCO	CONSONANTS TT VERSUS T			NUMBER=12
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
TT/T	2.21	0.36	0.10	0.23
TT-T	11.32	2.48	0.72	1.58

SUBJECT IH	CONSONANTS TT VERSUS T			NUMBER=11
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
TT/T	2.54	0.55	0.16	0.37
TT-T	13.97	3.77	1.14	2.53

SUBJECT RP	CONSONANTS TT VERSUS RT			NUMBER=10
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
TT/RT	0.99	0.06	0.02	0.04
TT-RT	-0.24	1.09	0.35	0.78

SUBJECT CCO	CONSONANTS TT VERSUS RT			NUMBER=10
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
TT/RT	1.01	0.06	0.02	0.04
TT-RT	0.25	1.15	0.36	0.82

TABLE V - continued

SUBJECT RP	CONSONANTS QQ VERSUS Q			NUMBER=18
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
QQ/Q	2.00	0.20	0.05	0.10
QQ-Q	10.40	2.30	0.54	1.15

SUBJECT CCO	CONSONANTS QQ VERSUS Q			NUMBER=18
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
QQ/Q	2.06	0.30	0.07	0.15
QQ-Q	10.02	1.85	0.44	0.92

SUBJECT IH	CONSONANTS QQ VERSUS Q			NUMBER=18
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
QQ/Q	2.36	0.32	0.08	0.16
QQ-Q	12.96	2.43	0.57	1.21

SUBJECT RP	CONSONANTS NN VERSUS N			NUMBER=10
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
NN/N	3.46	0.55	0.17	0.39
NN-N	11.97	2.24	0.71	1.60

SUBJECT CCO	CONSONANTS NN VERSUS N			NUMBER=10
	MEAN	S.D.	M.E.	95 PCT.CF.LM.
NN/N	2.71	0.53	0.17	0.38
NN-N	10.29	2.17	0.69	1.55

final clusters and geminates are impossible), but a syllable final consonant versus zero before an identical, syllable initial consonant:

VC-CV

V-CV

although on the articulatory level we get a longer versus a shorter closure hold. If we denote the syllable final, "mora-forming", consonant by C_a, and the syllable initial consonant by C_b, we may suggest that (i) the addition of C_a to the sequence gives an increment of duration which is not highly dependent upon the specific features of C_a, whereas (ii) the duration of C_b is rather dependent upon its specific features. (- How properties of the environments influence C_aC_b and C_a remains to be studied in detail.)

Geminates tend to be more than twice as long as single segments. This is not true of all word pairs taken individually, but it holds true for the averages of all results within each type of contrast except /qq/:/q/, although the average ratios do not exceed 2:1 very much. The results are:

/pp/:/p/	RP 2.21:1	CCO 2.21:1	IH 2.46:1
/tt/:/t/	RP 2.27:1	CCO 2.21:1	IH 2.54:1
/qq/:/q/	RP 2:1	CCO 2.36:1	IH 2.36:1
/nn/:/n/	RP 3.46:1	CCO 3.43:1	IH 3.43:1

These figures have limited validity since they are averages from arbitrary sets of word pairs, but they do show a tendency.

3.5.2. rC-clusters

If we now turn to the sequences representing underlying uvular plus consonant, the figures in the tables show conclusively that the stop or nasal segments of these sequences are equal in duration to the geminates treated above but entirely different from single segments between vowels.

I.e., it can be stated that sequences containing an underlying uvular plus a stop or nasal are reflected phonetically by sequences containing a long stop or nasal, which according to its duration might be the reflex of an underlying geminate or consonant cluster.

3.5.3. Coronal affricate

In all types discussed above the explosion of stops was included in the (total) duration of the consonant. We have calculated the averages for closure and explosion phase separately before adding them together, but we have not given the figures here because we did not find that they contributed significantly to this discussion. The explosion is generally short (and rather weak).

As stated earlier, t before i is reflected by the affricate /¢/, and tt in this position coalesces with t(+)s into /t¢/. Our measurements clearly show that the affrication phase is of essentially the same duration in /t¢/ and and /¢/, viz. about 5 cs in both (slightly less for IH than for the other persons), as against the short explosions of 2 cs or less found in other consonants. At the same time the occlusive portion of /t¢/ is considerably longer than a normal segment, so that the total duration of /t¢/ closely resembles that of /tt/. Thus the transcriptions /¢/ and /t¢/ seem adequate from a phonetic point of view: the former is an affricate of which roughly the last 5 cs are taken up by affrication, the latter is equal to the same affricate preceded by the homorganic stop /t/.

3.5.4. Duration of the whole word

In the preceding sections we have been concerned only with the durations of the specific segments under consideration. In the case of quantity contrasts (not involving a uvular component) one might, however, expect some influence on the remainder of the words. We have measured the total duration (minus initial or final consonant) and the durations of the single segments of some chosen word pairs, see Table VI. This showed no clear pattern, except for the durational difference of the contrasting segments.

Sequences involving an underlying uvular consonant pose a more important problem. We have not found any reason to assume that the long oral closure hold (whose existence can be inferred from the mingograms) is a succession of two different consonants sharing a manner of articulation but differing with respect to point of articulation (i.e. uvular stop or nasal plus non-uvular stop or nasal). However, one might a priori assume that there should be a clearly distinguishable "r"-segment of a certain duration before the stop or nasal, since this is the way phonetically trained scholars have almost invariably transcribed the sequences. We have measured the total duration (minus initial or final consonant) and the durations of the individual segments of some chosen words with and without underlying r̲ before consonant, see Table VII. The measurements show that vowels affected by an underlying uvular are either as long as or slightly longer than their non-uvularized counterparts. This tendency is, however, very weak indeed, and we cannot decide at present whether it is a matter of inherent extra duration of more open types of vowels (the uvularized vowels require a more pronounced displacement of the tongue than the other vowels, which might account for their tendency toward longer duration). At any rate, the total duration of the vocalic segment (including the transition to stop or nasal) cannot be said to bear evidence that it represents a succession of two segments: vowel and r̲. If one wishes to derive it that way phonologically, the underlying r̲ segment must be said to be incorporated phonetically into the vowel. It may instead be meaningful to speak of pre-uvularization as a secondary feature of the otherwise assimilated consonant.

TABLE VI

Durational relationships of some chosen words.
Averages of 5 measurements.

	i	p	a	(k)		
	i	pp	a	(k)		
CCO	4.7	12.2	5.6		=	22.5
	4.7	25.4	5.6		=	35.7
RP	5.6	1o.9	6.1		=	22.6
	4.1	26.5	5.4		=	36.o
IH	6.7	12.5	6.3		=	25.5
	4.7	32.o	6.1		=	42.8

	u	p	i	pp	u	(q)		
	u	pp	i	pp	u	(q)		
CCO	3.9	8.7	5.1	24.1	5.o		=	46.8
	4.6	21.8	6.o	21.5	5.6		=	59.5

	(t)	u	t	u	pp	uŋa		
	(t)	u	tt	u	pp	uŋa		
CCO		4.5	7.8	4.8	21.2	2o.4	=	58.7
		4.5	19.o	4.6	17.5	16.9	=	62.5
RP		4.9	9.o	4.8	19.8	17.9	=	56.3
		4.8	19.9	4.9	2o.5	17.2	=	67.2
IH		6.2	8.7	4.9	25.2	2o.9	=	65.9
		5.3	22.3	5.3	19.3	18.4	=	7o.6

TABLE VII

Durational relationships of some chosen words.
Averages of 5 measurements.

	u	pp	i	(k)		
	u	rp	i	(k)		
CCO	5.4	24.7	5.5		=	35.6
	6.6	24.7	5.6		=	36.9
RP	5.2	26.5	5.2		=	36.8
	5.4	25.5	4.9		=	35.9
IH	5.1	32.2	6.o		=	43.3
	6.5	32.o	6.o		=	44.5

	u	pp	i	ŋŋua	(q)		
	u	rp	i	ŋŋua	(q)		
CCO	4.9	21.6	5.2	3o.9		=	62.6
	5.3	21.5	5.5	3o.7		=	63.o

	i	n	u	tt	u	(t)		
	i	n	u	rt	u	(t)		
CCO	6.3	4.5	6.9	21.8	6.o		=	45.5
	5.5	5.1	7.2	22.6	5.9		=	46.3
RP	5.9	5.3	6.o	22.o	4.5		=	43.5
	6.1	3.8	7.5	22.4	5.2		=	44.9
IH	6.2	6.o	7.1	28.6	5.3		=	53.2
	6.6	5.9	7.1	29.9	5.2		=	54.7

	i	q	u	rp	u	(q)		
	i	qq	u	rp	u	(q)		
CCO	4.4	8.8	6.7	21.4	5.7		=	47.o
	5.2	2o.1	6.2	21.5	5.4		=	58.4
RP	5.4	1o.1	5.1	21.4	4.4		=	46.4
	6.4	22.o	6.4	22.o	5.4		=	62.1
IH	7.o	9.4	5.7	27.8	5.8		=	55.7
	5.7	23.6	6.2	22.3	5.8		=	63.6

	i	n	i	q	a	rp	u	(q)		
	i	nn	i	q	a	rp	u	(q)		
	i	rn	i	q	a	rp	u	(q)		
CCO	4.8	6.8	5.2	9.4	4.2	21.2	4.9		=	56.5
	5.2	15.8	5.5	9.4	4.2	2o.o	4.9		=	65.o
	6.8	15.6	4.2	9.3	4.5	19.6	5.6		=	65.6
RP	5.5	5.1	5.9	9.1	5.o	21.5	4.1		=	56.4
	6.4	16.3	5.5	9.o	5.3	22.6	4.4		=	69.4
	6.9	15.5	4.4	9.o	5.3	21.4	4.o		=	66.4
IH	5.3	6.1	7.8	8.4	5.2	23.2	5.o		=	61.o
	6.2	16.4	7.3	8.4	5.2	22.6	5.7		=	71.8
	7.9	15.o	5.9	9.1	5.4	22.3	6.o		=	71.5

3.6. Statistical treatment

Table V presents some specimens of a statistical treatment presenting overall ratios and differences. This approach is obviously of extremely limited validity since it involves a number of unwarranted (or contrary-to-fact) assumptions, such as (i) that for a given contrast there is a "true" overall figure for the ratio or difference as such, (ii) that this can be found from a "representative" sample of occurrences of the contrast, and (iii) that our sets of word pairs are such "representative" samples. This they are not very likely to be, particularly since the very concept of "representative" is anything but well-defined in this context. (For this reason we have left in a single repetition in RP's set of 18 word pairs with /pp/ versus /p/, although it is strictly speaking statistically dubious.)

The following figures are given for ratios as well as differences: (1) mean value of the total set (sample), (2) standard deviation, (3) standard error of the mean, and (4) maximum deviation of the sample mean from a hypothetical "true" mean estimated with a probability of 95 per cent.

Several of the sets which are not shown in Table V were too small for this kind of processing. However, they all show largely the same tendencies. Some of the information was given in section 3.5.1. above.

4. Conclusion

As shown in sections 3.5.1., 3.5.2., and 3.5.3. the phonetic measurements corroborate the assumptions about phonological rules made in the beginning of this paper. Firstly it is confirmed to be entirely adequate to describe geminates as clusters of two like consonants. Secondly clusters with an underlying uvular as the first segment behave phonetically as if the first segment is assimilated to the second as in other instances of assimilation, except that the feature of uvularity (uvularization) is retained somewhere in the sequence. Finally, the complex arising from t(+)t before i or from t(+)s can be described as a succession of stop and homorganic fricative, i.e. when derived from t(+)t it exhibits affrication before i, and when derived from t(+)s it exhibits partial assimilation of s to t.

References

Bergsland, Knut 1955: A Grammatical Outline of the Eskimo Language of West Greenland. Mimeographed (Oslo).

Chomsky, Noam and Morris Halle 1968: The Sound Pattern of English.

Egede, Paul 1760: Grammatica Grönlandica Danico-Latina. (In IDC microfiche editions.)

Hill, A.A. 1958: Introduction to Linguistic Structures, Appendix A: Eskimo (p. 419-440).

Holtved, Erik 1964: Kleinschmidts Briefe an Theodor Bourquin (= Meddelelser om Grønland 140,3).

Kiparsky, Paul 1968: How Abstract is Phonology? Mimeographed.

Kleinschmidt, Samuel 1851: Grammatik der grönländischen Sprache.

McCawley, James 1968: The Phonological Component of a Grammer of Japanese.

Petersen, Robert 1970: "On Phonological Length in the Eastern Eskimo Dialects", Folk 11-12, p. 329-344.

Slis, I. H. 1968: "Experiments on Consonant Duration Related to the Time Structure of Isolated Words", <u>IPO Annual Progress Report</u> 3, p. 71-8o.

Swadesh, Morris 1946: "South Greenlandic (Eskimo)", in: <u>Linguistic Structures of Native America</u>, <u>Viking Fund Publications in Anthropology and Linguistics</u>, p. 3o-54.

Ulving, Thor 1953: "Consonant Gradation in Eskimo", <u>IJAL</u> 19, p. 45-52.

Underhill, Robert 197o: "Greenlandic Morphophonemics", Paper read before the Conference on Eskimo Linguistics June 5-7 (available in the form of a tape recording kindly put at our disposal by Mr Robert Petersen).